



Área Académica de Administración de Tecnologías de Información

Propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax

Trabajo Final de Graduación para optar al grado de Licenciatura en Administración de Tecnología de Información

Elaborado por: Carlos H. Morales Porras

Prof. Tutor: Máster. Agustín Francesa Alfaro

Cartago, Costa Rica

II Semestre

Noviembre, 2022



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional](http://creativecommons.org/licenses/by-nc-nd/4.0). Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-nc-nd/4.0>

Hoja de Aprobación

AREA ACADEMICA DE ADMINISTRACIÓN DE TECNOLOGIA DE INFORMACIÓN

GRADO ACADEMICO: LICENCIATURA

Los miembros del Tribunal Examinador de área de Administración de Tecnologías de Información, recomendamos que el presente Informe Final del Proyecto de Graduación del estudiante Carlos Humberto Morales Porras sea aceptado como requisito parcial para obtener el grado académico de Licenciatura en Administración de Tecnología de información.

Agustín Francesa Alfaro
Profesor Tutor

Fauricio Conejo Navarro
Lector de Industria

Néstor Morales Rodríguez
Lector Académico

Yarima Sandoval
Coordinadora de Trabajo
Final de Graduación

Dedicatoria

A mi familia les dedico este trabajo ya que han sido las personas que me han apoyado en todos los ámbitos para llegar hasta este punto. Dándome siempre su amor, cariño y guía para llegar a ser la persona que soy hoy en día y siempre orientarme al desarrollo personal y profesional.

Agradecimientos

A todo el equipo de Servicios de Datos, especialmente a Luis, que me brindo la oportunidad de desarrollar el proyecto dentro del equipo. A Diego por siempre estar apoyándome en el proceso de ejecución del proyecto brindando comentarios, conocimiento y orientación de procesos del negocio.

Agustín, que fue una parte muy importante para la ejecución del proyecto. Brindando apoyo y guía para la elaboración del proyecto.

A mis compañeros dentro de la organización Ricardo y Carlos, los cuales fueron excelente compañía durante el TFG y estuvieron siempre dispuestos a colaborar y aclarar dudas. También a todos mis compañeros y profesores con quienes tuve la oportunidad de compartir y crecer a lo largo de la carrera para ser mejor persona y profesional.

Resumen

Esta investigación tiene como propósito proponer una solución para automatizar el proceso de elaboración de reportes de Business Care y Boarding en el departamento de Servicios de Datos.

El procedimiento metodológico se llevó a cabo utilizando un enfoque cualitativo, en el cual se realizaron entrevistas y revisiones literarias a fuentes primarias de información, relacionadas a la problemática. Por medio de estos instrumentos, se identificaron las actividades realizadas en este proceso, su situación actual y las opciones de mejora existentes, así como las herramientas utilizadas previamente y disponibles en la organización. Además, se emplearon pruebas de concepto para validar el nuevo modelo propuesto para la ejecución del proceso, con base en la información disponible y en los requerimientos recolectados. La realización de este procedimiento se ejecuta bajo una perspectiva de rediseño de procesos, debido a que se busca llegar al proceso mejorado, para así disminuir el tiempo y la frustración de los encargados. El marco metodológico que se toma como base para este rediseño fue el propuesto por Madison (2005). Asimismo, se revisaron las buenas prácticas existentes en el mercado, para utilizarlas como insumo dentro del proceso de rediseño.

Como resultado de la investigación, se concluyó que existen varios métodos para extraer datos, sin la necesidad de utilizar reportes precreados. Gracias a las características que presenta Python, se puede realizar por medio de las distintas librerías. La organización cuenta con conexiones directas a las fuentes de datos por medios de API, o bien, se ha migrado mucha información a su *data lake*. Así mismo, se determinó que no existía documentación o proceso estandarizado en la empresa, por lo cual, este documento es uno de los principales entregables a esta.

Palabras clave: Procesos, ETL, API, Base de datos, automatización, reportes.

Abstract

The purpose of this research is to propose a solution for the automation of the "Business Care" and "Boarding" reporting process in the data services department.

The methodological procedure carried out under a qualitative approach, in which interviews and literary reviews were conducted on primary sources of information related to the problem. Through these instruments, the activities inside the process, the current situation and the existing improvement options were identified, as well as the tools previously used and those available in the organization. In addition, proofs of concept were used to validate the new model proposed for the execution of the process, based on the information available and the requirements collected.

The methodological procedure is executed under a process redesign perspective, since it seeks to reach the improved process, reducing the time and frustration of those in charge. The methodological framework which is taken as the foundations for this redesign was the proposed by Madison (2005). In addition, a review of existing good practices in the market is carried out to be used as input in the redesign process.

As a result of the investigation, it was concluded that there are several methods for extracting information, without the need to use pre-created reports. Thanks to the characteristics that Python presents, data extraction can be conducted through the different libraries. Also, the organization has direct connections to data sources through APIs or a lot of information has been migrated to the organization's data lake. Likewise, it was determined that the non-existence of some documentation or standardized process, which led to the realization of this document as one of the main deliverables to the organization.

Keywords: Processes, ETL, API, Database, automation, reports.

Nota Aclaratoria

Género¹:

La actual tendencia al desdoblamiento indiscriminado del sustantivo en su forma masculina y femenina va contra el principio de economía del lenguaje y se funda en razones extralingüísticas. Por tanto, deben evitarse estas repeticiones, que generan dificultades sintácticas y de concordancia, que complican innecesariamente la redacción y lectura de los textos.

Este documento se redacta de acuerdo con las disposiciones actuales de la Real Academia Española con relación al uso del “género inclusivo”. Al mismo tiempo se aclara que estamos a favor de la igualdad de derechos entre los géneros.

¹ Recuperado de: <http://www.rae.es/consultas/los-ciudadanos-y-las-ciudadanas-los-ninos-y-las-ninas>

Tabla de contenidos

Capítulo 1.	Introducción	1
1.1.	Descripción general	2
1.2.	Antecedentes.....	3
1.2.1.	Descripción de la organización	3
1.2.1.1.	Misión	5
1.2.1.2.	Visión	5
1.2.1.3.	Valores.....	5
1.2.1.4.	Equipo de trabajo	6
1.2.2.	Trabajos similares realizados dentro y fuera de la organización.....	7
1.3.	Planteamiento del problema	9
1.3.1.	Situación problemática	9
1.3.2.	Justificación del proyecto	12
1.3.3.	Beneficios esperados o aportes del Trabajo Final de Graduación	13
1.3.3.1.	Beneficios directos	13
1.3.3.2.	Beneficios indirectos.....	13
1.4.	Objetivos del Trabajo Final de Graduación	13
1.4.1.	Objetivo general.....	13
1.4.2.	Objetivos específicos.....	13
1.5.	Alcance.....	14
1.6.	Supuestos	15
1.7.	Entregables.....	15
1.8.	Exclusiones del proyecto	16
1.9.	Limitaciones	16
Capítulo 2.	Marco conceptual	17
2.1.	Inteligencia de negocios	19
2.1.1.	Toma de decisiones.....	19
2.1.2.	Datos	19
2.1.3.	Información.....	19
2.1.4.	Bases de datos	19

2.1.5.	BigQuery	20
2.1.6.	Indicadores de desempeño (KPI)	20
2.1.7.	Integración de datos	21
2.1.8.	Herramientas y reportes	22
2.1.9.	Software	22
2.1.10.	Visualización de datos	22
2.1.11.	Interfaces de programación de aplicaciones (API)	22
2.2.	Gestión de procesos de negocio (BPM)	23
2.2.1.	Marcos metodológicos	23
2.2.1.1.	Metodología de BPM según Marlon Dumas	23
2.2.1.2.	Metodología de BPM según Dan Madison	24
2.2.2.	Notación BPMN	25
2.3.	Técnicas y análisis	26
2.3.1.	Análisis de costo-beneficio	26
2.3.2.	Prueba de concepto	27
2.3.3.	El principio de Pareto	27
2.4.	Plan de implementación	28
2.4.1.	PMBOK	28
Capítulo 3.	Marco metodológico	29
3.1.	Tipo de investigación	30
3.2.	Alcance de la investigación	31
3.3.	Diseño de la investigación	32
3.4.	Fuentes de datos e información	33
3.4.1.	Fuentes primarias	33
3.4.2.	Fuentes secundarias	34
3.5.	Sujetos de investigación	35
3.6.	Variables o categorías de la investigación	36
3.7.	Técnicas e instrumentos de recolección de datos	37
3.8.	Matriz de cobertura de las variables	39
3.9.	Procedimiento metodológico de la investigación	40
3.9.1.	Fase 1. Descubrimiento del proceso actual	41

3.9.1.1.	Identificación del estado actual.....	41
3.9.1.2.	Entrevistar a los involucrados.....	42
3.9.2.	Fase 2. Diseño del nuevo proceso.....	42
3.9.2.1.	Evaluación de mejores prácticas.....	42
3.9.2.2.	Identificación y documentación de las herramientas para los procesos de extracción, carga y transformación de datos	43
3.9.2.3.	Realizar primer rediseño	43
3.9.2.4.	Validación del nuevo diseño con involucrados y pruebas de concepto	43
3.9.2.5.	Realizar versión final del nuevo proceso.....	43
3.9.3.	Fase 3. Plan de implementación.....	43
3.9.4.	Fase 4. Análisis de costo-beneficio.....	45
3.10.	Operacionalización de las variables o categorías.....	46
3.11.	Tabla resumen del procedimiento metodológico o trazabilidad	48
Capítulo 4.	Análisis de resultados.....	50
4.1.	Primera fase: Descubrimiento del proceso actual.....	51
4.1.1.	Identificación del estado actual.....	51
4.1.2.	Identificación de necesidades del equipo y requerimientos del cliente	58
4.2.	Segunda fase: Diseño del nuevo proceso.....	61
4.2.1.	Evaluación de mejores prácticas.....	61
4.2.2.	Identificación y documentación de las herramientas para los procesos de extracción, carga y transformación de datos.....	63
4.2.3.	Validación del diseño y pruebas de concepto.....	64
4.3.	Tercera fase: Plan de implementación.....	69
4.4.	Cuarta fase: Análisis de costo beneficio.....	73
Capítulo 5.	Propuesta de solución.....	74
5.1.	Fase 2: Diseño del nuevo proceso.	75
5.1.1.	Realización del nuevo proceso.....	75
5.1.2.	Documentación del nuevo proceso.....	79
5.2.	Fase 3: Plan de implementación	81
5.3.	Fase 4: Análisis de costo-beneficio.....	82
5.3.1.	Recursos	82
5.3.2.	Licencias.....	83

5.3.3. Tiempo de ejecución	84
5.3.4. Costos totales	84
5.3.5. Beneficios	86
5.3.6. Resultados del análisis	87
Capítulo 6. Conclusiones	88
6.1. Conclusiones: objetivo específico 1	89
6.2. Conclusiones: objetivo específico 2	90
6.3. Conclusiones: objetivo específico 3	91
Capítulo 7. Recomendaciones	92
Capítulo 8. Referencias bibliográficas	94
Capítulo 9. Apéndices	97
9.1. Apéndice A. Plantilla de minutas	98
9.2. Apéndice B. Plantilla para control de cambios	98
9.3. Apéndice C. Plantilla de recopilación de datos en revisión documental	98
9.4. Apéndice D. Plantilla de recopilación de datos en la observación	99
9.5. Apéndice E. Plantilla de entrevista	99
9.6. Apéndice F. Plantilla de requerimientos	99
9.7. Apéndice G. Plantilla de la matriz <i>Two-by-Two</i>	100
9.8. Apéndice H. Plantilla de hoja de tiempos	100
9.9. Apéndice I. Resultados de la observación–Extracción de datos del reporte mensual 101	
9.10. Apéndice J. Resultados de la observación – Extracción de datos del reporte diario y semanal	102
9.11. Apéndice K. Resultados de la observación – Transformación de datos del reporte diario y semanal	102
9.12. Apéndice L. Resultados de la observación – Unión de datos para la creación de reportes 103	
9.13. Apéndice M. Resultados de la observación – Realización de cálculos y emparejar datos para la creación de reportes	104
9.14. Apéndice N. Resultados de la observación – Actualizar fuentes de datos de los reportes	105
9.15. Apéndice O. Resultados de entrevista #1 - Descubrimiento de la organización ... 106	
9.16. Apéndice P. Resultados de entrevista #2 - Descubrimiento del proceso	107

9.17. Apéndice Q. Resultados de entrevista #3 – Uso y requerimientos del proceso para el cliente.....	108
9.18. Apéndice R. Resultados de entrevista #4 – Limitaciones y requerimientos tecnológicos del departamento.....	108
9.19. Apéndice S. Resultados de Matriz <i>Two-by-Two</i>.....	110
9.20. Apéndice T. Resultados de la hoja de tiempos del proceso.	110
9.21. Apéndice U. Resultado de análisis revisión documental #1.....	111
9.22. Apéndice V. Resultado de análisis revisión documental #2.....	111
9.23. Apéndice W. Resultado de análisis revisión documental #3.....	112
9.24. Apéndice X. Resultado de análisis revisión documental #4.....	112
9.25. Apéndice Y. Resultado de análisis revisión documental #5.....	113
9.26. Apéndice Z. Resultado de análisis revisión documental #6.....	114
9.27. Apéndice AA. Diagrama de Pareto.....	115
9.28. Apéndice BB. Retorno de Inversión	115
9.29. Apéndice CC. Documentación del proceso	116
9.30. Apéndice DD. Plan de implementación	140
9.31. Apéndice EE. Minutas	147
9.31.1. Minuta 1. Presentación formal estudiantes y tutor	147
9.31.2. Minuta 2. Establecer roles, alcance y expectativas del proyecto.....	148
9.31.3. Minuta 3. Realizar ajustes del proyecto en términos del alcance y objetivos del proyecto. 149	
9.31.4. Minuta 4. Revisión del capítulo 1	151
9.31.5. Minuta 5. Revisión de cambios realizados en el capítulo 1 e idealizar el proceso metodológico que se utilizara.....	152
9.31.6. Minuta 6. Aclarar dudas con respecto al marco metodológico y sus partes.	153
9.31.7. Minuta 7. Revisión del capítulo 3.....	153
9.31.8. Minuta 8. 2da revisión del proyecto de graduación	153
9.31.9. Minuta 9. Recolección de requerimientos para estándar	155
9.31.10. Minuta 10. Validación del modelo “as-is”	156
9.31.11. Minuta 11. Realizar propuesta de estandarización.....	157
9.31.12. Minuta 12. Validar propuesta de estandarización	158

9.31.13. Minuta 13. Validación de pruebas de concepto y consultas sobre proyectos similares	159
9.31.14. Minuta 14. Atender dudas capítulos 4 y 5	160
9.31.15. Minuta 15. Revisión de propuesta de solución	161
9.31.16. Minuta 16. Validación de beneficios de la propuesta	162
9.31.17. Minuta 17. 3era revisión del proyecto de graduación	163

Índice de Figuras

Figura 1. Organigrama del departamento de Servicios de Datos	4
Figura 2. Organigrama del equipo de trabajo	7
Figura 3. Árbol del problema.....	11
Figura 4. Mapa de conceptos	18
Figura 5. <i>Metodología de rediseño de procesos según Dan Madison</i>	24
Figura 6. <i>Tipos de entrevistas</i>	39
Figura 7. Fases del proceso metodológico.....	40
Figura 8. Metodología de rediseño según Madison.....	41
Figura 9. Fases del PMBOK utilizadas.....	44
Figura 10. Proceso As-is de creación de reportes.....	56
Figura 11. Modelo del proceso To-be.....	77
Figura 12. Diagrama de flujo del sistema dentro proceso de elaboración de reportes	78

Índice de Tablas

Tabla 1. Proyectos internos.....	7
Tabla 2. Proyectos externos	8
Tabla 3. Cambios en las herramientas	9
Tabla 4. Descripción del proceso de ETL.....	21
Tabla 5. Elementos de la notación BPMN.....	25
Tabla 6. Tipos de Alcance.....	31
Tabla 7. Tipos de diseños para la investigación cualitativa.....	32
Tabla 8. Fuentes primarias	33
Tabla 9. Fuentes secundarias	34
Tabla 10. Sujetos de investigación: rol y relevancia.....	35
Tabla 11. Variables de la investigación identificadas.....	36
Tabla 12. Técnicas de recolección de datos	37
Tabla 13. Matriz de cobertura de variables.....	39
Tabla 14. Operacionalización de las variables.....	47
Tabla 15. Matriz de trazabilidad	49
Tabla 16. Participantes del proceso.....	52
Tabla 17. Actividades del proceso	52
Tabla 18. Fuentes de datos utilizadas dentro del proceso	53
Tabla 19. Herramientas implementadas dentro del proceso actual.....	53
Tabla 20. Reportes generados para el departamento Business Care y Boarding.....	54
Tabla 21. Ideas de mejoras-impacto vs. costo	58
Tabla 22. Tiempos del proceso	59
Tabla 23. Involucrados e interesados.....	59
Tabla 24. Fuentes de datos alternativas	64
Tabla 25. Pruebas de concepto realizadas.....	65
Tabla 26. Datos disponibles.....	66
Tabla 27. Resultados del análisis de Python y sus librerías.....	67
Tabla 28. Datos de los recursos necesarios.....	69
Tabla 29. Interesados en el proyecto.....	70

Tabla 30. Riesgos identificados	71
Tabla 31. Matriz de evaluación de riesgos.....	72
Tabla 32. Recursos necesarios para implementación del nuevo modelo.....	82
Tabla 33. Costos de licencias.....	83
Tabla 34. Tiempo de ejecución de la propuesta.....	84
Tabla 35. Costos totales	85
Tabla 36. Resumen de costos y beneficios	87
Tabla 37 Resumen de variables	87

Capítulo 1. Introducción

La presente sección contextualiza el proyecto. Por lo tanto, se describen los antecedentes de la organización, el área y el equipo de trabajo para el desarrollo de esta investigación, proyectos similares, la situación problemática con sus respectivos beneficios, el objetivo general y los objetivos específicos. También, se detallan las actividades del alcance, supuestos, limitaciones, exclusiones, y entregables de producto, académicos y de gestión.

1.1. Descripción general

En las organizaciones de la actualidad, los datos funcionan como el principal insumo para sus actividades diarias. Por esta razón, muchas optan por tomar las ventajas de la inteligencia de negocios para desarrollar reportes que faciliten el proceso de la toma de decisiones, la cual debe basarse en datos.

Equifax es una compañía con operaciones a nivel mundial, sus principales actividades se basan en la gestión de datos de distintos tipos. Debido a esto, el rol del departamento de operaciones ha tenido un alto impacto. Su principal labor es velar por la eficiencia de las actividades de la empresa, por lo cual, necesita contar con la mayor cantidad de datos, y que estos sean lo más actualizados y certeros posibles, para optimizar el rendimiento de los distintos equipos.

Actualmente, la organización está pasando por un proceso de renovación de aplicaciones, al tiempo que elimina otras por razones organizacionales. Esto ha sido un inconveniente para el departamento de Servicios de Datos, ya que se ha prescindido de herramientas que eran consumidas por este para la creación de reportes. En consecuencia, se han presentado atrasos en el proceso de ejecución de reportes e inconvenientes con la calidad de estos.

Dado lo anterior, el presente proyecto tiene como principal objetivo proponer un nuevo proceso para la creación de reportes, al aprovechar las herramientas ofrecidas por la inteligencia de negocios. De la mano con estas, se propone la implementación de un proceso que utilice las herramientas ETL (Extracción, Transformación y Carga, por sus siglas en inglés), que a su vez elimine el trabajo manual realizado por los encargados de elaborar los reportes. Esto con el fin de otorgar al departamento de operaciones (Business Care y Boarding) reportes actualizados, certeros y automatizados.

1.2. Antecedentes

En esta sección se presentan la organización, misión, visión y valores de la empresa. Además, se describen los antecedentes, es decir, proyectos que tienen un alto interés y relación con el presente proyecto.

1.2.1. Descripción de la organización

Equifax se define como una compañía global con exclusivos activos de datos, tecnología y analítica que transforman conocimientos en información, para impulsar decisiones que permitan fomentar el avance de las personas (Equifax, Inc. 2022). Esta organización brinda una serie de servicios, como lo son la defensa del consumidor, ser guardián de la educación financiera y permitir un avance económico. Al ser una empresa que permite el acceso al crédito es parte de colaboraciones e innovaciones que abordan complejos desafíos sociales, como relaciones con la comunidad, la educación financiera a jóvenes desfavorecidos y otros problemas sociales.

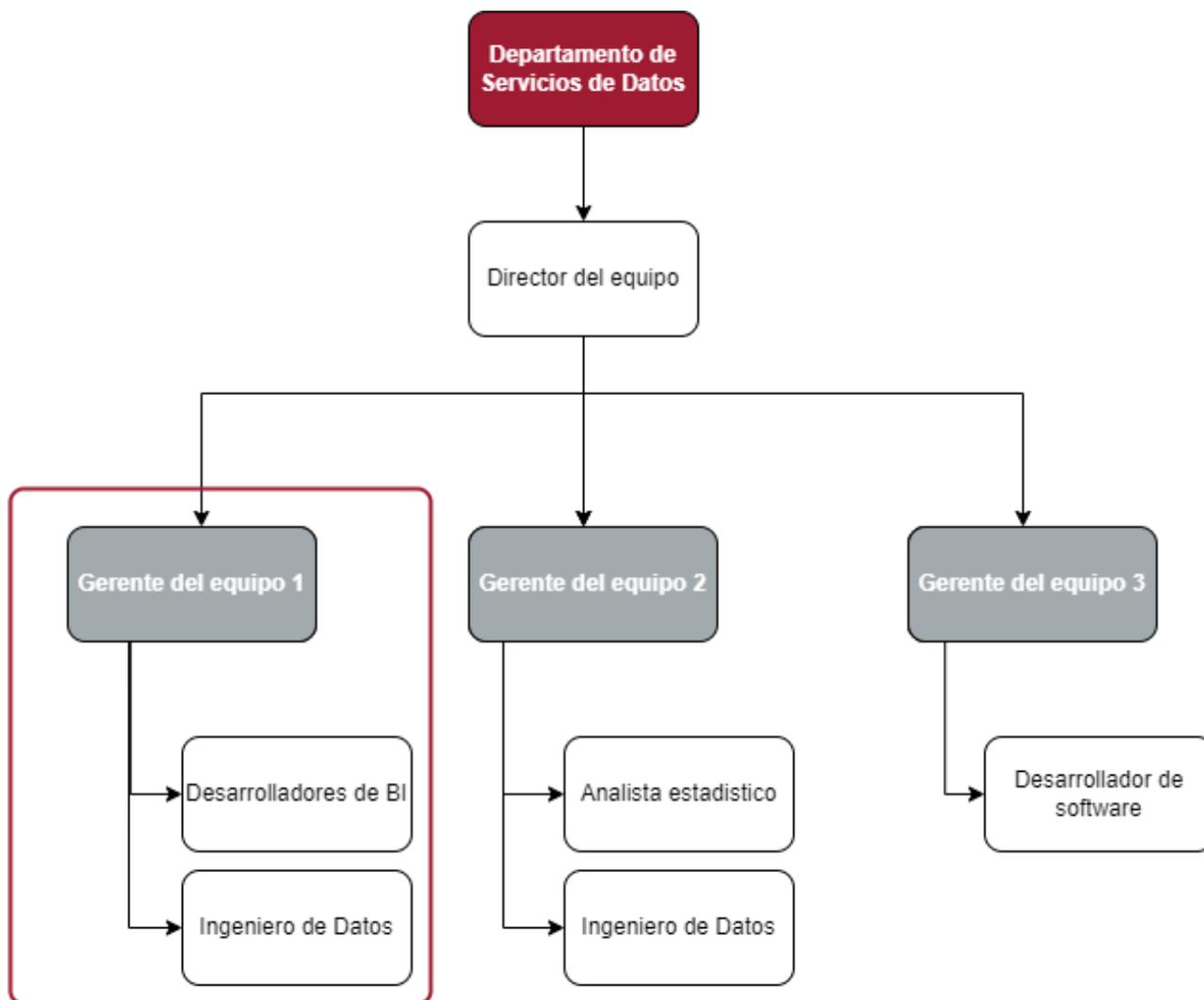
La empresa cuenta con 13,000 empleados distribuidos en 25 países en América del Norte, América Central, América del Sur, Europa, y las regiones de Asia y el Pacífico. Su sede principal está en Atlanta, Estados Unidos. Para el 2021, reportó ingresos de \$4.9 millones, los cuales permitieron que 13.5 millones de consumidores latinoamericanos accedieran a distintos créditos y 8.1 millones de personas dentro de Estados Unidos obtuvieron un crédito universitario (Equifax, Inc. 2022). Todos estos logros se cumplieron gracias a las cuatro principales unidades de negocio que tiene Equifax, las cuales se pueden definir como:

- **Unidad de negocio A:** el mercado en el que se enfocan es finanzas, telecomunicaciones, minoristas, gobiernos y empresas de cobranza a nivel estadounidense.
- **Unidad de negocio B:** el mercado en el cual se centran es finanzas, telecomunicaciones, minoristas, gobiernos, y empresas de cobranza a nivel latinoamericano, europeo y asiático.
- **Unidad de negocio C:** su mercado se basa en empresas de todo tamaño, prácticamente de todos los sectores.
- **Unidad de negocio D:** el mercado al que se dirige son los consumidores finales.

De acuerdo con lo mencionado por Tenorio (2022), Equifax tiene presencia en suelo costarricense desde 1995 con el Centro de Servicios Compartidos (SCC, por sus siglas en inglés), el cual se encuentra en la Zona Franca Metropolitana en Barreal de Heredia. En la actualidad, cuenta con aproximadamente 1200 empleados. En esta sede, se realizan actividades como operaciones del negocio, tecnologías y seguridad de la información, finanzas, y servicios de recursos humanos para Latinoamérica.

El área donde se desarrolla el proyecto es dentro de la unidad de operaciones de negocio, específicamente en el departamento de Servicios de Datos. Este se encuentra dividido en tres equipos, tal como se muestra en la **Figura 1**, los cuales se encargan de atender distintos tipos de clientes internos de la organización.

Figura 1. Organigrama del departamento de Servicios de Datos



Fuente: elaboración propia a partir de información brindada por el departamento de Servicios de Datos (2022).

1.2.1.1. Misión

La misión de una organización facilita comprender el motivo o la razón de ser de una organización. En el caso de Equifax es:

“Maximizar el valor de las inversiones de los accionistas de la compañía, manteniendo nuestros valores fundamentales” (Equifax, 2022).

1.2.1.2. Visión

La visión de una empresa permite tener claro lo que se espera que esta sea o alcance en un futuro. En el caso de Equifax es:

“Ser un líder global confiable en datos, análisis avanzados y tecnología que crea soluciones y perspectivas innovadoras para ayudar a los clientes a impulsar el crecimiento y fomentar el avance de las personas” (Equifax, 2022).

1.2.1.3. Valores

Los valores de Equifax (2022) son:

- Hacer lo correcto.
- Clientes primero.
- Innovar y crecer.
- Decir/Hacer.
- Dar lo mejor de nosotros.
- Un solo Equifax.

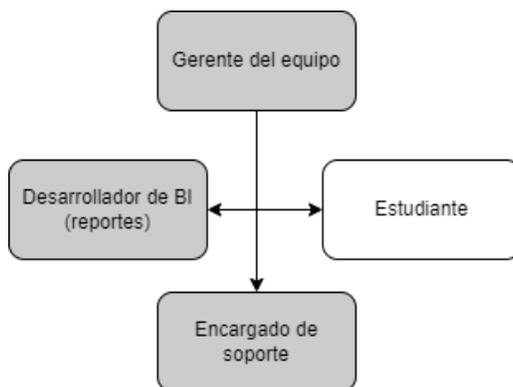
1.2.1.4. Equipo de trabajo

El equipo de trabajo involucrado en la ejecución del presente proyecto está conformado por los siguientes miembros:

- **Gerente del equipo:** cumple el rol de ser el patrocinador del proyecto. Él es su promotor, debido a que existe la necesidad de automatizar y eficientizar los procesos que realizan los miembros del equipo. Por otro lado, se encarga de firmar los documentos relacionados con la gestión del proyecto, tales como la carta de aceptación, evaluaciones por parte de la organización, documentos de control de cambios, entre otros.
- **Desarrollador de Inteligencia de Negocios:** es el encargado de ejecutar los reportes, por lo cual, es fundamental para el entendimiento del proceso *As-is* y la clase de la información que este deben incluir. Además, conoce los tiempos de entrega y es el principal punto de contacto con el cliente final.
- **Encargado de soporte:** es un integrante que forma parte del equipo, debido a que es importante que las soluciones implementadas sean sostenibles a largo plazo. Por lo cual, es necesario ajustarse a las políticas de la organización y a las herramientas que están en su lista blanca.
- **Estudiante:** forma parte del departamento de Servicios de Datos en el periodo definido de ejecución del proyecto. Además, es el encargado de desarrollar el proyecto y presentar los resultados tanto a la academia (ATI) como al equipo de Equifax.

En la **Figura 2**, se encuentra el organigrama de los miembros que componen el equipo de trabajo.

Figura 2. Organigrama del equipo de trabajo



Fuente: elaboración propia a partir de información brindada por el departamento de Servicios de Datos (2022).

1.2.2. Trabajos similares realizados dentro y fuera de la organización

Es importante indagar del tema previo a ejecutar esta investigación. Este paso permite encontrar fuentes y trabajos, los cuales pueden llegar a ser insumos del presente proyecto. A continuación, se presenta una descripción de investigaciones que se han realizado dentro y fuera de la empresa, y guardan relación con este proyecto.

Trabajos dentro de la organización

En la **Tabla 1**, se detallan los proyectos internos realizados, que sirven como insumo para el desarrollo de la presente investigación.

Tabla 1. *Proyectos internos*

Autor / Encargado	Título	Objetivo
Jianbo Jin	Recredentialing Project.	Desarrollar un modelo el cual agrupe insumos y genere un reporte del estado de los clientes cada tres meses.
Douglas Ricardo Pérez Gago	Desarrollo de herramienta de inteligencia de negocios para el área de ventas.	Configurar una herramienta de inteligencia de negocios para la generación de reportes personalizados y permitir la actualización automática de los datos para eliminar el trabajo manual de los colaboradores de la empresa, y así minimizar el riesgo de errores humanos en el proceso.

Fuente: Vindas (2022) y Pérez (2017).

Trabajos fuera de la organización

En la **Tabla 2**, se detallan los trabajos externos realizados que sirven como insumo para el desarrollo del presente proyecto. Estos fueron realizados por estudiantes del Instituto Tecnológico de Costa Rica.

Tabla 2. *Proyectos externos*

Autor	Título	Objetivo
Luis David Acuña Campos	Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes.	En este documento, se detallan dos propuestas de solución frente a las necesidades y problemática presentes para la generación de reportes. La empresa tiene necesidades específicas de reportes, pues integra información de tres sistemas, además, solo un colaborador es responsable y tiene los conocimientos para realizarlos. Este proceso es manual, lo cual implica una serie de desventajas, como el tiempo para generar los reportes y la entrega oportuna de estos.
José Daniel Navarro Romero	Propuesta de implementación de una solución de inteligencia de negocios que apoye el análisis de información y la toma de decisiones para Unity Service Delivery and Electronic Payment.	Este proyecto tiene como propósito implementar una solución de inteligencia de negocios (BI), a través de un Datamart y un Tablero de Mando, que apoye el análisis de información y la toma de decisiones estratégicas y operativas.
Minor Andrés Umaña Murillo	Implementación de una solución de inteligencia de negocios que apoye el análisis de información y la toma de decisiones en Greenpay.	Este trabajo busca implementar una solución de inteligencia de negocios (BI) mediante un Datamart, para facilitar la generación de visualizaciones orientadas a la gerencia comercial y operaciones de la empresa Greenpay.

Fuente: Acuña (2018), Navarro (2019) y Umaña (2020).

1.3. Planteamiento del problema

En este apartado, se describe la situación problemática identificada dentro del equipo de Servicios de Datos, la cual motiva el desarrollo del presente proyecto. Además, se mencionan los beneficios esperados con la solución del problema.

1.3.1. Situación problemática

El problema central identificado para este proyecto es **la deficiencia en el proceso de extracción, transformación, carga y visualización de datos realizado por el equipo de Servicios de Datos**. Esto se encuentra detallado en la **Figura 3**. Una serie de inconvenientes dentro del proceso de generación de reportes permitieron identificar el problema; esta se describe a continuación.

Como se mencionó en la sección 2.1, Equifax es una empresa transnacional dedicada a brindar servicios de información orientada a los datos crediticios. Cuenta con 4 unidades de negocio, las cuales se encargan de generar ganancias a la organización. Una de las más importantes es la **Unidad de negocio A**, en esta, los departamentos de Business care y Boarding gestionan los niveles de atención de los agentes dentro del centro de llamadas brindan soporte de ventas y verifican el rendimiento de las operaciones de Equifax.

Debido a políticas de la organización y el cese de relaciones con proveedores, el Departamento de Datos ha sufrido cambios en sus principales herramientas, las cuales se encuentran en la **Tabla 3** enlistadas.

Tabla 3. *Cambios en las herramientas*

Herramienta	Última fecha de uso
Microsoft Office	1/2020
Java	6/2019
Pentaho	9/2022
Bamboo	9/2022
SAS	10/2022

Fuente: Vindas (2022).

Para los departamentos de Business Care y Boarding, se generan un total de 15 reportes, los cuales consumen información de distintas fuentes de datos. Estas se encuentran aisladas entre sí, debido a que están ubicadas en distintos silos. Algunas de bases de datos utilizadas son:

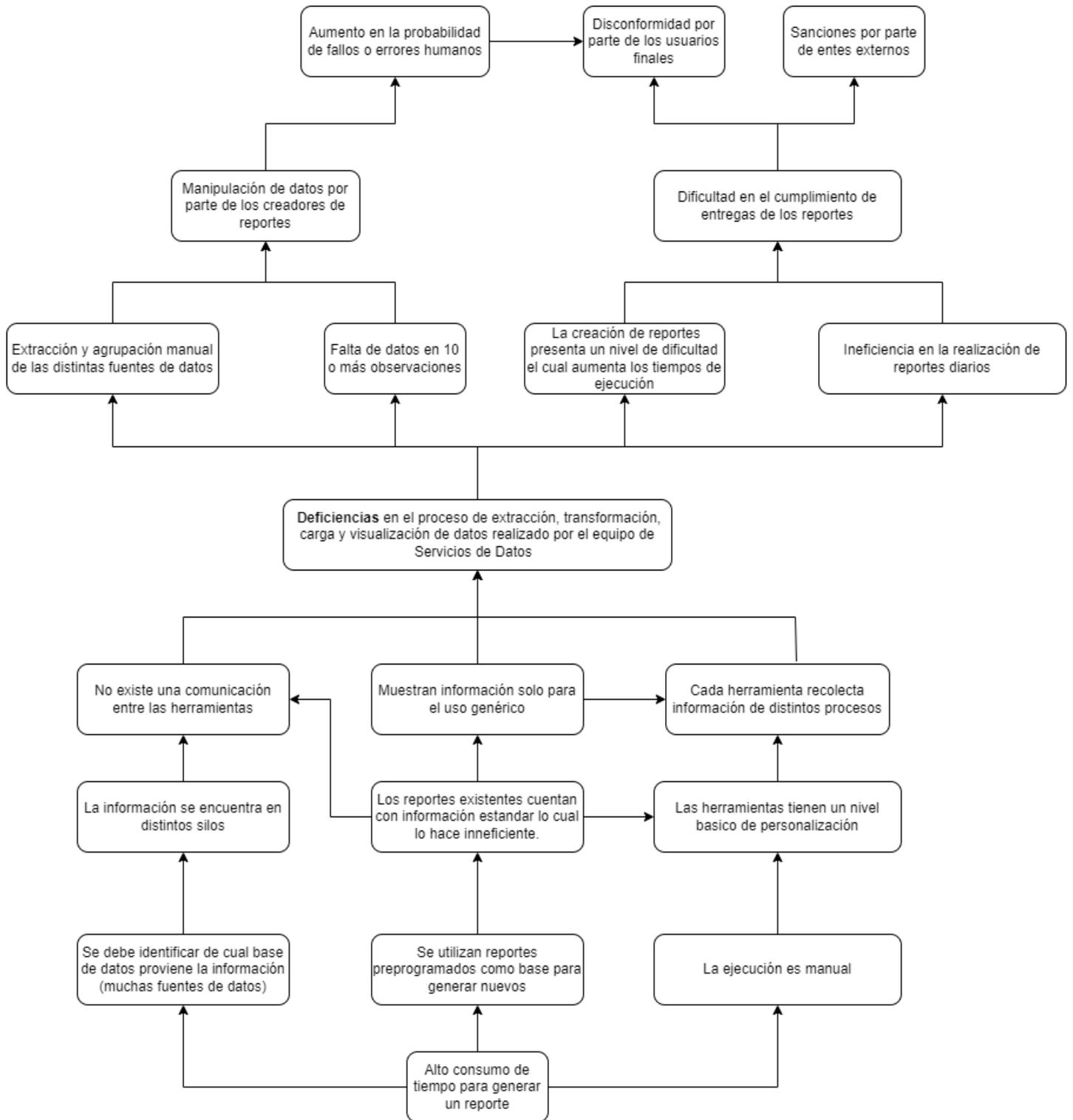
- *SalesForce* (versión actual)
- *SalesForce* (versión anterior)
- *Qualtrics*
- *AODocs*
- *Looker*
- Base de datos PostgreSQL

Estas herramientas cuentan con reportes o vistas precreadas por los dueños del producto o *product owners*, los cuales deben ser utilizados por el equipo como principal insumo para generar los reportes solicitados por los clientes. Esto debido a la baja capacidad de personalización que presentan las herramientas que emplean.

El equipo genera quince reportes para tales departamentos. De estos, seis deben realizarse a diario y ser entregados en las primeras horas del día, ya que son un insumo primario para los representantes y encargados de dichos departamentos para su programación diaria de tareas. El resto, se crean semanalmente. Los reportes fueron creados utilizando las herramientas mencionadas en la **Tabla 3**, por lo cual, el equipo ha tenido que implementar métodos manuales dentro de su desarrollo.

Actualmente, el equipo uno del departamento de Servicios de Datos cuenta con dos personas encargadas de realizar estos reportes y es una tarea recurrente. Por esta razón, generar los reportes se ha transformado en un proceso que consume una gran cantidad de tiempo, según lo mencionado por Vindas (2022): más de cuatro horas de cada uno de los colaboradores que se encarga de ejecutarlos. En consecuencia, hay problemas en el cumplimiento de entregas. Otro aspecto mencionado es que, por ser un proceso manual, la probabilidad de encontrar errores en los reportes ha aumentado.

Figura 3. *Árbol del problema*



Fuente: elaboración propia a partir de información brindada por el departamento de Servicios de Datos (2022).

1.3.2. Justificación del proyecto

Este proyecto busca mejorar el proceso actual de creación de reportes, por medio del uso de inteligencia de negocios y la gestión de procesos. La inteligencia de negocios (en adelante BI, por sus siglas en inglés) ayuda a las organizaciones a analizar datos históricos y actuales, para descubrir rápidamente información útil que permita tomar decisiones estratégicas. Las herramientas de inteligencia lo hacen posible mediante el procesamiento de grandes conjuntos de datos, a través de múltiples fuentes y la presentación de hallazgos en formatos visuales que son fáciles de entender y compartir (IBM, 2022). Por otro lado, la gestión de procesos (en adelante, BPM, por sus siglas en inglés) es una disciplina que utiliza varios métodos para descubrir, modelar, analizar, medir, mejorar y optimizar los procesos de negocio (Gartner, s.f.).

Dentro de una organización como Equifax, BI tiene un rol importante, al permitir gestionar los datos de las distintas unidades de negocios, y facilitar la centralización de la información y la creación de reportes. Como se ha mencionado en las secciones anteriores, el departamento de Servicios de Datos ha sufrido cambios en los últimos años, lo cual ha ocasionado que los procesos originales se tornen ineficientes. Ante estas situaciones, se han presentado los siguientes inconvenientes:

- **Consumo de más de 4 horas en la extracción de información:** los encargados deben ejecutar los reportes precreados en cada una de las fuentes de datos y luego descargarlos en un formato universal (CSV), para manipularlos de forma sencilla.
- **Manipulación de la información:** una vez que tienen los datos, deben agruparlos. Algunos de los métodos utilizados son copiando y pegando segmentos de los otros reportes, o empleando funciones como BUSCARV, para obtener el resultado final en un mismo archivo, que por lo general es de tipo Excel.
- **Carga de datos para la generación de reportes:** debido a las políticas de la empresa, solo ciertos colaboradores cuentan con licencias de Microsoft, por esta razón, la información debe ser cargada en hojas de datos de Google o en alguna base de datos.

Así, al ejecutar este proyecto, se busca mejorar la realización de reportes por parte del departamento de Servicios de Datos, desde una perspectiva de calidad y de tiempos de entrega. Como se ha mencionado, Equifax ha recibido multas y sanciones debido a la falta de control de la información. El encargado de presentar los reportes que van para el CFBP (Oficina para la Protección Financiera del Consumidor) es de Business Care y Boarding. Además, se pretende brindar satisfacción a los clientes con mejores tiempos de entrega y mayor calidad en los productos entregados.

1.3.3. Beneficios esperados o aportes del Trabajo Final de Graduación

Al identificar las problemáticas actuales del equipo y al desarrollar una solución de inteligencia de negocios que permita analizar e interpretar datos de manera automatizada, se pueden alcanzar los siguientes beneficios:

1.3.3.1. Beneficios directos

1. La estandarización del proceso de realizar los reportes para el departamento de Business Care y Boarding.
2. Disminución de los errores humanos en los reportes generados debido a la estandarización de proceso.
3. Documentación en los procesos de generación de reportes en Business Care y Boarding.
4. Favorecer el tiempo de respuesta a los clientes.

1.3.3.2. Beneficios indirectos

1. Satisfacción del cliente y del usuario.
2. Nuevas oportunidades de negocios.
3. Desarrollo de una cultura de innovación de procesos.

1.4. Objetivos del Trabajo Final de Graduación

En esta sección, se detalla el objetivo general y los objetivos del proyecto, con el propósito de resolver la problemática expuesta.

1.4.1. Objetivo general

Proponer un plan para la automatización de las actividades realizadas en el proceso de generación de reportes correspondiente a los grupos de Business Care y Boarding, en Equifax, durante el segundo semestre del año 2022.

1.4.2. Objetivos específicos

1. Analizar el estado actual del proceso de creación de reportes para la identificación de necesidades del equipo y el modelado del nuevo proceso.
2. Diseñar un nuevo proceso de extracción, transformación y carga de datos (ETL), su respectiva documentación y visualización de reportes para ser evaluados por medio pruebas de concepto.
3. Desarrollar un plan de implementación para la automatización del proceso de creación de reportes y su respectivo análisis de costo-beneficio para la valoración.

1.5. Alcance

El presente proyecto busca desarrollar un nuevo proceso que permita al equipo de Servicios de Datos de Equifax sacar provecho a las herramientas de la inteligencia de negocios. Para lograrlo, se pretende realizar las siguientes actividades:

- Análisis de la situación actual

Esta actividad se pretende conocer el proceso *As-is* de la creación de reportes. Es decir, identificar tanto las actividades realizadas como los participantes, y a su vez las herramientas, fuentes de datos y reportes generados para el equipo de Business Care y Boarding. Para finalmente, modelar el proceso actual utilizando la notación BPMN.

- Generar un modelo para el proceso de creación de reportes

Durante esta actividad, se elabora el nuevo proceso de creación de reportes. Se modela un nuevo procedimiento por medio del cual el equipo logra extraer, transformar y cargar los datos de forma sencilla, con el fin de crear los reportes de una mejor forma. Para esto, se analiza el proceso según lo descrito por Madison (2005) y Dumas (2019) en sus respectivos libros con una notación BPMN. Adicionalmente, se realiza una documentación la cual permita a los colaboradores conocer el proceso y entender tanto sus datos como sus pasos.

- Realizar pruebas de concepto

Se desarrollan pruebas de concepto con el fin de comprobar la factibilidad de implementación del nuevo proceso. Estas se realizan bajo el esquema ofrecido por Intel para desarrollarlas. Se utiliza el principio de Pareto para la selección de las pruebas de concepto que se emplean. Cabe destacar que las que estas pruebas de concepto abarcan el proceso ETL, dejando de lado la creación de los tableros.

- Realizar un plan de implementación

Se realiza un plan que permita a la organización crear un proyecto cuando desee implementar el proceso. Este cuenta con valores de tiempo estimado, debido a la limitación existente para su implementación completa y procesos burocráticos dentro de Equifax.

- Análisis de costo-beneficio

Finalmente, se realiza un análisis de costo-beneficio, para obtener los resultados en términos financieros de la factibilidad e implementación del proyecto.

1.6. Supuestos

En este apartado, se indican los factores y elementos que se pretenden cumplir con la realización del proyecto. Son los siguientes:

1. Apoyo por parte del equipo de Servicios de Datos para la realización completa y exitosa del proyecto.
2. Disposición del gerente del equipo, desarrollador y encargado de soporte para atender reuniones y consultas relacionadas con el proyecto.
3. La información necesaria para realizar el proyecto será entregada al estudiante a tiempo, conforme esta sea solicitada.
4. Exactitud y confiabilidad de la información brindada por el equipo de Servicios de Datos para llevar a cabo el proyecto.
5. La información brindada por parte del equipo del Departamento de Datos, sobre aspectos de costos operativos y salarios está sujeta a políticas de confidencialidad, por esto, se definieron montos estimados para el análisis financiero de la propuesta.

1.7. Entregables

Los entregables del proyecto facilitan llevar un control de cada actividad realizada. Estos proporcionan valor al cliente durante el desarrollo del proyecto y permiten que se culmine con lo establecido en las secciones anteriores. Algunos de estos entregables son solicitados por la organización y otros por la academia; son los siguientes:

1. **Análisis del proceso de generación de reportes:** pretende describir de forma explícita la situación actual del proceso de generación de reportes actual y sus respectivas mejoras. Se hace por medio de una evaluación, caracterización y diagramación del proceso.
2. **Análisis de costo beneficio:** busca brindar análisis desde una perspectiva financiera de la propuesta realizada para la organización.
3. **Plan de implementación del nuevo proceso:** se ofrece el plan de implementación del proceso realizado con sus respectivos ETL y automatizaciones.
4. **Documentación del proceso:** corresponde a un documento con la información del nuevo proceso para generar los reportes. En este, se encuentran los detalles del nuevo proceso y las pruebas de concepto realizadas.

1.8. Exclusiones del proyecto

A continuación, se enlistan aquellas actividades, detalles y aspectos en general que son excluidos del alcance de este proyecto:

- **Implementación de la solución en un ambiente de producción o desarrollo:** el proyecto consiste en una propuesta, por este motivo, la integración de la solución a las plataformas existentes le corresponderá a la empresa. La carga de datos se realiza en el ambiente de pruebas autorizado únicamente para hacer una prueba de concepto.
- **Capacitaciones sobre la propuesta:** no se brindaron capacitaciones al personal del Departamento de Datos sobre la solución propuesta, debido a las limitaciones de tiempo y alcance del proyecto.

1.9. Limitaciones

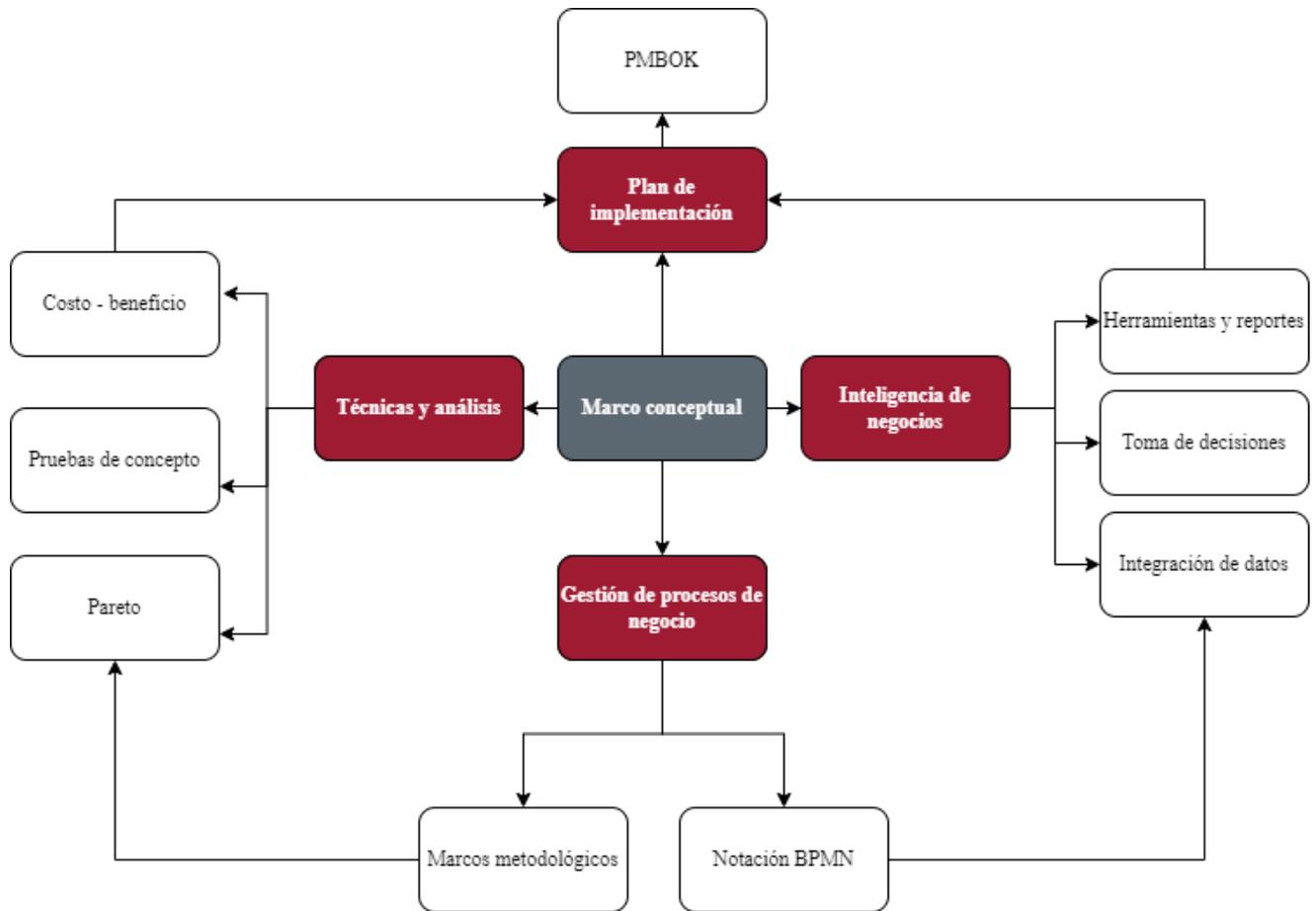
En esta sección, se indican explícitamente cuáles son los factores o elementos que en alguna medida restringen la realización del proyecto.

- Disponibilidad inmediata del gerente para atender consultas de este trabajo.
- Inexistencia de documentación sobre el proceso de generación de reportes u otra información necesaria para realizar el proyecto.
- Acceso de información considerada como confidencial para elaborar la presente investigación.
- Realizar y mostrar las visualizaciones del proceso planteado para el departamento de Servicios de Datos.

Capítulo 2. Marco conceptual

En el presente capítulo, se definen las bases conceptuales relevantes, ya sea teóricas o prácticas, para sustentar tanto el problema identificado como la propuesta de solución de este proyecto. Se indagan conceptos relacionados con procesos, inteligencia de negocios, proyectos y automatización; asimismo, se describen términos relacionados con análisis de costos y beneficios, los cuales serán ejecutados como parte de la propuesta de solución. En la **Figura 4**, se muestra gráficamente la relación entre conceptos.

Figura 4. Mapa de conceptos



Fuente: elaboración propia.

2.1. Inteligencia de negocios

El concepto de inteligencia de negocios corresponde a un proceso que ayuda a las organizaciones a analizar datos históricos y actuales, para descubrir rápidamente información útil y, en consecuencia, tomar decisiones estratégicas. Las herramientas de inteligencia empresarial lo hacen posible mediante el procesamiento de grandes conjuntos de datos a través de múltiples fuentes y la presentación de hallazgos en formatos visuales fáciles de entender y compartir (Microsoft, 2022). Dentro de este, existe una serie de conceptos que se definen a continuación.

2.1.1. Toma de decisiones

De acuerdo con Burns (2020), la toma de decisiones es “una serie de pasos que toma un individuo para determinar la mejor opción o curso de acción para satisfacer sus necesidades” (párr. 1). Este proceso es realizado por todos los seres humanos día a día, desde decidir la comida hasta la ropa. La toma de decisiones tiene un rol muy importante en un contexto empresarial cuando lo combina con el concepto de inteligencia de negocios, pues lleva a la toma de decisiones basada en datos. De acuerdo con el sitio web de Tableau (2022), “La toma de decisiones basadas en los datos se define como el uso de hechos, métricas y datos para guiar decisiones de negocios estratégicas que se alineen con las metas, los objetivos y las iniciativas de una organización” (párr. 1). En la actualidad, debido a la gran cantidad de información que se genera y a su complejidad, ha tomado un rol importante el uso de herramientas de inteligencia de negocios.

2.1.2. Datos

De acuerdo con Carrión (2017), “Un dato es un conjunto discreto, de factores objetivos sobre un hecho real. Dentro de un contexto empresarial, el concepto de dato es definido como un registro de transacciones” (párr. 3). Así, este puede llegar a ser cualquier cosa y, por lo general, se ocupa una agrupación de datos para generar conocimiento.

2.1.3. Información

Carrión (2017) define información como “un mensaje, normalmente bajo la forma de un documento o algún tipo de comunicación audible o visible. Como cualquier mensaje, tiene un emisor y un receptor” (párr. 6). A diferencia de los datos, la información tiene significado y está organizada para algún propósito. Por su parte, los datos se convierten en información cuando su creador les añade significado.

2.1.4. Bases de datos

Según Oracle (s.f.), “Una base de datos es una recopilación organizada de información o datos estructurados, que normalmente se almacena de forma electrónica en un sistema informático” (párr. 1). Este tipo de herramientas suelen contar un gestor, el cual ayuda a los encargados a controlarlas. Las bases de datos han sido las sucesoras de las hojas de cálculos, dentro de las principales diferencias entre ambas están:

- La forma en la que se almacena y manipula la información: en una hoja de cálculos, es sencillo realizar modificaciones y almacenar nuevos valores, lo cual tiene sus ventajas y desventajas. Dentro de las bases de datos, se brinda integridad a la información y un nivel mayor de seguridad.
- Acceso a los datos: si bien las hojas de cálculos pueden contener claves, no lo hace un método fácil de gestionar por posibles filtraciones de las claves; por otro lado, las bases de datos brindan la facilidad de asignar roles dentro de sí a personas determinadas, para limitar el uso que les den a los datos.
- Cantidad de datos que se pueden almacenar: en las hojas de Excel existen límites de la cantidad de datos que se pueden tener, además, se pierde mucho desempeño cuando se realizan operaciones en hojas con gran cantidad de datos. Por su parte, las bases de datos almacenan grandes cantidades de datos y brindan una fácil manipulación de estos con cuatro funciones básicas (crear, modificar, eliminar y seleccionar, CRUD, por sus siglas en inglés).

2.1.5. BigQuery

De acuerdo con Google, BigQuery “es un almacén de datos de Google de bajo coste y totalmente administrado que permite extraer analíticas de *petabytes* de datos” (Google, s.f., párr. 1). Es una de las opciones de base de datos ofrecida por Google, con la capacidad de gestionar grandes cantidades de datos.

2.1.6. Indicadores de desempeño (KPI)

De acuerdo con Martins (2022), estos son “una métrica cuantitativa que muestra cómo tu equipo o empresa progresa hacia tus objetivos empresariales más importantes” (párr. 2). Estos indicadores son muy importantes para la inteligencia de negocios, ya que funcionan como puntos claves para las herramientas mencionadas. Cuando se utiliza la inteligencia de negocios al querer saber los resultados, se encuentran KPIs en la elaboración de tableros. Algunas características importantes destacadas por Martins (2022) son:

- Los KPIs ayudan a establecer objetivos estratégicos.
- Son medibles.
- Informan sobre la planificación de recursos.
- Brindan información para la toma de decisiones.

2.1.7. Integración de datos

Otro concepto importante relacionado con la inteligencia de negocios es la integración de datos. Google (2022) lo define como “un proceso que consiste en reunir datos de diferentes fuentes para obtener una vista unificada y más valiosa de ellos” (párr. 6). Esto permite a las organizaciones recolectar la mayor cantidad de datos generados, unificarlos en un mismo punto para relacionarlos entre sí y generar distintos insumos, los cuales pueden ser utilizados para la toma de decisiones o validación en estrategias. Existen distintos métodos para integrar, el más conocido, y que es implementado en el presente proyecto, es el ETL (extracción, transformación y carga, por sus siglas en inglés). Este proceso proporciona la base para el análisis de datos y flujos de trabajo para automatización. En la **Tabla 4**, se describe cada una de las etapas y tareas realizadas en tal proceso.

Tabla 4. Descripción del proceso de ETL

	Extracción	Transformación	Carga
Descripción	Durante la extracción de datos, los que están sin procesar se copian o exportan desde las ubicaciones de origen a un área de preparación.	Los datos sin procesar se someten a procesamiento para prepararlos a su caso de uso analítico previsto.	Los datos transformados se mueven desde el área de ensayo a un almacén de datos de destino.
Tareas por realizar	Conexión a fuentes de datos (bases de datos, archivos planos, páginas web).	<ul style="list-style-type: none"> - Se filtra la información. - Se realizan cálculos. - Se brinda formato a tablas. 	Se almacenan los datos procesados para ser utilizados fácilmente. Usualmente, en este paso se realizan operaciones de actualización o inserción para nuevos datos, o borrado para reemplazar los datos existentes.

Fuente: adaptado de “What is ETL (Extract, Transform, ¿Load)?”, de IBM Cloud Education (2020).

2.1.8. Herramientas y reportes

De acuerdo con la Real Academia Española (2022), una herramienta es un instrumento que permite realizar una actividad determinada. En la inteligencia de negocios, se encuentra una serie de herramientas de BI, las cuales corresponden a “tipos de software de aplicación que recopilan y procesan grandes cantidades de datos no estructurados que proceden de sistemas internos y externos, incluidos libros, publicaciones, documentos, historias clínicas, imágenes, archivos, correo electrónico, vídeos y otros orígenes empresariales” (Microsoft, s.f., párr. 1). Con estas herramientas, se toman los datos del proceso ETL y se elaboran los reportes. De acuerdo con Qlik (2022), estos últimos se definen como “el proceso de usar una herramienta de BI para preparar y analizar datos para encontrar y compartir información procesable” (párr. 3). Por medio de estos reportes, se hacen cálculos para conocer el estado de la organización, según los indicadores clave de desempeño o KPI (por sus siglas en inglés).

2.1.9. Software

Según la RAE (s.f.), “el software es un conjunto de programas, instrucciones y reglas informáticas que permiten ejecutar distintas tareas en una computadora”. Este tipo de programas facilitan realizar tareas diarias, por medio de la utilización de todos los beneficios y capacidades de las computadoras actuales.

2.1.10. Visualización de datos

La visualización de datos se puede definir como representación gráfica de información y datos. Al utilizar elementos visuales como cuadros, gráficos y mapas, las herramientas de visualización de datos proporcionan una manera accesible de ver y comprender tendencias, valores atípicos y patrones en los datos (Tableau, s.f.). Con este tipo de herramientas podemos mostrar información importante de una forma más sencilla. Por medio de los colores, formas y gráficos podemos generar tableros los cuales puedan ser captados de mejor forma y así compartir los datos e información efectivamente.

2.1.11. Interfaces de programación de aplicaciones (API)

Una interfaz de programación de aplicaciones (API, por sus siglas en inglés) “hace referencia a cualquier software con una función distinta. La interfaz puede considerarse como un contrato de servicio entre dos aplicaciones” (Amazon, s.f., párr. 2). Con esta herramienta, se pueden generar conexiones para transferir información, o bien, aprovechar ciertas funcionalidades ajenas a otras aplicaciones.

2.2. Gestión de procesos de negocio (BPM)

La gestión de procesos de negocio (en adelante BPM), de acuerdo con Dumas (2019), “es el arte y la ciencia de supervisar cómo se realiza el trabajo en una organización para garantizar resultados consistentes y aprovechar las oportunidades de mejora” (p. 1). Esta actividad, por lo general, se realiza debido a tres principales razones: reducción de costos, reducción en tiempos de ejecución y reducción de errores (Dumas et al. 2019). Esta actividad favorece a las organizaciones modificar sus procesos de una forma integral, al tomar el proceso completo, y realizar un análisis y caracterización de cada una de sus etapas. De esta forma, se logra determinar cuáles actividades o eventos dentro del proceso agregan valor al producto final y permite reconocer las tareas que pueden llegar a automatizarse.

2.2.1. Marcos metodológicos

Para la ejecución de iniciativas BPM, se encuentran varias metodologías de trabajo. Dos de las más conocidas y estudiadas en la academia son las planteadas por Marlon Dumas y por Dan Madison. Ambas buscan un fin común, pero sus enfoques son diferentes.

2.2.1.1. Metodología de BPM según Marlon Dumas

De acuerdo con el libro *Fundamentals of Business Process Management* realizado por Marlon Dumas et al. (2019), la gestión de procesos o BPM debe realizarse de forma heurística. El modelo de rediseño de procesos que mencionan abarca una serie de criterios, los cuales permiten analizar los procesos en distintas perspectivas:

- Perspectiva financiera.
- Perspectiva del cliente.
- Perspectiva interna.
- Perspectiva de aprendizaje y crecimiento.

Este modelo es funcional cuando realiza un análisis de varios procesos, es decir, al rediseñar más de un proceso, ya que permite caracterizarlos de acuerdo si son centrales de negocio; conocer sus relaciones si se descomponen en más procesos, o bien, si son la base otros; e identificar en qué nivel esta cada uno, si son procesos operacionales, centrales o de negocio. Por otro lado, se realiza un análisis cualitativo del proceso, para comprender cuáles actividades generan desperdicios (retrabajo, traslado, o en espera) y las causas raíz de los problemas dentro del proceso, etc. Posteriormente, se analiza de forma cuantitativa las actividades realizadas dentro del proceso por medio de pruebas de tiempos, cuellos de botella, costos, ciclos de tiempo, etc. Todo esto con el fin de rediseñar un proceso que optimice cada una de las actividades, para obtener mejores resultados.

2.2.1.2. Metodología de BPM según Dan Madison

De acuerdo con el libro *Process Mapping, Process Improvement, and Process Management*, de Dan Madison, se establece una metodología más práctica para el rediseño de procesos de negocio. Primeramente, se brinda una serie de síntomas para identificar los procesos con problemas. Estos, a su vez, pueden ser analizados desde distintas perspectivas, según Madison (2005):

- Perspectiva de frustración.
- Perspectiva de tiempo.
- Perspectiva de calidad.
- Perspectiva de costos.

Una vez identificados los procesos rotos, este autor establece cuatro fases para el rediseño de procesos. En la **Figura 5**, se encuentran tales fases y los diez pasos definidos para concluir efectivamente el rediseño.

Figura 5. Metodología de rediseño de procesos según Dan Madison

Fase #1: Preparación

- Introducción al rediseño
- Creación de equipo de trabajo

Fase #2: Análisis del proceso

- Proceso *As-is*
- Entrevistas
- Mejores prácticas

Fase #3: Rediseño

- Primer borrador del proceso
- Validación y pruebas
- Diseño final
- Implementación

Fase #4: Mejora continua

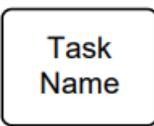
- Métricas y mejora continua

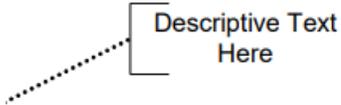
Fuente: adaptado de *Process Mapping, Process Improvement, and Process Management*, por Dan Madison (2005).

2.2.2. Notación BPMN

Existe una notación estandarizada que brinda herramientas para realizar de forma gráfica el proceso de BPM, se denomina BPMN. De acuerdo con la Object Management Group® Standards Development Organization (2011), “El estándar BPMN (Business Process Model and Notation) permite a las empresas visualizar sus procedimientos internos de negocio de forma gráfica y proporciona la notación estándar para la comunicación de procesos” (p. 1). Este es útil cuando se desea transmitir visualmente cada una de las tareas, eventos y decisiones realizadas dentro de un proceso. En la **Tabla 5**, se detallan algunos de los elementos utilizados dentro de la notación BPMN.

Tabla 5. Elementos de la notación BPMN

Elemento	Descripción	Notación
Evento	Este es algo que “sucede” durante el transcurso de un proceso. Afecta el flujo del modelo y por lo general tiene una causa (disparador) o un impacto (resultado).	
Inicio	Evento de inicio que dónde comenzará un proceso en particular.	
Intermedio	Este ocurre entre un evento de inicio y uno de finalización. Afecta el flujo del proceso.	
Final	El evento de fin indica dónde termina un proceso o coreografía.	
Actividad	Una actividad es un término genérico para el trabajo que la empresa realiza en un proceso. Puede ser atómica o no atómica (compuesta).	
Tarea	Esta corresponde a una actividad atómica que se incluye dentro de un proceso. Se usa cuando el trabajo en el proceso no se desglosa a un nivel más fino de detalle.	
Compuerta	Una puerta de enlace se utiliza para controlar la divergencia y la convergencia de los flujos de secuencia en un proceso.	
Flujo	El flujo normal se refiere a las rutas del flujo de secuencia, que no comienzan desde un evento intermedio adjunto al límite de una actividad.	
Flujo condicional	Un flujo de secuencia puede tener una expresión de condición que se evalúa en tiempo de ejecución, para determinar si se utilizará o no. Si el flujo condicional sale de una actividad,	

Elemento	Descripción	Notación
	entonces, el flujo de secuencia tendrá un mini diamante al comienzo de la conexión.	
Flujo de mensaje	Este se utiliza para mostrar el flujo de mensajes entre dos participantes que se encuentran preparados para enviarlos y recibirlos.	
Anotaciones de texto	Estas corresponden a un mecanismo para que un modelador proporcione información de texto adicional para el lector de un diagrama BPMN.	

Fuente: adaptado de *Business Process Model and Notation* (BPMN), por OMG (2011).

2.3. Técnicas y análisis

Al momento de realizar una investigación, se toman en cuenta distintas técnicas y análisis que ya han sido probados, para ponerlos en práctica. Estos permiten realizar comprobaciones o tomar decisiones que impactan de una u otra forma el proyecto, como facilitar su ejecución. En esta sección se conceptualizan el análisis costo-beneficio, técnicas como las pruebas de concepto y la técnica de Pareto.

2.3.1. Análisis de costo-beneficio

Las organizaciones con fines de lucro tienen como fin incrementar sus niveles de rentabilidad a partir de la obtención de utilidades. Por esta razón, utilizar técnicas como el análisis de costo-beneficio es de suma utilidad. De acuerdo con Aguilera (2017), “el análisis del costo-beneficio es un proceso que, de manera general, se refiere a la evaluación de un determinado proyecto, de un esquema para tomar decisiones de cualquier tipo” (p. 328). Esto quiere decir que involucra determinar el total de costos y beneficios de todas las alternativas para seleccionar la mejor o más rentable. Según Aguilera (2017), esta técnica se relaciona de manera directa con la teoría de la decisión, debido a que el determinar la factibilidad de un proyecto a partir de los costos y beneficios es una de las decisiones más influyentes al momento de la elección final.

2.3.2. Prueba de concepto

De acuerdo con Halicki, Mirabal, Price y Buck (2022) “Las pruebas de concepto, es una oportunidad para proporcionar la evidencia de que la solución propuesta resuelve los problemas empresariales” (párr. 1). Esta herramienta favorece la probabilidad de éxito de un software o proceso, debido a que impulsa una decisión técnica y segura a las organizaciones. Algunas de las principales características destacadas por los autores citados son:

1. Demuestra que un modelo de negocio o idea es factible y funcionará para resolver el problema empresarial.
2. Normalmente, implica de una a tres características o funcionalidades.
3. Puede estar en una o varias tecnologías.
4. Normalmente, está orientado hacia un escenario determinado y demuestra lo que el cliente necesita saber para tomar la decisión técnica o de seguridad.
5. Solo se usa como demostración y no entrará en producción.
6. Se basa en las TI y la habilitación. (párr. 2)

2.3.3. El principio de Pareto

El principio de Pareto o la regla 80/20 fue descrito, según Kruse (2016), por Vilfredo Federico Dámaso Pareto en 1848, quien realizó una investigación en la cual notó que el 80% de las tierras correspondían al 20% de la población. Este principio, poco a poco, fue trasladado a otras áreas. Por ejemplo, menciona Kruse (2016), que el 20% de los representantes de ventas de una organización generan el 80% de las ventas totales, o bien, el 20% de los errores más informados de software causan el 80% de las fallas de este. De acuerdo con Dumas et al. (2018), “el objetivo del análisis de Pareto es identificar qué problemas o qué factores causales de un problema deben tener prioridad” (p. 232). Básicamente, este análisis sostiene que un pequeño número de factores son responsables de la mayor parte de un efecto dado. Según los autores mencionados, un enfoque típico para realizar el análisis de Pareto sería:

1. Definir el efecto a analizar y la medida a través de la cual se cuantificará.
2. Identificar todos los temas relevantes que contribuyen al efecto a analizar.
3. Cuantificar cada tema según la medida elegida. Este paso se puede realizar sobre la base del registro de problemas, en particular, la columna de impacto cuantitativo del registro.

2.4. Plan de implementación

De acuerdo con la Real Academia Española (s.f.), se entiende como implementación “Poner en funcionamiento o aplicar métodos, medidas, etc. para llevar a cabo algo”. Contextualizando este término al presente proyecto, se corresponde a la puesta en práctica del nuevo proceso realizado. El objetivo de esta investigación es desarrollar un plan de implementación que sea guía para la empresa. Para lograrlo, se emplea como guía el PMBOK propuesto por el “Project Management Institute”.

2.4.1. PMBOK

El Project Management Body of Knowledge (en adelante PMBOK) es una guía de buenas prácticas para la mayoría de los proyectos, además, brinda aspectos que deben conocerse al momento de realizarlos. Contiene diez distintas áreas de conocimiento, las cuales pueden ser implementadas parcial o totalmente, según la necesidad del proyecto (Hillson, 2017). Si bien es una guía de ejecución, no es totalmente cerrada, pues no todos los proyectos son iguales. En términos del presente trabajo, esta se tomó como marco de referencia para ejecutar el plan de implementación, ya que es considerado un proyecto dentro del equipo de Equifax.

Todos estos conceptos son necesarios para una mejor comprensión y realización del procedimiento metodológico mostrado en el siguiente capítulo. Este guía las fases para construir la solución.

Capítulo 3. Marco metodológico

En este capítulo, se desarrollan todos los elementos necesarios contemplados dentro del marco metodológico para el trabajo final de graduación. Se exponen aspectos como el tipo, enfoque, alcance y diseño de la investigación, fuentes de información, sujetos de investigación, muestras, variables, e instrumentos utilizados para completar el proyecto. Asimismo, se define el procedimiento metodológico utilizado para guiar su desarrollo, con una especificación de cada una de sus fases.

3.1. Tipo de investigación

Para la ejecución de este proyecto, es necesario establecer una metodología que permita ajustar las actividades a su objetivo. Esta investigación es de tipo científica, la cual Hernández Sampieri y Mendoza (2018) definen como “el conjunto de procesos sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno o problema con el resultado (u objetivo) de ampliar su conocimiento” (p. 4). Este tipo de trabajos permite llevar a cabo una serie de fases para obtener el resultado final definido en el objetivo del proyecto.

Según Hernández Sampieri y Mendoza (2018), las investigaciones se pueden clasificar en tres enfoques: cualitativo, cuantitativo y mixto. De acuerdo con lo que mencionan los autores en su libro, el cuantitativo “pretende describir, explicar y predecir los fenómenos investigados, buscando regularidades y relaciones causales entre elementos. Significa que la meta principal es la prueba de hipótesis y la formulación y demostración de teorías” (pp. 5-7). Por otro lado, el enfoque cualitativo “se basa en métodos de recolección de datos no estandarizados al inicio ni completamente predeterminados. Los datos cualitativos consisten fundamentalmente en narrativas de diferentes clases” (p.9). Esto quiere decir que es posible utilizar herramientas verbales o visuales como un apoyo para la investigación. Finalmente, el enfoque mixto es descrito como la “integración sistemática de los métodos cuantitativo y cualitativo en un solo estudio, con el propósito de obtener una ‘fotografía’ más completa del fenómeno” (p.10). Con este último, se busca sacar provecho a lo mejor de los dos enfoques anteriores.

Por otro lado, se encuentra el enfoque alternativo, propuesto por Luis Naranjo (2021), el cual se caracteriza por que “No se necesita definir qué tan cuantitativo o cualitativo es el proyecto [...] y se puede realizar combinaciones en el diseño de la investigación” (Canal Yarima Sandoval, 2022, 47m32s). Así, busca facilitar las investigaciones, sin limitarlas a un diseño en específico. Esto debido a que se basa en un paradigma pragmático.

Una vez conocidos los distintos tipos de investigación y sus características, se determina que el **enfoque cualitativo** es el pertinente la presente investigación. Esto debido a la necesidad de realizar un planteamiento abierto, fundamentado en experiencias. Este proyecto busca aprender de las experiencias de los desarrolladores y encargado del equipo, para conocer el proceso actual, las necesidades y deficiencias de este y desarrollar un proceso que permita una mejora en el departamento de Servicios de Datos. La razón por la cual se descartan los enfoques cuantitativos y mixto es porque esta investigación no busca realizar un análisis estadístico o numérico ni trabajar

únicamente con datos cuantitativos, lo cual deja de lado la amplia selección de diseños que brinda el enfoque alternativo.

3.2. Alcance de la investigación

El alcance de una investigación se mantiene en constante cambio, ya que se ajusta al contexto en el cual se realiza. Este aspecto permite establecer los límites conceptuales y metodológicos. Existen cuatro tipos de alcances, los cuales se describen en la **Tabla 6**.

Tabla 6. *Tipos de Alcance*

Tipo	Descripción
Exploratorio	<ul style="list-style-type: none">• Se realiza cuando el objetivo es examinar un tema o problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes.• Sirve para familiarizarse con fenómenos relativamente desconocidos y obtener información sobre la posibilidad de llevar a cabo una investigación más completa.
Descriptivo	<ul style="list-style-type: none">• Busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis.
Correlacional	<ul style="list-style-type: none">• Tiene como finalidad conocer la relación o grado de asociación que exista entre dos o más conceptos, categorías o variables en una muestra o contexto en particular.
Explicativo	<ul style="list-style-type: none">• Pretende establecer las causas de los sucesos o fenómenos que se estudian.

Fuente: adaptado de *Metodología de la investigación*, por Hernández Sampieri y Mendoza (2018).

Para la presente investigación, se selecciona el alcance de tipo **descriptivo**. Esto porque se busca especificar las propiedades, características y los perfiles de personas, grupos, comunidades, procesos, objetos, o cualquier otro fenómeno que se someta a un análisis. Los análisis descriptivos son sumamente funcionales para demostrar con precisión las dimensiones de un fenómeno, suceso, comunidad, contexto o situación. Por lo cual, se debe definir qué se medirá y sobre qué se recolectarán los datos. Por medio de este alcance, en el presente proyecto, se identifican aspectos relacionados con las actividades claves y con qué herramientas se va a extraer la información que sirva de insumo en el desarrollo de los reportes de los grupos Business Care y Boarding.

3.3. Diseño de la investigación

A partir del tipo de enfoque seleccionado, en este caso cualitativo, es necesario señalar el diseño para desarrollar el presente proyecto. Según Hernández Sampieri y Mendoza (2018), el diseño de la investigación “se refiere al plan o estrategia concebida para obtener la información que deseas con el propósito de responder al planteamiento del problema” (p. 150). Para la elección de diseño, se evaluaron los distintos tipos planteados para la investigación cualitativa en el libro de Hernández Sampieri y Mendoza, los cuales se encuentran en la **Tabla 7**.

Tabla 7. *Tipos de diseños para la investigación cualitativa*

Diseño	Descripción
Teoría fundamentada	Su propósito es inducir una teoría basada en datos empíricos; se aplica a áreas específicas. El investigador produce una explicación o teoría respecto a un fenómeno, proceso, acción o interacciones, que se aplica a un contexto concreto y desde la perspectiva de diversos participantes.
Etnográfico.	Pretende explorar, examinar y entender sistemas sociales (grupos, comunidades, culturas y sociedades), así como producir interpretaciones profundas y significados culturales, desde la perspectiva o punto de vista de los participantes o nativos.
Narrativo	Busca entender la sucesión de hechos, situaciones, fenómenos, procesos y eventos donde se involucran pensamientos, sentimientos, emociones e interacciones, a través de las vivencias contadas por quienes los experimentaron.
Fenomenológico	Su propósito principal es explorar, describir y comprender las experiencias de las personas con respecto a un fenómeno, y descubrir los elementos en común de tales vivencias.
Investigación-acción	Su objetivo es comprender y resolver problemáticas específicas de una colectividad, vinculadas a un ambiente (grupo, programa, organización o comunidad). Se centra en aportar información que guíe la toma de decisiones para proyectos, procesos y reformas estructurales.

Fuente: adaptado de *Metodología de la investigación*, por Hernández Sampieri y Mendoza (2018).

Por la naturaleza de este proyecto, se selecciona como diseño la investigación-acción. Este se basa en las fases cíclicas de actuar, pensar y observar. Sus perspectivas son: visión técnico-científica, visión deliberative y visión emancipadora. Por otro lado, se encuentran las fases, las cuales son tres: observar (construir un bosquejo del problema y recolectar datos), pensar (analizar e interpretar) y actuar (resolver problemáticas e implementar mejoras). Estas se dan de manera cíclica, hasta que todo quede resuelto. Así, el cambio se logra o la mejora se introduce satisfactoriamente.

3.4. Fuentes de datos e información

Al realizar una investigación, se debe recolectar distintas fuentes de información, las cuales tendrán un rol importante. De acuerdo con Ulate y Vargas (2014), estas son aquellas fuentes consultadas a lo largo del proyecto; se clasifican en primarias, secundarias y terciarias. En esta sección, se enlistan las fuentes de información que fueron utilizadas, se brinda una clasificación por tipos y se conceptualizan de acuerdo con su importancia en el proyecto.

3.4.1. Fuentes primarias

De acuerdo con Monje (2011), las fuentes son consideradas primarias “cuando se trata de un escrito personal referente a las propias experiencias, investigaciones y resultados” (p. 74). Estas permiten obtener criterios y parámetros para las investigaciones con base a trabajo realizados en el pasado. En la **Tabla 8**, se encuentran las fuentes primarias utilizadas para el presente proyecto.

Tabla 8. Fuentes primarias

Fuente	Relevancia
Encargado del departamento	Esta persona dirige y coordina el grupo. Brinda la información necesaria para recolectar las necesidades y limitaciones que presenta el equipo.
Encargado de creación de reportes	Él ejecuta los reportes y conoce el proceso actual, así como las distintas fuentes de información.
Libro de marca de la organización	Este contiene los lineamientos necesarios para el uso de la tipografía oficial y aspectos visuales asociados a la imagen de Equifax.
PMBOK	Este aporta las etapas necesarias para realizar una implementación.
Fundamentos de la gestión de procesos de negocio	Estos proporcionan etapas, conceptualización y desarrollo de proyectos relacionados con la mejora de procesos.

Fuente: elaboración propia.

3.4.2. Fuentes secundarias

Con respecto a estas fuentes, Monje (2011) explica que “El autor obtiene la información a través de otras personas, libro o material”. En la **Tabla 9**, se encuentran las fuentes secundarias utilizadas para la presente investigación.

Tabla 9. Fuentes secundarias

Fuente	Relevancia
Repositorio de anteproyectos del ITCR	Los anteproyectos sirven como referencia para elaborar la investigación y brindan una guía para el desarrollo de la solución.
Sitios web con información sobre inteligencia de negocios	Estos sitios brindan información técnica relacionada con el área que se desea mejorar.
Libros y revistas sobre inteligencia de negocios y administración de procesos de negocio	En estos, se presentan procesos metodológicos y etapas necesarias para el desarrollo de soluciones o proyectos relacionados.
Sitio de Equifax	Acá se encuentra información pública relevante para el proyecto.

Fuente: elaboración propia.

3.5. Sujetos de investigación

En esta sección se describirán los sujetos claves de este proyecto. Este apartado permite la recolección de datos, los cuales sirven como insumo para el análisis de la situación actual. En la **Tabla 10**, se definen los sujetos de investigación y se justifica su relevancia dentro del proyecto.

Tabla 10. *Sujetos de investigación: rol y relevancia*

Rol	Descripción del rol	Relevancia
Gerente del departamento	Esta persona se encarga de dirigir el día a día de la organización, así como de comunicar la información entre la alta gerencia y el resto del personal.	Toma las decisiones gerenciales del equipo. Conoce los procesos de negocio. Cuenta con una visión completa del negocio.
Creadores de reportes	Esta persona realiza diariamente los reportes.	Conoce el proceso de realizar reportes. Conoce las limitaciones y posibles mejoras. Conoce las fuentes de datos que utilizan los distintos reportes.
Encargado de soporte de software	Esta persona gestiona los softwares utilizados dentro del equipo y revisa diariamente el estado de los ETL.	Provee información sobre herramientas disponibles, alcance y limitaciones que presentan estas. Brinda herramientas que son permitidas dentro de la organización.
Clientes finales.	Estas personas son los consumidores de los reportes generados por el equipo.	Proveen los requerimientos de los reportes finales y las métricas que estos deben contener.

Fuente: elaboración propia.

3.6. Variables o categorías de la investigación

En este apartado, se detallan las variables de la investigación asociadas a cada uno de los objetivos específicos definidos en el presente proyecto. Según Carlos Mata (2022), las variables son “características que no se dan igual en los distintos objetos o unidades que cubren el estudio” (Canal Yarima Sandoval, 2022, 5m10s). Son la síntesis de lo que se pretende descubrir con la investigación. En la **Tabla 11**, se encuentran tanto las variables identificadas en relación con cada objetivo específico, como su relevancia para el proyecto.

Tabla 11. Variables de la investigación identificadas

Objetivo	Variables	Relevancia
Analizar el estado actual del proceso de creación de reportes para la identificación de necesidades del equipo y el modelado del nuevo proceso.	Situación actual del proceso de creación de reportes.	Determinar el contexto actual de la organización, el equipo y las limitaciones de este.
	Deficiencias existentes en el proceso actual.	Identificar las necesidades y requisitos que deben ser cumplidos.
Diseñar un nuevo proceso y su respectiva documentación para la extracción, transformación y carga de datos, y realización de reportes.	Fuentes de datos utilizadas.	Documentar las distintas fuentes de datos existentes y utilizadas para el nuevo proceso estandarizado y automatizado de creación de reportes.
	Estandarización del proceso automatizado.	
Diseñar pruebas de concepto del proceso anterior para su validación.	Funcionalidad del rediseño propuesto.	Permite la validación con los principales involucrados del proceso.
Desarrollar un plan de implementación para la automatización del proceso de creación de reportes.	Plan de implementación para aplicar el modelo.	Permite brindar a la organización un camino a seguir para implementar el nuevo proceso.
Desarrollar un análisis de costo-beneficio para la valoración de viabilidad del plan de implementación.	Costo-beneficio.	Determinar los costos y beneficios en términos económicos que conlleva esta implementación.

Fuente: elaboración propia.

3.7. Técnicas e instrumentos de recolección de datos

En este apartado, se explican los instrumentos utilizados para recolectar los datos que se requieren para elaborar el proyecto. De acuerdo con Monje (2011), en esta sección se debe “determinar la utilización racional de los procedimientos, técnicas e instrumentos de investigación, su implementación y uso” (p. 132). Para esto, es fundamental contar con el entendimiento de las técnicas existentes para realizar una investigación. De acuerdo con Delgado, Paitán, Vilela y Dueñas et al. (2018), las técnicas se pueden definir como “un conjunto de normas que regulan el proceso de investigación, en cada etapa, desde el principio hasta el fin”. En la **Tabla 12**, se encuentran las técnicas definidas por dicho autor.

Tabla 12. *Técnicas de recolección de datos*

Técnica	Descripción
Conceptual	Son procedimientos mentales y reglas lógicas que se utilizan en todo el proceso de la investigación. Refieren a las técnicas que hacen posible las operaciones racionales, de abstracción, generalización, análisis, síntesis, clasificación, comparación; y a las reglas lógico-formales y lógico-dialécticas necesarias en la investigación científica.
Descriptivas	Sirven para la recolección de datos y para la verificación de las hipótesis. Entre estas se encuentran la observación, la entrevista, la prueba, el análisis de contenido, etc.
Cuantitativas	Son aquellas que se refieren a magnitudes o cantidades y se expresan mediante números, fórmulas, algoritmos numéricos.

Fuente: adaptado de *Metodología de la investigación: cuantitativa-cualitativa y redacción de la tesis*, por Delgado, Dueñas, Paitán y Vilela (2018).

Una vez comprendidas las distintas técnicas, se destaca la utilizada para el presente proyecto, la cual es la **descriptiva**. Esto se debe a que la presente investigación recolecta datos a través de experiencias de usuarios, puntos de vista de investigadores, reconstrucciones de hechos y documentaciones.

Por otro lado, se encuentran los instrumentos de recolección de datos, los cuales son esenciales para llevar a cabo el proyecto. Los utilizados en este trabajo son:

- **Revisión documental**

La revisión documental es uno de los instrumentos de recolección de datos que permite consultar textos, materiales y artefactos diversos, con el propósito de entender el fenómeno central de estudio. De acuerdo con Delgado (2018), “En esta fase exploratoria la información básica se va incrementando, el investigador va logrando un conocimiento cada vez más cabal sobre la idea o tema que le interesa estudiar” (p. 390). Para este proyecto, se realiza una revisión artículos científicos y académicos, y sitios web con el fin identificar proyectos similares en los cuales se realizaron mejoras de procesos relacionados con la inteligencia de negocios. Además, se revisan las herramientas para determinar costos, alcance y factibilidad de uso. La plantilla utilizada para este instrumento se encuentra en el **Apéndice C. Plantilla de recopilación de datos en revisión documental**

- **Observación**

La observación como método de recolección de datos permite comprender el comportamiento y experiencias de un grupo de personas en un determinado ambiente (Monje, 2011). Asimismo, facilita comprender los problemas que enfrentan en el día a día las personas encargadas de realizar una determinada actividad. En este proyecto se lleva a cabo una observación participativa, cuyo objetivo es “observar y registrar información de las personas en sus medios con un mínimo de estructuras y sin inferencia del investigador” (Monje, 2011). Con este instrumento, se recolecta información sobre el proceso de creación de reportes, sus limitaciones y las necesidades de los usuarios. En el **Apéndice D. Plantilla de recopilación de datos en la observación** se encuentra el aplicado durante este trabajo.

- **Entrevistas**

Las entrevistas se definen como “una reunión para conversar e intercambiar información entre una persona (el entrevistador) y otra (el entrevistado) u otras (entrevistados)” (Hernández Sampieri y Mendoza, 2018, p. 449). Este tipo de instrumento es útil para recolectar opiniones detalladas, pensamientos y experiencias, debido a que las preguntas suelen ser abiertas y neutrales. En un enfoque cualitativo, existen tres tipos de entrevistas, las cuales se encuentran definidas en la **Figura 6**.

Figura 6. *Tipos de entrevistas*



Fuente: adaptado de *Metodología de la investigación*, por Hernández Sampieri y Mendoza (2018).

Para este proyecto, se realizaron **entrevistas semiestructuradas**. Estas permitieron evaluar la situación actual de la problemática identificada. Se aplicaron al encargado y personal del equipo de Servicios de Datos. Las plantillas utilizadas para estos instrumentos se encuentran en el **Apéndice E. Plantilla de entrevista**

3.8. Matriz de cobertura de las variables

A continuación, en la **Tabla 13**, se asocia cada uno de los instrumentos de recolección de datos con las variables identificadas a partir de los objetivos específicos del proyecto.

Tabla 13. *Matriz de cobertura de variables*

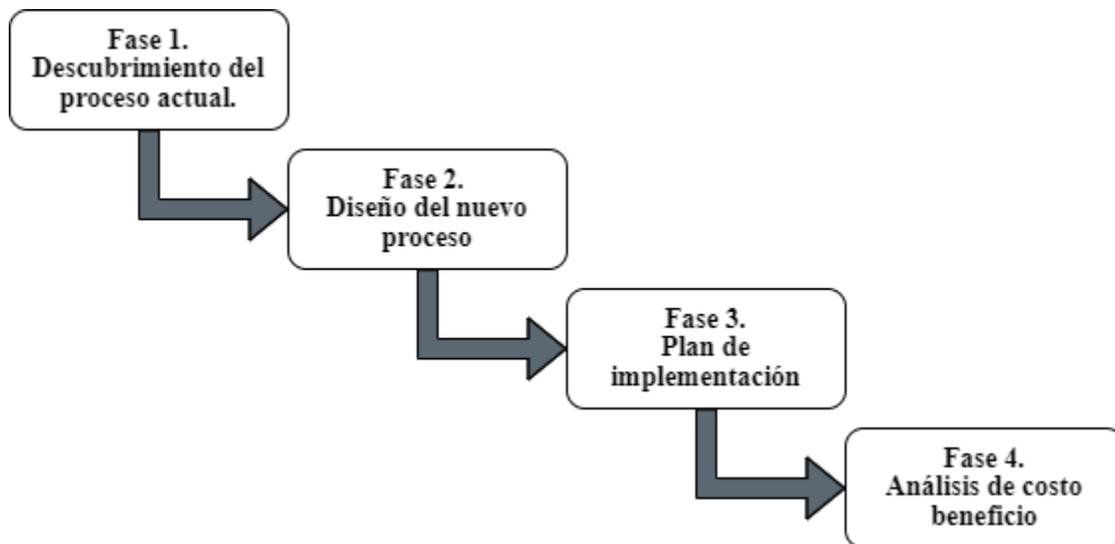
Variable	Técnica		
	Revisión Documental	Entrevista	Observación
Situación actual del proceso de creación de reportes.		X	X
Necesidades presentes en el proceso.		X	X
Documentación del nuevo modelo de extracción, transformación y carga de datos.		Propuesta	
Pruebas de concepto del modelo propuesto.	X	X	X
Plan de implementación para la aplicación del modelo.	X		x
Análisis de costo-beneficio para la valoración del modelo.		Propuesta	

Fuente: elaboración propia.

3.9. Procedimiento metodológico de la investigación

En esta sección, se desglosa y explica el procedimiento utilizado para llevar a cabo el presente proyecto. Este se encuentra dividido en cuatro fases que involucraron el proceso metodológico para cumplir los objetivos establecidos, las cuales se representan en la **Figura 7**. En cada una de estas, se establecieron tareas, las cuales se describen a continuación.

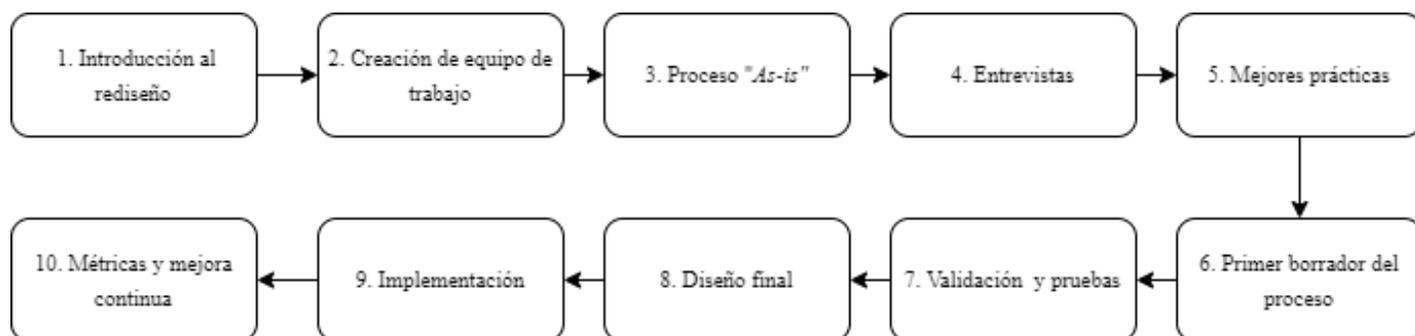
Figura 7. Fases del proceso metodológico



Fuente: elaboración propia.

Al ser un proyecto de mejora de procesos, se pueden emplear dos metodologías muy conocidas: la de Dumas y la propuesta por Madison (descritas en apartado **2.2.1 más atrás**). Si bien es cierto, ambas son funcionales en términos de rediseño y mejora de procesos, cabe destacar que la de Madison se enfoca en el rediseño de un proceso aislado. Por su lado, la metodología propuesta por Dumas y compañía realiza una serie de pasos de identificación, análisis y mejora, lo cual permite llevar a cabo un análisis completo de los todos los procesos, eventos y subtareas de estos. Por esta razón, para para las primeras dos fases de este proyecto, se emplea una adaptación de la metodología propuesta por Madison. Tal adaptación consiste en tomar desde la segunda hasta la octava etapa de la mejora de procesos, debido al tiempo disponible para la ejecución y el alcance del proyecto. De acuerdo con Madison (2005), se deben realizar diez actividades claves para rediseñar un proceso, estos pasos se observan en la **Figura 8**.

Figura 8. Metodología de rediseño según Madison



Fuente: adaptado de *Process Mapping, Process Improvement and Process Management: A Practical Guide to Enhancing Workflow and Information Flow*, por Madison (2005).

3.9.1. Fase 1. Descubrimiento del proceso actual

En esta fase, se trabaja el objetivo específico uno, es decir, se busca comprender la situación actual del proceso de generación de reportes dentro del equipo de Servicios de Datos, para identificar las necesidades presentes. Para cumplirlo, se ejecutan los pasos tres y cuatro de Madison (2005). Por medio de estos, se conforma el equipo de involucrados e interesados del proyecto y a su vez se comprende el estado actual del proceso, el cual se describe ampliamente en el capítulo #1, y las necesidades que presenta. Las tareas realizadas durante esta fase se desarrollan a continuación.

3.9.1.1. Identificación del estado actual

En esta actividad, se realiza un diagramado del proceso *As-is* y corresponde al paso tres establecido por Madison en su metodología. Así, se lleva a cabo un descubrimiento del proceso actual y se cumple con el primer objetivo específico. Según Madison (2005), este paso es previo a las entrevistas con los involucrados, pues durante el proceso pueden surgir preguntas, que pueden ser de mucha utilidad para el siguiente paso.

Para generar el diagrama de esta actividad, es necesario tomar en cuenta ciertas consideraciones. De acuerdo con Madison (2005), existen cuatro tipos de enfoques o razones por las cuales se busca la mejora o rediseño de un proceso: frustración, tiempo, costos y calidad. Tomando estos y comparándolos con el contexto del presente proyecto, se considera que se presentan el de **tiempo y el de frustración**, según explica Vindas (2022) en la entrevista que se le realiza:

- **Lente del tiempo:** “El tiempo es una dimensión crítica de la satisfacción del cliente y la reducción de costos. Llevar productos y servicios a los clientes de manera rápida es una actividad que agrega valor” (Madison, 2005, p. 91). Este es uno de los principales “lentes” o razones para realizar el presente proyecto. Se desean resolver los problemas con tiempos de ejecución y de entrega.

- **Lente de frustración:** “El lente de la frustración diagnostica el proceso desde la perspectiva de quienes trabajan en él. El propósito es aprender qué frustraciones experimentan las personas al hacer su trabajo” (Madison, 2005, p. 85). Según obtenido en las entrevistas con Vindas (2022) y Alfaro (2022), la necesidad de estar realizando tareas repetitivas que toman tiempo y los problemas debido a la poca integridad de datos causa frustración al realizar estos reportes.

Para este proceso, se utilizan las herramientas mencionadas en el apartado de **Técnicas e instrumentos de recolección de datos**, con el fin de establecer el estado actual de la realidad de la empresa.

3.9.1.2. Entrevistar a los involucrados

Durante esta actividad, se entrevista a los distintos involucrados dentro del proceso. Esto corresponde al paso cuatro de la metodología de Madison (2005), un subproducto de esta fase son los requerimientos necesarios, ya que, en este paso, el objetivo es averiguar qué necesita y desea el cliente. Se realizan las preguntas concebidas a partir del diagrama *As-is* en el paso tres, este se genera por medio de la notación BPMN, la cual se encuentra descrita en apartado **2.2.2 del Marco conceptual**. Para esta actividad, se utilizaron técnicas de recolección como los grupos focales, entrevistas y observaciones.

3.9.2. Fase 2. Diseño del nuevo proceso

En esta fase, se abordan los objetivos dos y tres del presente proyecto, los cuales abarcan diseñar una mejora del proceso. Para esto, se realiza un estudio de las mejores prácticas del mercado para la gestión y diseño del proceso propuesto. Con respecto a la metodología de mejora de Madison, se toman los pasos cinco, seis y ocho expuestas en la **Figura 8**. En esta fase, se realiza el diagramado del *To-be*, el cual utiliza la misma notación BPMN descrita en apartado **2.2.2**. Para esto, se realizó un análisis de las herramientas y de los datos disponibles para la elaboración del nuevo proceso. Una vez se obtuvo el nuevo modelo y su aprobación, se llevaron a cabo las pruebas de concepto, con el fin de validar el funcionamiento del proceso propuesto.

3.9.2.1. Evaluación de mejores prácticas

En esta actividad, se revisan las mejores prácticas del mercado de procesos relacionados tanto con inteligencia de negocios como con gestión de la información. El objetivo es que sirvan como guía para el rediseño del proceso. Aquí se extrae la información más relevante de este proyecto, para destacar aspectos claves que serán utilizados como insumos en las actividades siguientes.

3.9.2.2. Identificación y documentación de las herramientas para los procesos de extracción, carga y transformación de datos

Para cumplir con este apartado, se entrevista al gerente del departamento, así se conocen las herramientas disponibles dentro de la organización que permiten mejorar el proceso de creación de reportes. Por otro lado, se identifican los datos disponibles, con el fin de documentar la información encontrada y utilizarla como insumo en los entregables.

3.9.2.3. Realizar primer rediseño

En esta actividad, se hace una primera versión del rediseño del proceso. De este modo, se validan todos los requerimientos y conocimientos obtenidos en las actividades pasadas. Se utilizan como insumo las entrevistas, requerimientos y buenas prácticas para desarrollar un proceso más certero. De acuerdo con Madison (2005), en esta actividad se pueden plantear más de un diseño, los cuales serán evaluados, para depurar el modelo hasta obtener una última versión. La finalidad es tener un primer diseño que sirva como guía para la solución de la problemática.

3.9.2.4. Validación del nuevo diseño con involucrados y pruebas de concepto

Una vez que se depuran los diseños realizados, son validados por los principales involucrados. Esta validación se acompaña de unas pruebas de concepto, que se hace con una adaptación de la guía propuesta por Intel para aplicar pruebas de concepto para inteligencia artificial; la cual cuenta con cinco pasos principales: confirmar la oportunidad, caracterizar el problema, desarrollar la solución, evaluación para valor del negocio y escalar el PoC. Para el presente proyecto se utilizan los primeros cuatro pasos.

Debido al contexto, estas pruebas se realizan por medio del principio de Pareto (ver 2.3.3 **El principio de Pareto**), donde se emplean las fuentes de datos presentes en una mayor cantidad de reportes. Posteriormente, se pasa de la teoría a la práctica: se llevan a cabo estas pruebas utilizando las fuentes de datos de prueba que representan datos similares a los reales (por motivos de confidencialidad) y Python como principal lenguaje de programación.

3.9.2.5. Realizar versión final del nuevo proceso

Esta tarea corresponde al paso ocho de la metodología de Madison. Una vez depurados los detalles encontrados en las fases previas, se presenta la última versión del proceso rediseñado. De acuerdo con Madison (2005), el diseño final debe compartirse con todos los involucrados dentro del proceso, para validar su efectividad y verificar que satisface las necesidades y da solución a la problemática. Entre los principales resultados de esta actividad están el modelo *To-be* y la documentación del proceso, la cual respondería al objetivo específico dos de este trabajo.

3.9.3. Fase 3. Plan de implementación

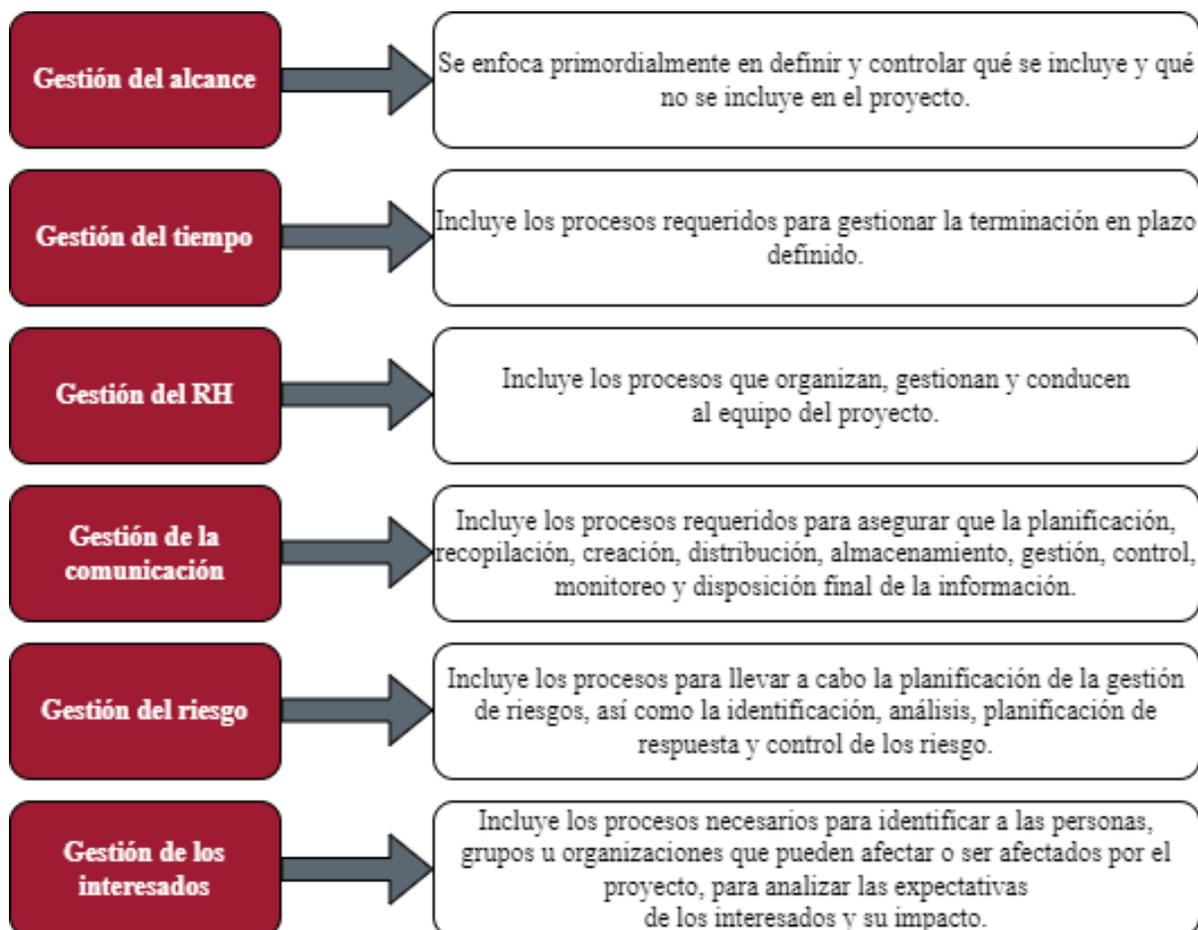
Esta fase comprende el objetivo cuatro del proyecto, en el cual se desarrolla un plan de implementación. Para este, se utilizan como referencia las etapas establecidas por el PMI en su guía del PMBOK. No obstante, se adaptan a la guía y a las distintas gestiones que establece esta

organización. Las gestiones del PMBOK adoptadas para el desarrollo del plan de implementación se definen en la **Figura 9**. Fases del PMBOK utilizadas.

Para ejecutar el plan de implementación, se define el alcance por medio de los requerimientos previamente recolectados en la **Fase 1. Descubrimiento del proceso actual**. Con esta información, se establecen las limitaciones y el objetivo del presente plan. Después, se realiza un cronograma de implementación, donde se plantea la duración de las actividades en horas, debido a los procesos burocráticos de la empresa que dificultan tener certeza en cuanto a las fechas de implementación.

En esta sección del plan, se definen tanto los tiempos como las actividades necesarias. A continuación, se establece la sección de colaboradores, sus roles, responsabilidades y habilidades requeridas para la implementación del proyecto. Finalmente, se define la sección de controles, donde se determinan las gestiones de comunicación, los riesgos e interesados.

Figura 9. Fases del PMBOK utilizadas



Fuente: adaptación de *Guía de los fundamentos para la dirección de proyectos*, por Project Management Institution. (2014)

3.9.4. Fase 4. Análisis de costo-beneficio

En esta fase, se abarca el quinto objetivos específico del proyecto. Contando con el plan de implementación, se evalúa si es factible para la organización en términos financieros. Para esto, se analiza el costo-beneficio, el cual “es un proceso sumamente útil para determinar los beneficios económicos que pueden derivar de una decisión y determinar si vale la pena avanzar con esa opción” MacNeil (2022). Esta herramienta brinda una perspectiva cuantitativa del problema, la cual permite tomar una decisión. Dentro de las actividades realizadas, se encuentran identificar costos asociados, así como determinar tanto de los beneficios financieros como de los que no lo son. A continuación, se describen detalladamente.

- Identificación de costos asociados

Acá se determinan los costos asociados con la propuesta de solución, mediante una revisión documental. Esta permite conocer los aspectos que deben considerarse al realizar una cotización.

- Determinación de beneficios financieros y no financieros

Con el instrumento de revisión documental, se identifican los beneficios tanto financieros como no financieros de la propuesta planteada por el investigador. Los primeros determinan la viabilidad de la solución.

3.10. Operacionalización de las variables o categorías

En este apartado, se detalla la operacionalización de variables y se asocian los componentes metodológicos principales del proyecto (ver **Tabla 14**); esto incluye las fases de la investigación, los objetivos específicos, los instrumentos, las variables y los sujetos de investigación.

Tabla 14. *Operacionalización de las variables*

Fase	Objetivo específico	Variable	Sujeto	Instrumento
1. Descubrimiento del proceso actual.	Analizar el estado actual del proceso de creación de reportes para la identificación de necesidades del equipo y el modelado del nuevo proceso.	Situación actual del proceso de creación de reportes. Deficiencias existentes en el proceso actual.	Encargado de generar reportes.	Apéndice D. Plantilla de recopilación de datos en la observación
			Gerente del departamento.	Apéndice E. Plantilla de entrevista
			Encargado de soporte. Clientes finales.	
2. Diseño del nuevo proceso.	Diseñar un nuevo proceso de extracción, transformación y carga de datos (ETL), su respectiva documentación y visualización de reportes para ser evaluados por medio pruebas de concepto.	Fuentes de datos utilizadas. Estandarización del proceso automatizado. Funcionalidad del rediseño propuesto.	Gerente del departamento.	Apéndice C. Plantilla de recopilación de datos en revisión documental
			Encargado de generar reportes.	Apéndice E. Plantilla de entrevista
3. Plan de implementación	Desarrollar un plan de implementación para la automatización del proceso de creación de reportes y su respectivo análisis de costo-beneficio para la valoración.	Plan de implementación para la aplicación del modelo.	Gerente del departamento.	.Apéndice C. Plantilla de recopilación de datos en revisión documental Apéndice E. Plantilla de entrevista
4. Análisis de viabilidad.		Costo-beneficio.	Gerente del departamento.	Apéndice C. Plantilla de recopilación de datos en revisión documental Apéndice E. Plantilla de entrevista

Fuente: elaboración propia.

3.11. **Tabla resumen del procedimiento metodológico o trazabilidad**

Finalmente, en la **Tabla 15**, se muestra la relación entre los distintos objetivos específicos de este proyecto, el marco teórico, la metodología, el análisis de resultados, las conclusiones y las recomendaciones.

Tabla 15. Matriz de trazabilidad

Objetivo específico	Marco conceptual	Metodología	Análisis y propuesta	Conclusiones	Recomendaciones
Analizar el estado actual del proceso de creación de reportes para la identificación de necesidades del equipo y el modelado del nuevo proceso.	Sección 2.1.1	Sección 3.1	Sección 4.1.1	Sección 6.1	
	Sección 2.1.2	Sección 3.2	Sección 4.1.2		
	Sección 2.1.3	Sección 3.3			
	Sección 2.1.8	Sección 3.4			
		Sección 3.5			
Diseñar un nuevo proceso y su respectiva documentación para la extracción, transformación y carga de datos y realización de reportes.	Sección 2.1.6	Sección 3.5	Sección 4.2.1	Sección 6.2	
	Sección 2.1.7	Sección 3.6	Sección 4.2.2		
	Sección 2.1.10	Sección 3.7	Sección Capítulo 5		
	Sección 2.1.11	Sección 3.9.2			
Diseñar pruebas de concepto del proceso anterior para su validación.	Sección 2.3.2	Sección 3.6	Sección 4.2.3	Sección 6.3	Sección Capítulo 7
	Sección 2.3.3	Sección 3.7	Sección Capítulo 5		
		Sección 3.9.2			
Desarrollar un plan de implementación para la automatización del proceso de creación de reportes.	Sección 2.4	Sección 3.7	Sección 4.3	Sección ¡Error! No se encuentra el origen de la referencia.	
	Sección 2.4.1	Sección 3.9.3	Sección 5.2		
Desarrollar un análisis de costo-beneficio para la valoración de viabilidad del plan de implementación.	Sección 2.3.1	Sección 3.7	Sección 4.4	Sección ¡Error! No se encuentra el origen de la referencia.	
		Sección 3.9.4	Sección 5.3		

Fuente: elaboración propia.

Capítulo 4. Análisis de resultados

En este capítulo, se lleva a cabo el análisis de los resultados obtenidos al aplicar los instrumentos de recolección de datos seleccionados para las fases 1 y del proceso metodológico, que se evidencian en la sección **Marco metodológico**. Se emplearon entrevistas, observaciones y revisiones documentales, con el propósito de identificar las oportunidades de mejora y hacer una propuesta de solución adecuada.

Las técnicas utilizadas para el análisis de resultados de las fases 1 y 2 son las siguientes:

- Entrevista al gerente del departamento de Servicios de Datos.
- Entrevista a los desarrolladores de inteligencia de negocios.
- Entrevista al encargado de brindar soporte a las herramientas del equipo de Servicios de Datos.
- Entrevista a los clientes que consumen los reportes generados por el equipo.
- Revisión documental sobre buenas prácticas de procesos.
- Revisión documental sobre buenas prácticas de inteligencia de negocios.
- Proceso de observación de los distintos pasos dentro del proceso de creación de reportes.

4.1. Primera fase: Descubrimiento del proceso actual

Como se mencionó en el capítulo anterior, el objetivo de esta fase es descubrir el estado actual del proceso de creación de reportes y recopilar los requerimientos que se tienen para este, esto permite identificar posibles soluciones para mejorar el proceso estudiado, y disminuir así los efectos que produce a la operación del negocio.

4.1.1. Identificación del estado actual

Para identificar el proceso actual de creación de reportes, se utiliza la observación, durante dos días distintos. Así se conoce a fondo cada una de las actividades realizadas y a los participantes durante el proceso. En la **Tabla 16**, se evidencian tanto los participantes identificados como la descripción de su actividad dentro del proceso actual de crear reportes.

Tabla 16. *Participantes del proceso*

Participante	Descripción
Desarrollador de inteligencia de negocios	<ol style="list-style-type: none"> 1. Extrae la información de las fuentes de datos. 2. Realiza un proceso de integración de datos. 3. Actualiza reportes. 4. Notifica a los clientes.
Cliente	Encargado de revisar y consumir los reportes generados.

Fuente: elaboración propia.

Las principales actividades para realizar los reportes para el departamento de Business Care y Boarding, así como su descripción y los participantes, se detallan en la **Tabla 17**.

Tabla 17. *Actividades del proceso*

Actividad	Participante	Descripción
Extracción de datos	Desarrollador de inteligencia de negocios	Se extraen los datos de las distintas fuentes. Por lo general, son descargados al pc local en archivos planos.
Agrupar información	Desarrollador de inteligencia de negocios	Se agrupa la información de cada uno de los archivos en un mismo documento (Excel) utilizando distintas pestañas.
Transformación de datos	Desarrollador de inteligencia de negocios	Se procesan datos manuales. El encargado utiliza funciones disponibles en Excel para integrar los datos y hacer cálculos de tiempos. Elimina a mano datos que no deben ser tomados en cuenta.
Actualizar fuentes de datos asociadas a los reportes	Desarrollador de inteligencia de negocios	Se actualizan las fuentes de datos que funcionan como base de datos para los reportes. Estas en su mayoría son hojas de cálculo de Google.
Actualizar reportes	Desarrollador de inteligencia de negocios	Existen distintos tipos de reportes, pero el procedimiento es el mismo. Si es un reporte en una herramienta como Tableau o Data Studio, se actualiza manualmente la fuente; en el caso de reportes en hojas de cálculo de Google, se realiza una nueva hoja con los nuevos valores.
Entrega de reportes	Desarrollador de inteligencia de negocios	Vía mail (Gmail) se notifica a los clientes que ya han sido actualizados los datos y se adjuntan los enlaces.
Revisión de reportes	Cliente	El cliente hace una revisión de los reportes. Si existen algún problema, envía la información encontrada al encargado para que se realice los cambios.

Fuente: elaboración propia.

Además, se identifican las fuentes de datos, cómo los encargados extraen estos y las herramientas utilizadas dentro del proceso de entregas de los reportes. En la **Tabla 18**, se detallan.

Tabla 18. Fuentes de datos utilizadas dentro del proceso

Fuente	Descripción	Método de extracción
Salesforce	Contiene información de los casos generados para el proceso de venta y gestión de clientes.	Archivos planos (CSV o Excel)
Genesis (Qualtrics)	Cuenta con la información de los datos generados por las encuestas enviadas por los agentes del centro de llamadas.	Hoja de cálculos de Google
AOdocs	Contiene información de un proceso realizado por el departamento, el cual busca actualizar los datos y permisos de acceso de los consumidores de sus productos.	Hoja de cálculos de Google
Looker	Cuenta con información correspondiente a las consultas y solicitudes de información realizadas por los clientes de la organización.	Archivos planos (CSV o Excel)
Gmail	Medio para compartir, en correos encriptados, información perteneciente a la lista de clientes activos y la cantidad de transacciones realizadas.	Archivos planos (CSV o Excel)

Fuente: elaboración propia.

En la **Tabla 19**, se expone tanto la herramienta como el uso que se le brinda dentro del proceso.

Tabla 19. Herramientas implementadas dentro del proceso actual

Herramienta	Uso
Excel	Con esta herramienta, se realiza el proceso de integración de datos, tanto por medio de algunas fórmulas como con procesamiento manual de datos.
Hojas de cálculo de Google	Esta herramienta funciona como base de datos para los reportes, o bien, para crearlos por medio de tablas que el cliente puede consultar.
Tableau y Google Data Studio	Ambas son utilizadas para generar las visualizaciones, las cuales son generadas a criterio del desarrollador.
SAS	Esta herramienta permite realizar análisis e integración de datos. Generalmente, emplea para hacer el reporte trimestral, debido a la cantidad de datos.

Fuente: elaboración propia.

Por medio de este proceso, se realiza un total de dieciséis reportes, los cuales se detallan en la **Tabla 20**. Aquí se describen el reporte, fuentes de datos que se utiliza, tiempo promedio de ejecución y la periodicidad en la cual se generan. El tiempo promedio de ejecución contempla desde que se accede a los datos hasta la finalización del reporte. De acuerdo con Alfaro (2022), estos reportes son consumidos por gerentes o líderes de los equipos para validar las métricas en diferentes ventanas de tiempo.

Tabla 20. Reportes generados para el departamento Business Care y Boarding

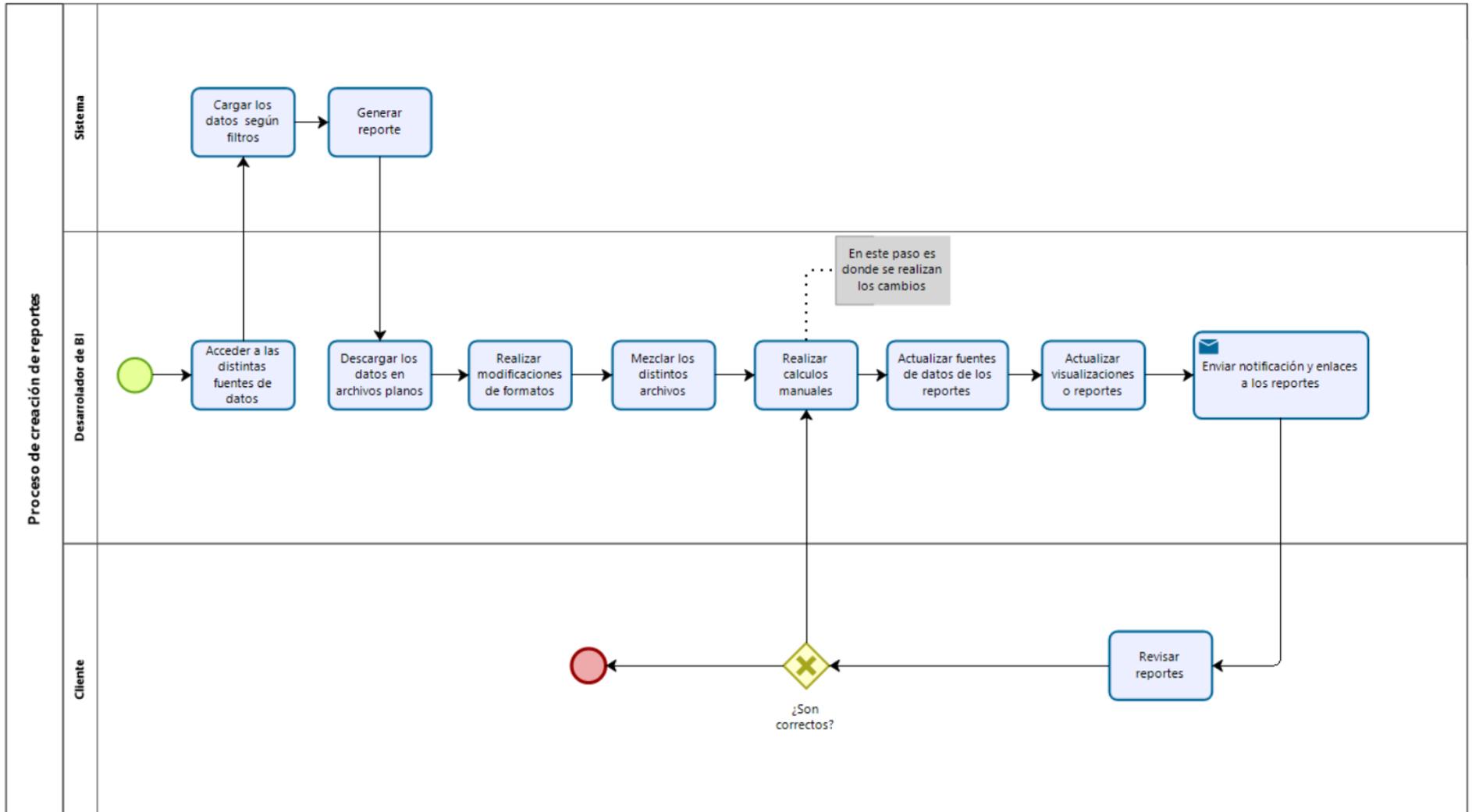
Reporte	Descripción	Fuentes	Periodicidad
VARVAB	Se muestran los casos asignados a los tres principales clientes revendedores. Se evidencia el número del caso, nombre del cliente, tiempo de apertura, estatus, actividad y tiempo que lleva abierto el caso.	Salesforce	Diario
Productividad-acreditación	Se presentan los casos asignados, tiempo promedio de resolución, cantidad de casos terminados y enviados a Estatus Remediation, por parte de los agentes asignados.	Salesforce	Diario
Productividad-provisionamiento	Se muestran los casos asignados, tiempo promedio de resolución, cantidad de casos terminados y cantidad de casos enviados a Estatus Remediation, por parte de los agentes asignados.	Salesforce	Diario
Productividad-reacreditación	Se evidencian los casos asignados, tiempo promedio de resolución, cantidad de casos terminados y enviados a Estatus Remediation de los agentes asignados.	AODOcs	Diario
Vejez de todos los casos (Boarding)	Se señalan todas las actividades abiertas, cuánto llevan en progreso y tiempo en Estatus Remediation de los casos de Boarding.	Salesforce	Diario
Vejez de todos los casos (Business Care)	Muestra todos los casos abiertos tanto en cola como asignados. Para estos, se muestra el tiempo que llevan abiertos y los comentarios que tienen.	Salesforce	Diario
Casos NCC	Corresponde a uno de los clientes revendedores más grandes de la organización; se muestran los casos abiertos, el estado del caso, los agentes encargado y asociado, compañía asociada y tiempo que lleva abierto el tiquete.	Salesforce	Diario
Calificación del agente	Muestra la nota promedio semanal por gerente, asignada a los agentes	Qualtrics	Semanal

Reporte	Descripción	Fuentes	Periodicidad
Casos de mantenimiento	Este reporte muestra todos los casos de mantenimiento de cuentas realizados mensualmente. Además, señala el tiempo de ejecución promedio, cantidad de tiempo enviado al estatus Remediation” y tiempo que llevan abiertos los casos.	Salesforce	Mensual
Nuevos clientes	Se evidencian todos los clientes nuevos que han sido acreditados durante el mes.	Salesforce	Mensual
Reporte de operaciones mensual	Se reportan todos los casos que han sido procesados, cerrados o abandonados. También, el promedio del tiempo de respuesta de los agentes y el nivel de satisfacción de los clientes.	Salesforce Genesis (Qualtrics) Tracker	Mensual
Reporte de tipo comercial	Se indica el volumen de llamadas recibidas, la calidad asignada a cada una de estas, el tiempo de espera del cliente, el tiempo promedio de duración en la llamada y el promedio de la satisfacción del cliente.	Salesforce Genesis (Qualtrics) Tracker	Mensual
Reporte de ACRO	Se muestran las consultas generadas por clientes al mes, cuántas fueron resueltas, tiempo promedio de resolución, volumen por tipo, y porcentajes atendidos por el <i>bot</i> y por un agente.	Salesforce RPA (CSV).	Mensual.
Reacreditación	Se utiliza para saber cuáles clientes deben ser reacreditados durante el siguiente trimestre. Cada uno de los clientes son asignados a un nivel (T1-T4) y se señala si estos son directos o indirectos.	AODocs (GoogleSheet) Looker Gmail (CSV)	Trimestral

Fuente: elaboración propia.

Una vez identificados estos datos, se obtuvo el diagrama As-is, representado en la **Figura 10**. Proceso As-is de creación de reportes En este se, describe el proceso actual para la creación de reportes por parte del departamento de Business Care y Boarding.

Figura 10. Proceso As-is de creación de reportes



Fuente: elaboración propia.

El modelo del estado actual del proceso representado en la **Figura 10**, fue validado por el encargado de crear estos reportes (ver **Minuta 10. Validación del modelo “as-is”**), con el fin de realizar los ajustes necesarios para estar más ajustados a la realidad actual. Por medio de esta reunión, se logra validar los siguientes puntos:

- Documentación

Se determina que no existe una documentación del proceso, ni de los datos, ni de los indicadores claves de desempeño (en adelante KPI), que se deben calcular en cada uno de los reportes. Se depende de la persona encargada de realizar los reportes, pues es la única que domina el proceso.

- Estandarización

El proceso no se encuentra estandarizado. De acuerdo con lo observado, algunas personas realizan los reportes según se les haga más sencillo. Por un lado, algunos cuentan con hojas de cálculo con espacios para facilitar la tarea. Por otro, los reportes se realizan con el formato deseado por el encargado.

- Nuevas solicitudes

No existe un proceso definido para la recepción de nuevos reportes. Según menciona Alfaro (2022, comunicación personal), “cuando se realizan nuevas solicitudes, el encargado de los reportes es el que recibe la solicitud y valida si es posible realizarla”. Debido a que no existe una definición de requerimientos o indicadores de desempeño, para hacer las visualizaciones, se puede generar retrabajo.

4.1.2. Identificación de necesidades del equipo y requerimientos del cliente

Como parte de los objetivos de esta fase, se encuentra identificar las necesidades tanto del equipo de Servicios de Datos como del proceso de creación de reportes. Esto surge debido a la inexistencia de una documentación al respecto en la empresa. Para identificar las necesidades, se realiza un análisis a partir de los lentes de frustración y tiempo, con los instrumentos mencionados por Madison (2005). Para la frustración, se implementa la matriz *Two-by-Two*, la cual se encuentra en el **Apéndice S**. Resultados de Matriz *Two-by-Two*, los valores obtenidos se encuentran resumidos en la **Tabla 21**.

Tabla 21. Ideas de mejoras-impacto vs. costo

Mejoras	Impacto y valor	Costo y nivel de complejidad
Actualización de los datos en los reportes de forma automática	Alto	Bajo
Mejorar las visualizaciones utilizadas en vez de tablas	Alto	Bajo
Documentar el proceso de creación de reportes	Bajo	Bajo
Extracción automatizada de datos	Alto	Alto
Calendarización del ETL	Alto	Alto
Brindar acceso a los reportes a los usuarios	Bajo	Alto
Uso de herramientas especializadas para generar ETL	Bajo	Alto

Fuente: elaboración propia.

Con respecto al lente del tiempo, se implementa el instrumento “Hoja de tiempos en el proceso” (**Apéndice T. Resultados de la hoja de tiempos del proceso.**), para esto se recolectan los datos del tiempo de ejecución de las actividades y, por medio de un documento compartido por Alfaro (**Minuta 10. Validación del modelo “as-is”**), se validan los tiempos de espera, revisión y retrabajo. Así, en la **Tabla 22**, se encuentra resumida dicha información: el ID asignado a cada paso, una breve descripción de este y el tiempo total que se toma en cada uno.

Tabla 22. *Tiempos del proceso*

# de paso	Descripción	Tiempo total
1	Se accede a las distintas fuentes de datos y se descargan los datos en archivos planos.	35 min.
2	Se realizan cambios en el formato de los datos.	26 min.
3	Se agrupan varios documentos en uno solo, para manipular y mezclar datos.	18 min.
4	Se actualizan fuentes.	28 min.
5	Se actualizan reportes.	10 min.

Fuente: elaboración propia.

Por otro lado, para identificar los requerimientos, se entrevista a los interesados e involucrados del proceso, ver la **Tabla 23**.

Tabla 23. *Involucrados e interesados*

Interesado o involucrado	Rol
Luis Vindas.	Encargado del departamento
Diego Alfaro	Encargado de realizar los reportes
Mateo Marín	Encargado de brindar soporte y desarrollo
Kimberly Long	VP de Business Care y Boarding.
Kevin Durio	Manager de Business Care

Fuente: elaboración propia.

Las entrevistas pueden consultarse en los **Apéndices**. De esta forma, se conocen las necesidades, desde distintas perspectivas dentro del proceso, así como los principales requerimientos que se deben satisfacerse con el nuevo proceso. Para especificar los requerimientos, se emplea la plantilla definida en el **Apéndices F**. Así, se detalla el identificador, el tipo de requerimiento (funcional o no funcional), su descripción, el objetivo y el nivel de prioridad.

Requerimientos por parte del equipo

ID del requerimiento	R001
Tipo de requerimiento	Funcional.
Descripción	Debe actualizar las fuentes de datos que utilizan los reportes de forma automática.
Objetivo	Automatización de reportes.
Prioridad	Alta.

ID del requerimiento	R002
Tipo de requerimiento	No funcional.
Descripción	Para la creación de reportes, se debe contar con documentación que contenga KPIs, fuentes, y los procesos de automatización para estandarizar y de creación de visualizaciones.
Objetivo	Documentación y estandarización.
Prioridad	Alta.

ID del requerimiento	R003
Tipo de requerimiento	Funcional.
Descripción	El proceso de extracción, transformación y carga de datos debe ser generado por medio de Python, debido a necesidades del negocio.
Objetivo	Crear una base de trabajo para proyectos futuros.
Prioridad	Media.

ID del requerimiento	R004
Tipo de requerimiento	Funcional.
Descripción	Se deben utilizar los procesos disponibles dentro de la organización para la extracción de datos de las herramientas.
Objetivo	Cumplir las políticas.
Prioridad	Alta.

ID del requerimiento	R005
Tipo de requerimiento	Funcional.
Descripción	Para la creación de reportes, se deben utilizar solamente las herramientas disponibles y aceptadas por las políticas de la organización.
Objetivo	Cumplimiento de políticas.
Prioridad	Alta

4.2. Segunda fase: Diseño del nuevo proceso

Para realizar esta fase, se emplean las entrevistas y se lleva a cabo una revisión documental. El fin principal es generar la propuesta de mejora para la creación de reportes, sin embargo, parte fundamental de este paso es estudiar las prácticas consideradas en el mercado como las mejores para administrar procesos de este tipo.

4.2.1. Evaluación de mejores prácticas

Para esta actividad, se hace una revisión documental en tres temas relacionados con la ejecución del proyecto. Dentro de las fuentes consultadas, se destacan buenas prácticas en inteligencia de negocios, diseño de tableros, y rediseño de procesos y gobernanza de datos. Para profundizar, se puede consultar el **Marco conceptual** y los **Apéndices**.

- Buenas prácticas en el rediseño procesos

En términos del rediseño de procesos, se realiza una revisión documental de dos fuentes relacionadas con la mejora o rediseño de procesos:

- *Process Mapping, Process Improvement, and Process Management*, de Dan Madison.
- *The Power of Business Process Improvement*, de Susan Page.

La importancia de estos dos libros radica en que definen mecanismos para el rediseño o mejora de procesos. Por lo tanto, sirven como guía para evaluar el proceso actual e identificar actividades pueden mejorarse, o bien, identificar cuáles son las que más tiempo consumen. Algunos aspectos relevantes para tomar en cuenta para el rediseño de procesos destacados por estas fuentes son:

- Las opiniones de las personas involucradas dentro del proceso, es decir, de los encargados de realizar estas actividades de forma diaria, pues los pros y contras de estas.
- Abarcar los conceptos de:
 - Eficacia
 - Eficiencia
 - Adaptabilidad

- Buenas prácticas en la inteligencia de negocios

Para realizar el presente proyecto se hace una revisión documental de las buenas prácticas en la inteligencia de negocios, enfocadas para el proceso ETL y para la integración de datos; tales prácticas son AWS, IBM, y de Kimball. Dentro de los aspectos más relevantes encontrados al analizarlas se encuentran:

- Para optimizar los procedimientos dentro de la inteligencia de negocios, se debe:
 - Copiar datos de múltiples archivos de tamaño uniforme.
 - Utilizar la gestión de la carga de trabajo para mejorar los tiempos de ejecución de ETL.
 - Realizar el mantenimiento regularmente.
 - Llevar a cabo múltiples pasos en una sola transacción.
 - Un procedimiento almacenado realiza tanto la carga inicial como la incremental para una tabla agregada. Si la tabla agregada está vacía, se realiza una carga inicial. Si hay datos en dicha tabla, se hacer una carga incremental.
- De acuerdo con Kimball, el proceso de cuatro pasos para el modelado de datos es el siguiente:
 - Elegir un proceso de negocio para modelar.
 - Decidir la granularidad, esto es el nivel de datos para almacenar como la tabla de hechos principal. Debe ser el nivel más atómico posible, es decir, un nivel de datos que no se pueda dividir más.
 - Elegir las dimensiones que se aplican a cada fila de la tabla de hechos.
 - Identificar los hechos numéricos que llenarán cada fila de la tabla de hechos.
- Es importante realizar un mapeo de relaciones: una relación externa entre dos modelos de datos.

- Buenas prácticas para el diseño de tableros digitales

Para crear un documento que permita estandarizar los diseños de las visualizaciones, se realiza una revisión documental de buenas prácticas para el proceso de diseño y del libro de marca de la organización, con el fin de generar un estándar acorde a esta. Dentro de los puntos más importantes encontrados, se destacan:

- Para realizar los tableros digitales, es importante tener en cuenta la siguiente información.
 - Los gráficos de barras son efectivos para comparar categorías dentro de una sola medida. Son especialmente eficaces cuando tiene datos que se pueden dividir en varias categorías.

- Los gráficos de viñetas muestran el progreso con respecto a un objetivo, al comparar las medidas. Se diseñaron para reemplazar los indicadores, medidores y termómetros del tablero.
- El gráfico de líneas conecta varios puntos de datos distintos y los presenta como una evolución continua. Es una forma sencilla y directa de visualizar los cambios en un valor en relación con otro.
- Los histogramas y los diagramas de caja muestran dónde se agrupan sus datos y se pueden comparar categorías.
- Los mapas son una obviedad para visualizar preguntas específicas de la ubicación o para ayudar en la exploración geográfica.
- Los gráficos circulares son útiles para agregar detalles a otras visualizaciones, pero no son tan efectivos por sí solos.
 - Contar historias de datos rápidamente con señales de color claras.
 - Usar estratégicamente del tamaño para visualizar valores.

4.2.2. Identificación y documentación de las herramientas para los procesos de extracción, carga y transformación de datos

Para realizar esta, se entrevista al gerente del departamento de datos Luis Vindas (**Apéndice R. Resultados de entrevista #4 – Limitaciones y requerimientos tecnológicos del departamento**). Como resultado, se conocen las alternativas disponibles para obtener los datos. Según lo mencionado por Vindas, si bien existen medios por los cuales se puede acceder a la información, esta no ha sido utilizada por el poco tiempo para revisar los datos y realizar los cambios en el proceso. Los medios referidos para extraer la información son:

- **BigQuery:**

Dentro de la organización, se utiliza como principal almacenamiento de datos. En esta herramienta, se replican los datos generados por los agentes dentro de Salesforce y en AODocs.

- **APIs:**

Existen aplicaciones con sus propios APIs, los cuales pueden consumirse. Para utilizar estas herramientas, es necesario solicitar accesos que permiten obtener credenciales. Algunas de las fuentes de datos que poseen APIs son Salesforce y Qualtrics, estas se pueden consultar para extraer datos directamente desde sus propias bases.

Identificados los métodos disponibles para la extracción de datos, se realiza una revisión de estos para identificar:

- Datos disponibles.
- Accesos requeridos.
- Método de conexión.

Los resultados obtenidos de esta revisión se encuentran detallados en la **Tabla 24**.

Tabla 24. Fuentes de datos alternativas

Método de extracción	Datos disponibles	Accesos requeridos	Método de conexión
BigQuery	Casos de Salesforce	<ul style="list-style-type: none"> • Acceso a GCP. • Acceso al proyecto de EDH. 	Extracción directa de la fuente.
	Agentes de Salesforce	<ul style="list-style-type: none"> • Acceso a GCP. • Acceso al proyecto de EDH. 	Extracción directa de la fuente.
	AODocs cases.	<ul style="list-style-type: none"> • Acceso a GCP. • Acceso al proyecto de AODocs. 	Extracción directa de la fuente.
API	Extracción de GoogleSheet	<ul style="list-style-type: none"> • Acceso a las hojas de Google. • Llave para el uso del API. • Cliente secreto generado por Google. 	Extracción directa a los reportes generados en GoogleSheet.
	Qualtrics	<ul style="list-style-type: none"> • Llave para el uso del API. • ID de los cuestionarios. 	Extracción directa de la fuente.

Fuente: elaboración propia.

4.2.3. Validación del diseño y pruebas de concepto

Una vez realizado el estudio de las buenas prácticas y marcos de trabajo existentes en el mercado, se lleva a cabo el diseño del nuevo proceso. Este busca tomar las ventajas que traen las herramientas de inteligencia de negocios para facilitar las tareas manuales realizadas por los encargados. Una de las ventajas es el uso de herramientas como los ETL, para realizar los reportes, lo cual favorece los tiempos de ejecución y disminuye el trabajo manual. Dicho esto, en la **Propuesta de solución**, se encuentra el nuevo modelo de creación de reportes. Este fue validado por los involucrados e interesados en el proceso (**Minuta 13. Validación de pruebas de concepto y consultas sobre proyectos similares**). Una vez aprobado, se realizan las pruebas de concepto. Para esto, se emplea el principio Pareto, el cual permite determinar cuáles eran las fuentes más utilizadas (**Apéndice AA. Diagrama de Pareto**). Se establece que en el orden de mayor a menor uso son:

- Salesforce
- Genesis (Qualtrics)

- Archivos CSV
- Looker

Las pruebas de concepto realizadas se detallan en la **Tabla 25**, en esta se menciona el ID, el reporte utilizado, la descripción y el apéndice que contiene la prueba de concepto. Para la ejecución de tales pruebas, se utilizó como lenguaje de programación Python, pues, según menciona Vindas (2022), por políticas de la organización es el aceptado. Existen herramientas para realizar ETL de una forma sencilla, dentro de las más conocidas se pueden destacar Pentaho y las herramientas que provee Microsoft (SQL Server Integration Services). No obstante, estas fueron mencionadas en medio de una reunión (**Minuta 8. 2da revisión del proyecto de graduación**), pero ambas fueron descartadas por Vindas debido a las políticas de Equifax. Esta se decisión se debe a que la empresa rompió relaciones comerciales con Microsoft y es prohibido el uso de herramientas de esta organización (se hacen excepciones con Excel) y Pentaho es descartado por estar escrito en el lenguaje de programación Java, además fue uno de los causantes del hackeo sufrido por la organización hace unos años. Dichas pruebas de conceptos se encuentran en el en el siguiente capítulo (**Propuesta de solución**).

Tabla 25. Pruebas de concepto realizadas

ID	Reporte	Descripción	Mejora	Fuente
1	Vejez de todos los casos (Business Care)	Indica todos los casos abiertos, en cola y abiertos asignados. Se muestra el tiempo que llevan abiertos y los comentarios.	Extracción de datos de Salesforce por medio BigQuery y creación de hoja de cálculos actualizada automáticamente.	Salesforce
2	Reporte comercial	Muestra el volumen de llamadas recibidas, la calidad asignada a cada llamada, tiempo de espera del cliente, tiempo promedio de duración en la llamada y promedio de la satisfacción del cliente.	Extracción por medio de un API.	Qualtrics
3	Productividad de casos, tipo acreditación diaria	Señala los casos asignados, tiempo promedio de resolución, cantidad de casos terminados y enviados a estatus Remediation de los agentes asignados a este equipo.	Extracción directa de las hojas de cálculo.	Gmail

Fuente: elaboración propia.

Para realizar las pruebas de concepto, se aplican las tres actividades claves de la inteligencia de negocios las cuales son: extracción, transformación y carga de datos. En cada una se llevan a cabo tareas de investigación y análisis documental de las fuentes internas.

- Extracción de datos

La extracción de datos es uno de los principales problemas del proceso actual, para lo cual se investigaron las herramientas mencionadas por Vindas en el punto 4.2.2. Los resultados se muestran en la **Tabla 26**.

Tabla 26. *Datos disponibles*

Fuente	Tabla	Datos
BigQuery	Cases	<ul style="list-style-type: none"> - ID del caso - Fecha y hora de apertura - Fecha y hora de última modificación - Fecha y hora de cierre - Estado del caso - Causa - Tipo de caso - ID de la cuenta asociada - Descripción - ID del agente
	Case History	<ul style="list-style-type: none"> - ID del caso - Fecha y hora de apertura - Fecha y hora de última modificación - Fecha y hora de cierre - Estado antiguo - Nuevo estado - ID de agente que realiza el cambio
	Users	<ul style="list-style-type: none"> - ID del agente - Nombre del agente - Apellido del agente - Rol - Departamento - Supervisor
	AODocs	<ul style="list-style-type: none"> - ID del caso - Fecha y hora de apertura - Fecha y hora de última modificación - Fecha y hora de cierre - Nivel de reacreditación

Fuente	Tabla	Datos
		<ul style="list-style-type: none"> - Estado - Decisión - Razón de la decisión
API	Google API	Todos los valores dentro de las hojas de cálculo.
	Qualtrics API	<ul style="list-style-type: none"> - ID de encuestas - Resultado de encuestas - Fecha y hora de envió - Fecha y hora de finalización - País - ID del cliente - Tipo - Representante de ventas - Nombre de compañía - Tipo de contacto - Razón

Fuente: elaboración propia.

Todos los datos obtenidos a través de BigQuery tienen un periodo promedio de una hora de tardanza. Esto se debe a que la replicación, tal y como está programada, toma tiempo para actualizar los valores de las tablas. Sin embargo, esto es mucho más real que la información brindada por el proceso original. En cuanto a los datos provenientes de los APIs, se actualizan en periodos definidos, debido a que solo basta con ejecutar la tarea o código que extrae los datos, según a las necesidades del negocio.

- Transformación de datos

La transformación de datos implementada en estas pruebas de concepto se realiza utilizando Python, herramienta establecida por la organización. Se hace una investigación documental de Python y de las librerías disponibles para realizar operaciones. Los resultados obtenidos se detallan en la **Tabla 27**.

Tabla 27. Resultados del análisis de Python y sus librerías

Librería	Descripción
CSV	Permite crear, leer y escribir sobre archivos de tipo CSV.
Json	Permite realizar operaciones de lectura y escritura en formato Json.
Dateutil	Permite realizar distintas transformaciones en valores de fechas.
Numpy	Permite trabajar con arreglos dentro de Python, para la manipulación de matrices y realizar operaciones algebraicas.
Google api	Permite hacer uso de las herramientas de Google a través de Python.

Librería	Descripción
Pandas	Permite realizar análisis de datos a través de Python, por medio de métodos especializados.
Psycopg2	Permite realizar conexiones con bases de datos PostgreSQL.
Re	Permite realizar análisis por medio de expresiones regulares (Regex).

Fuente: elaboración propia.

- Carga de datos

Esta actividad fue una de las más sencillas en términos de elección de herramientas, debido a que, por políticas internas, la única disponible para el equipo de Servicios de Datos es una base de datos PostgreSQL. Como se menciona en la **Tabla 27**, Python dispone de una librería especial para las conexiones con una base de datos PostgreSQL; en esta, se encuentran métodos que facilitan la carga de datos por lotes, es decir, realizar inyecciones de datos en grandes grupos, para favorecer la velocidad del proceso.

Obtenida la información sobre el proceso ETL que se desea implementar, se llevan a cabo las pruebas de concepto (pueden consultar en **Apéndice AA. Diagrama de Pareto**). Estas son validadas por medio de una comparación de los valores del proceso antiguo contra el propuesto, con ayuda del encargado actual de realizar los reportes (**Minuta 13. Validación de pruebas de concepto y consultas sobre proyectos similares**). Algunos de los aspectos más relevantes de este nuevo proceso mencionados por Alfaro (2022) fueron:

- Velocidad de ejecución del proceso:

Con el nuevo proceso, tanto el tiempo como el trabajo manual, se ven reducidos, gracias a la automatización por medio del ETL.

- Correctitud de datos:

La probabilidad de errores humanos se reduce, ya que no existe la necesidad de realizar tareas diarias. Los problemas podrían generarse por el código del ETL, pero son considerados como de la fuente de datos.

- Ajuste a las necesidades:

El nuevo proceso disminuye la cantidad de trabajo manual realizado, pues pasa a ser automatizado.

- Realización de cambios:

Debido a que muchas de las tareas se cambian al código, si existiera la necesidad de hacer modificaciones específicas, la dificultad de realizarlas sería mayor, debido a que no se llevaría a cabo en un solo espacio.

4.3. Tercera fase: Plan de implementación

En esta sección, se detallan los principales hallazgos de las actividades relacionadas con el plan de implementación de la propuesta. Se llevan a cabo entrevistas para recolectar información importante a la hora de definir el alcance, recursos, interesados, medios de comunicación y riesgos. Para todos estos puntos, se extrae la información por medio de entrevistas con Luis Vindas (**Apéndice P. Resultados de entrevista #2 - Descubrimiento del proceso.** y **Apéndice Q. Resultados de entrevista #3 – Uso y requerimientos del proceso para el cliente.**).

- Recursos

De acuerdo con lo mencionado por Vindas (2022), solo se asigna a una persona para realizar esta implementación. Los datos del recurso se encuentran en la **Tabla 28**. Datos de los recursos necesarios, donde se especifican tanto el personal como el equipo necesario para realizarla. Cabe destacar que se colocan 2 recursos ya que se ocupa el conocimiento del encargado de BI, el cual estará involucrado al inicio de la implementación y para la revisión. Por otro lado, el segundo recurso se mantendrá en la organización migrando los otros procesos.

Tabla 28. Datos de los recursos necesarios

Recurso	Cantidad	Descripción
Personal	2	Persona encargada de realizar la implementación.
Softwares	6	Los softwares necesarios para realizar la implementación son: <ul style="list-style-type: none"> • Python • Excel • BigQuery • Postman • GCP • PostgreSQL
Hardware	2	Equipo físico necesario para la implementación (computadora portátil).

Fuente: elaboración propia.

- Tiempo de implementación

Para el cálculo del tiempo necesario para la implementación, se toma en cuenta aspectos como: planeación (el presente proyecto), ejecución y evaluación. Basado en estos parámetros e investigando el alcance de cada una de las fuentes y herramientas, se establecen los siguientes periodos:

- Planeación: 16 semanas.
- Ejecución: 3 semanas.
- Evaluación: 2 semanas.

Alfaro (encargado de reportes) y Vindas (gerente) validan estos valores en una reunión (**Minuta 13. Validación de pruebas de concepto y consultas sobre proyectos similares**). Ellos mencionan que la implementación debe ser posterior al inicio del trimestre, debido a las altas cargas de trabajo durante este periodo.

- Interesados

Los principales beneficiarios e interesados, tanto personas como grupos relacionados de forma directa o indirecta con la implementación del proceso, se encuentran identificados en la **Tabla 29**.

Tabla 29. *Interesados en el proyecto*

Interesados		Expectativa	Importancia
Director del departamento de Servicio de Datos	del del	Colaborar con temas de herramientas, licencias y presupuesto.	Baja
Gerente del departamento del Servicio de Datos		Colaborar con accesos, información, conexiones, y conocimiento.	Alta
Desarrollador de BI del departamento de Servicio de Datos		Brindar la información completa del proceso y cumplir un rol de control de calidad.	Alta
Encargado de soporte del equipo		Ofrecer soporte a las herramientas utilizadas.	Media
Departamento de TI y seguridad		Proveer llaves para el uso de conexiones como BigQuery o el API de Google.	Media
Gerente del departamento de "Business Care"		Brindar conocimiento e información sobre los reportes generados y los indicadores esperados.	Alta

Fuente: elaboración propia.

- Riesgos

Los principales riesgos del proyecto a implementar se encuentran identificados en la **Tabla 30**.

Tabla 30. *Riesgos identificados*

ID	Descripción
1	Problemas con las licencias
2	Problemas de hardware
3	Aumento de costos operacionales
4	Problemas con la resistencia al cambio
5	Retrasos en instalaciones de software
6	Mala configuración de softwares
7	Poca comunicación
8	Cambios en políticas de la organización
9	Problemas con los accesos

Fuente: elaboración propia.

Es importante complementar los riesgos identificados, para esto, se detallan los criterios de probabilidad e impacto que se utilizan para clasificar cada uno.

- **Probabilidad:** grado que tiene el riesgo para materializarse durante la ejecución de las iniciativas.

Nivel	Probabilidad	Descripción
1	Baja	Ocurre en solo ciertas ocasiones.
2	Media	No tiene ninguna condición especial, pero es posible que ocurra.
3	Alta	Tiene una alta ocurrencia.

- **Impacto:** grado de influencia positiva o negativa que posee un riesgo al materializarse.

Nivel	Impacto	Descripción
1	Bajo	El riesgo puede tener un impacto bajo o nulo.
2	Medio	Si se presenta, debe atenderse rápidamente, no tiene un alto nivel de urgencia.
3	Alto	Presenta un impacto muy grave dentro de las operaciones de la organización.

Con base en los criterios establecidos, se elabora una matriz de calor para los riesgos del proyecto, los cuales se encuentran evaluados en el siguiente cuadro.

	Bajo	Medio	Alto
Alta	3	6	9
Media	2	4	6
Baja	1	2	3

Una vez definidos estos criterios, se evalúan los riesgos identificados. En la **Tabla 31**, se encuentran los resultados obtenidos.

Tabla 31. *Matriz de evaluación de riesgos*

ID	Nivel de probabilidad	Nivel de impacto	Resultado
1	Baja	Medio	
2	Baja	Alto	
3	Baja	Bajo	
4	Baja	Alto	
5	Medio	Alto	
6	Baja	Alto	
7	Baja	Medio	
8	Baja	Alto	
9	Media	Alta	

Fuente: elaboración propia.

4.4. Cuarta fase: Análisis de costo beneficio

En esta sección, se detallan los principales hallazgos de las actividades relacionadas con el análisis financiero de la propuesta. Se realizan entrevistas para recolectar información importante relacionada con cotización del proyecto, costos de salarios, impuestos y márgenes de ganancia. Posteriormente, en el capítulo 5, estos datos emplean para cotizar el plan de implementación.

- Costos asociados

En esta actividad, se determinan los costos asociados con la propuesta de solución. Se revisa documentación, adicionalmente, se entrevista al gerente del departamento, para conocer los aspectos que deben ser considerados al realizar el análisis.

- Beneficios directos e indirectos identificados

Durante esta actividad, mediante el instrumento de revisión documental, se identifican los beneficios tanto financieros como aquellos que no lo son de la propuesta planteada. Se emplean indicadores financieros para determinar la viabilidad de la solución propuesta.

Una vez comprendidas las fases del proceso metodológico, se pone en práctica cada una de las cuatro fases mencionadas, para recopilar toda la información necesaria y poder aplicar la propuesta de solución. En el siguiente capítulo, se encuentran los resultados obtenidos.

Capítulo 5. Propuesta de solución

En este capítulo se procede con el desarrollo de la propuesta de solución ante la problemática planteada para este trabajo. A partir de la información obtenida y analizada en el capítulo anterior, se desarrollan las siguientes tres fases: diseño del nuevo proceso, plan de implementación y análisis de costo beneficio de la propuesta

5.1. Fase 2: Diseño del nuevo proceso.

En esta fase se desarrolla la propuesta de un nuevo proceso para la creación de reportes para los equipos de Business Care y Boarding y se realiza un documento el cual contiene la documentación respectiva. El fin de esta fase es brindar una actualización al departamento de Servicios de Datos y estandarizar el proceso de creación de reportes.

5.1.1. Realización del nuevo proceso.

Una vez realizada la actividad de evaluación de mejores prácticas se realizó el rediseño del proceso de creación de reportes. Los criterios implementados para la selección de tareas que fueron automatizadas por medio del sistema se fundamentan en las tareas de extracción, transformación y carga de datos, esto debido a que estas tareas son:

- Ejecutadas dentro de un sistema, por lo cual existen procedimientos para automatizarlas.
- Son tareas las cuales tienen alto tiempo de ejecución de forma manual (Apéndice T. Resultados de la hoja de tiempos del proceso.).
- Son tareas susceptibles a errores humanos, debido a la ejecución de distintas actividades manuales.

En la **Figura 11**, muestra el proceso *To-be* propuesto, el cual fue diseñado para atender las necesidades planteadas por el departamento de Servicios de Datos de Equifax. Respecto al proceso *As-is*, se identifican las siguientes mejoras:

- Extracción de datos automatizada

La extracción de datos se realiza de forma automatizada diariamente. Esto quiere decir que, se obtienen los datos más nuevos por medio de conexiones directas a las distintas fuera de datos. Un aspecto importante para mencionar es que la extracción de datos puede llegar a ser calendarizada con ayuda de alguna otra herramienta de como Bamboo o Airflow.

- Carga de datos automatizada

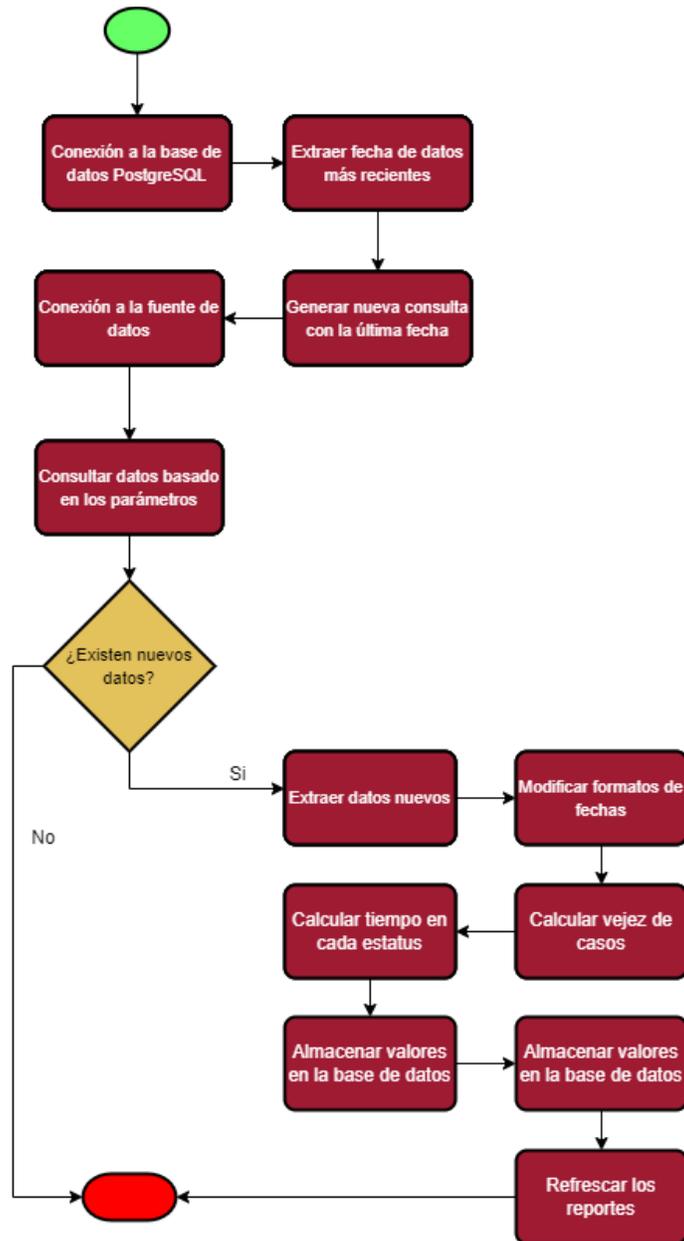
La carga de datos a las fuentes de los reportes se realiza de forma automática de forma diaria. Esto removiendo los pasos manuales en los cuales se debían almacenar los documentos en folder específicos para que fueran cargados por otra persona. Gracias a este proceso

- Conexión en vivo del reporte a la base de datos

Aprovechando las actualizaciones en el proceso de extracción y carga, se obtiene el beneficio de realizar conexiones en vivo desde los reportes a la base de datos. Con esto se ahorra el paso de tener que abrir cada uno de los reportes y actualizar las fuentes de datos. Ahora se tiene una conexión directa por lo cual, cualquier dato nuevo en la base de datos se ve reflejado en los reportes.

En la **Figura 12**, se muestra cómo, el sistema haría las distintas actividades de extracción, transformación y carga de datos (ETL). Se ejemplifica cómo toma una gran cantidad de las tareas que se realizaban manualmente, y permitiendo el uso de herramientas que ejecutan el proceso de forma automatizada.

Figura 12. Diagrama de flujo del sistema dentro proceso de elaboración de reportes



Fuente: elaboración propia.

Como se aprecia en las figuras anteriores, la extracción de datos se realiza directamente de sus fuentes, al no tomar los reportes que brinda la herramienta. La gran ventaja de esto es que se elimina la búsqueda información independiente. Esto quiere decir que solo se extrae información que está totalmente relacionada. También, se aprovecha la característica de que los reportes son tareas repetitivas y buscan siempre un mismo resultado.

5.1.2. Documentación del nuevo proceso

Uno de los puntos destacados por Vindas en la primera reunión fue la inexistencia de una documentación de este proceso. Por este motivo, se agregó un entregable para la empresa. Con base en la información de los capítulos anteriores, se generan estos documentos, con lo cual existirá una referencia que puede ser consumida y consultada por los colaboradores, además, le facilitará realizar sus tareas. Al elaborar este tipo de documentos, se obtienen beneficios tales como:

- Generar un estándar: se define una sola forma de realizar las tareas, con la posibilidad de seleccionar un procedimiento probado y funcional para las necesidades del negocio.
- Transferencia de conocimiento: las documentaciones permiten acelerar el proceso de transferencia de conocimiento: cuando ingresan nuevos colaboradores, es sencillo orientarlos a través de un proceso ya redactado.

Esta documentación cuenta con un total de seis secciones, en las cuales se desarrolla la información asociada con la ejecución del nuevo proceso. Se detallan aspectos relacionados con las solicitudes del negocio, así como con la ejecución y configuración del proceso. Cada uno de los contenidos de las secciones se describe a continuación:

- Introducción

Este apartado describe de forma general el documento donde se registra el proceso de creación de reportes para los departamentos Business Care y Boarding. Se detalla el objetivo y se desglosan los apartados.

- Reportes

En esta sección, se detallan los reportes realizados para cada unidad de negocios:

- Descripción: se describen los datos y visualizaciones utilizadas para generar el reporte.
- Periodicidad: se detalla el periodo en el cual debe actualizarse cada reporte.
- Fuentes de datos: se describe desde la información que las fuentes de datos contienen hasta los accesos necesarios para acceder a estas.

- Herramientas

En este apartado se detallan las herramientas necesarias para el nuevo proceso, así como las librerías y permisos necesarios para ejecutar cada una de las tareas.

- Funcionamiento del proceso

Acá se brinda una explicación del funcionamiento del proceso: las actividades realizadas, el diagrama de flujo correspondiente al nuevo proceso y la configuración necesaria de cada una de las herramientas implementadas.

- Estándar de visualizaciones

Para realizar un estándar del proceso completo, es necesario abarcar las visualizaciones. Por esta razón, este brinda a los encargados una guía para su elaboración. Morales y Oliver (2022) elaboraron una propuesta para el departamento. Algunos de los aspectos contemplados para la realización de este estándar son:

- Granularidad: las visualizaciones, por lo general, se realizan de acuerdo con el movimiento natural del ojo, el cual es de arriba hacia abajo y de izquierda a derecha. Debido a esto, se priorizan las visualizaciones de mayor importancia al lado superior izquierdo para posteriormente ir colando las de menor relevancia.
- Estructura: se establecen cuatro principales secciones dentro del *dashboard*:
 - Información: se coloca el nombre del reporte con una breve descripción de este; se acompaña de información de la organización y fechas.
 - Filtros: se emplean los filtros que permitan al usuario una mejor interacción con los datos. Debido a tamaño del *dashboard*, se estableció un máximo de cinco filtros.
 - Indicadores: se colocan los indicadores de mayor impacto para el *dashboard*; en general, valores numéricos, los cuales muestran totales, porcentajes, promedios, etc.
 - Visualizaciones: se sitúan los gráficos que conforman el *dashboard*, con base en la lógica mencionada.
- Paleta de colores: para ajustarse a las políticas de la organización, se utiliza la paleta establecida en su libro de marca.
- Gráficos: estos se recomiendan para distintos tipos de datos; se brinda información acerca los gráficos pastel, de barras, viñetas, líneas, mapas e histogramas.

- Pruebas de concepto

Estas pruebas se incluyeron como parte de la documentación, pues cumplen con la función de orientar a los encargados en el desarrollo de nuevos procedimientos, al automatizar más reportes. Dentro de la documentación, se incluyeron las pruebas de concepto para realizar las conexiones con las distintas fuentes de información, también se incluyeron algunas funciones importantes.

5.2. Fase 3: Plan de implementación

Debido a limitaciones de tiempo y procesos burocráticos que están fuera del alcance de la ejecución de este proyecto, se propone un plan de implementación que pueda ser puesto en práctica cuando el equipo esté preparado para realizar el cambio en el proceso de elaboración de reportes para Business Care y Boarding. Dentro este plan, se abarcan seis de las gestiones propuestas por el PMI en el PMBOK, con el fin de aplicarlo como un proyecto interno. Para esto se generó el entregable (**Apéndice DD. Plan de implementación**), el cual cuenta con los siguientes apartados.

- Introducción

Esta parte describe de forma general el documento realizado para la implementación del nuevo proceso de creación de reportes, para los departamentos Business Care y Boarding.

- Alcance

Este apartado presenta el alcance del proyecto a implementar: desde la generación de los ETL, hasta la calendarización de las tareas para que sean ejecutadas de forma diaria, sin la necesidad de contar con un recurso que lo ejecute manualmente.

- Gestión del tiempo

En esta sección se establece el tiempo necesario para ejecutar este proyecto: se establecen tiempos de ejecución de tareas, debido a que existen limitaciones con este factor. Se abarcan desde las actividades para generar la extracción de datos, hasta su actualización en la base datos para que sea consumidos por los clientes finales.

- Gestión de recursos

En esta parte, se definen los recursos necesarios para la ejecución del proyecto. De cada uno se establece su descripción, cantidad, tipo y función. Todos los recursos identificados son necesarios paralelamente para la ejecución del proyecto; como se obtienen de la organización, no hay necesidad de realizar un entrenamiento específico, lo cual impacta tanto en el tiempo como en los costos de la implementación.

- Gestión de interesados

En este apartado, se establece a los interesados e involucrados en el proyecto, es decir a las personas con algún nivel de influencia en las decisiones de este. Acá, se identifica el rol de cada una y el grado de poder sobre las decisiones.

- Gestión de comunicación

En esta sección, se establecen los medios de comunicación utilizados por cada uno de los involucrados e interesados. También, el tipo de información que estarían recibiendo y la periodicidad de su distribución.

- Gestión de riesgo

En esta parte, se enlistados todos los riesgos identificados en las etapas anteriores. De cada uno se presenta su evaluación de acuerdo con los parámetros establecidos y cómo se reaccionará si alguno se materializa.

5.3. Fase 4: Análisis de costo-beneficio

El presente apartado corresponde al análisis financiero de la propuesta de solución desarrollada en la sección anterior. A partir de rangos de datos proporcionados por la empresa, se calculan los montos necesarios para determinar la inversión de la propuesta. Este análisis se realiza con un promedio de datos proporcionados por la empresa, y con revisiones en el Ministerio de Trabajo y Seguridad Social o Instituto Nacional de Seguros, pues mucha información es confidencial y no puede detallarse.

5.3.1. Recursos

En el capítulo anterior, se definió la cantidad de recursos necesarios para esta implementación. En este ámbito, se abarca tanto a las personas necesarias para realizar el cambio en el modelo, como el equipo que estarán utilizando. En la **Tabla 32**, se detallan los recursos, el tipo de colaboración que brindan y el costo de cada uno, y se hace una descripción de estos.

Tabla 32. Recursos necesarios para implementación del nuevo modelo

Recurso	Cantidad	Descripción	Tipo de colaboración	Costo / Salario
Desarrollador de BI	1	Encargado de realizar e implementar el nuevo modelo	100%	€1,000,000
	1	Encargado de validar el modelo	50%	€1,000,000
Computadoras	2	Hardware necesario para realizar las tareas durante el proyecto de implementación	100%	\$550

Fuente: elaboración propia a partir de información brindada por Vindas (2022).

5.3.2. Licencias

Si bien la principal herramienta para realizar esta propuesta es Python, se deben considerar las licencias necesarias para la visualización de datos. Para los reportes elaborados en Tableau, se debe adquirir un total de 11 licencias. Estas pueden ser de dos tipos:

1. Licencia de visualización: permite a los usuarios a acceder a los reportes desde el servidor de Tableau. De acuerdo con la información proporcionada por Vindas (2022), Equifax cuenta con un contrato de precios diferenciados por consumir dos herramientas de la misma organización: Salesforce y Tableau, esta última fue adquirida por Salesforce en los últimos años. El precio de cada una de estas licencias es de \$97 al año.
2. Licencia de creador: son necesarias para realizar y publicar los reportes nuevos dentro del servidor de Tableau. Estas, al igual que las anteriores, tienen un precio especial para la organización: \$502 por licencia al año.

Los costos totales por concepto de licencias se detallan en la **Tabla 33**.

Tabla 33. *Costos de licencias*

Licencia	Costo	Cantidad	Costo total anual.
Licencia de visualización	\$97	9	\$873
Licencia de creador	\$502	2	\$1,100

Fuente: elaboración propia a partir de información brindada por Vindas (2022).

5.3.3. Tiempo de ejecución

El tiempo de ejecución se dividió en tres fases: planeación, ejecución y pruebas. Con esto, se busca hacer cálculos más certeros, se abarcan las 16 semanas que se tomaron para la planeación de este documento. A esto, se debe sumar el tiempo de ejecución determinado en el capítulo 4. En la **Tabla 34**, se exponen las actividades y el tiempo planteado para cada una de las fases. Dicho esto, la propuesta debe estar concluida a las seis semanas

Tabla 34. *Tiempo de ejecución de la propuesta*

Fase	Actividades	Tiempo
Planeación	<ul style="list-style-type: none"> - Descubrimiento del proceso - Revisión de fuentes - Revisión de herramientas - Plantear solución - Realizar documentos 	16 semanas
Ejecución	<ul style="list-style-type: none"> - Validación de requerimientos - Solicitar accesos - Crear instancias necesarias - Implementar nuevo modelo 	5 semanas
Pruebas	<ul style="list-style-type: none"> - Realizar pruebas y comparación para validar nuevo proceso 	1 semanas

Fuente: elaboración propia.

5.3.4. Costos totales

Una vez abarcados los puntos anteriores, se procede a realizar el cálculo del costo total de la propuesta. Se tomaron en cuenta los siguientes puntos para calcular los costos:

- Aportes del patrono en Caja Costarricense del Seguro Social (CSSS), en otras instituciones como cuota patronal Banco Popular, Asignaciones Familiares, el Instituto Mixto de Ayuda Social (IMAS), el Instituto Nacional de Aprendizaje (INA) y la Ley de Protección del Trabajador, equivalente a un 26.50% del salario.
- Beneficios que brinda la empresa a cada uno de estos trabajadores: subsidio alimenticio, transporte a la oficina y café. Estos suman un monto total promedio de ₡86,000 mensuales. Por otro lado, los costos promedio por colaborador para el mantenimiento es de ₡40,000 mensuales.
- El tipo de cambio establecido para realizar este cálculo fue ₡634.1, este monto ha sido el promedio en los últimos 30 días según el Banco Central.

Los costos totales para poner en marcha la propuesta son un total de **₡9,538,218.7** para el primer año y de **₡1,190,205.7** a partir del segundo año por concepto de licencias y **₡16,692,000** por concepto de salario y beneficios del recurso. Todos los detalles de los costos se encuentran en la **Tabla 35**. Cabe destacar que tanto el costo de las laptops como el de las licencias se encuentran sujetos a modificaciones, debido al tipo de cambio del dólar, ya que su base es en esta moneda.

Tabla 35. Costos totales

Investigación y desarrollo de la propuesta				
Concepto	Costos mensuales	Plazo total	Inversión total	Inversión a partir del segundo año
Desarrollador de BI	₡1,265,000 mensual (Salario + cargas sociales)	5 meses	₡6,325,000	₡15,180,000
Desarrollador de BI	₡1,265,000 mensual (Salario + cargas sociales)	0.5 meses	₡632,500	N/A
Beneficios	₡86,000 mensuales	5.5 meses	₡473,000	₡1,032,000
Mantenimiento	₡40,000 mensuales	5.5 meses	₡220,000	₡480,000
2 laptops	₡348,755	Se paga una sola vez	₡697,510	

Licencias			
Concepto	Costo	Cantidad	Inversión
Licencias de visualización	₡61,507.7 anuales	9	₡553,569.3
Licencias de creador	₡348.755 anuales	2	₡697,510

Fuente: elaboración propia.

De acuerdo con los resultados obtenidos a través del ROI (Apéndice BB. Retorno de Inversión). A los dos años posteriores del desarrollo de la propuesta se ha recuperado la inversión 0.81 veces. Al ser un ROI positivo al segundo año, la inversión en la implementación del nuevo proceso es viable.

5.3.5. Beneficios

Para realizar el cálculo de los posibles beneficios que se obtendrían con esta propuesta, se hace una reunión con Vindas (**Minuta 16. Validación de beneficios de la propuesta**), cuyo fin es mostrar el contenido propuesto y los costos que tendría. Una vez se realizada esta actividad, se plantean los beneficios en conjunto con la parte de la organización, los resultados fueron:

- Estandarización: gracias a la existencia de un proceso comprobado y documentado, es posible estandarizar la forma en que se realizan los procesos ETL dentro del equipo. Esto abre las puertas a su implementación para otros departamentos y clientes.
- Disminución de errores: si bien es cierto los errores son muy difíciles de quitar en su totalidad, al eliminar los procesos manuales, se favorece la disminución de errores humanos, pues la información se tiene controlada y accesible. Esto, de acuerdo con Vindas (2022), mejora la credibilidad y desempeño del departamento, y mitiga la posibilidad de recibir sanciones por parte de entes reguladores.
- Uso efectivo del tiempo: debido a que muchas tareas se realizan a través de sistemas, el tiempo disponible para que los agentes tomen tareas se ve favorecido. De acuerdo con Vindas (2022), al desocupar a los encargados de tareas manuales, se puede utilizar su tiempo para la migración de otros procesos e implementación de este modelo.
- Ahorro en licencias: según Vindas (2022), la implementación de este nuevo proceso facilitaría la ejecución de los reportes. Además, se eliminaría la necesidad de contar con licencias de sistemas especializados como SAS, el cual tiene un costo de \$15,450 cada licencia al año.
- Material para entrenamiento: una de las necesidades mencionadas desde el día uno, es contar con una documentación. Gracias a la implementación de esta propuesta, el departamento estaría contando con una documentación que permita transferir el conocimiento a nuevos recursos, lo cual favorecería la curva de aprendizaje.

5.3.6. Resultados del análisis

A modo de resumen, en la **Tabla 36**, se realiza un cuadro comparativo entre los costos de la propuesta contra los beneficios que esta brindaría a la organización con los cuales se obtuvo un ROI de un 0.81 al segundo año.

Tabla 36. *Resumen de costos y beneficios*

Costos	Beneficio
Costo inicial: ₡9,538,218.7	Estandarización del procedimiento de elaboración de reportes.
	Documentación del proceso para transferencia del conocimiento a nuevos colaboradores.
	Ahorro de \$15,450 en licencias de SAS.
Costo anual a partir del segundo año: ₡ 1,190,205.7	Uso efectivo del tiempo gracias a la automatización de tareas.
	Favorece la disminución de al momento de realizar los reportes.

Fuente: elaboración propia.

Para resumir la relación entre las variables y el trabajo ejecutado durante estos últimos capítulos, en la **Tabla 37** se resumen las variables y cada resultado obtenido y propuesta realizada.

Tabla 37 *Resumen de variables*

Variable	Análisis de resultados	Propuesta
Situación actual del proceso de creación de reportes.	4.1.1. Identificación del estado actual	
Deficiencias existentes en el proceso actual.	4.1.2. Identificación de necesidades del equipo y requerimientos del cliente	
Fuentes de datos utilizadas.	4.2.2 Identificación y documentación de las herramientas para los procesos de extracción, carga y transformación de datos	5.1 Fase 2: Diseño del nuevo proceso.
Estandarización del proceso automatizado.	4.2.1 Evaluación de mejores prácticas	
Funcionalidad del rediseño propuesto.	4.2.3 Validación del diseño y pruebas de concepto	
Plan de implementación para la aplicación del modelo.	4.3 Tercera fase: Plan de implementación	5.2 Fase 3: Plan de implementación
Costo-beneficio.	4.4 Cuarta fase: Análisis de costo beneficio	5.3 Fase 4: Análisis de costo-beneficio

Fuente: elaboración propia.

Capítulo 6. Conclusiones

En este capítulo, se muestran las conclusiones obtenidas a partir de la investigación realizada. Estas agrupan según el objetivo específico al cual corresponde cada una.

6.1. Conclusiones: objetivo específico 1

En referencia al objetivo específico 1: “Analizar el estado actual del proceso de creación de reportes para la identificación de necesidades del equipo y el modelado del nuevo proceso”, se concluye lo siguiente:

- Con respecto a la situación actual,
 - El departamento de servicios de datos cuenta con un proceso manual para la realización de tareas las cuales pueden llegar a ser automatizadas por medio de procedimientos computacionales.
 - En el departamento no existe un estándar para realizar procesos ETL, ni para recibir solicitudes para nuevos reportes; tampoco hay alguna documentación sobre estos.
 - Los reportes son generados sin ningún tipo de estándar o buena práctica. Cabe destacar que, entre 5 a 10 de estos eran tablas resumen entregadas a través de hojas de cálculo de Google.
- Con respecto a las deficiencias del proceso, necesidades del equipo y el modelado del nuevo proceso,
 - El proceso actual cuenta con una dependencia a tareas manuales, las cuales aumentan el tiempo de ejecución.
 - Existe una dependencia a reportes creados por sistemas los cuales son la fuente de datos para generar nuevos, debido a que no existe una conexión directa a la data cruda.
 - Se identifican cinco actividades manuales realizadas por los desarrolladores de BI las cuales pueden ser realizadas por los sistemas disponibles en la organización.
 - En Equifax existen métodos de extracción de datos los cuales pueden ser utilizados por los departamentos. Se determina que el principal problema es la falta de distribución de información y la poca documentación existente.

6.2. Conclusiones: objetivo específico 2

En cuanto al objetivo específico 2: “Diseñar un nuevo proceso de extracción, transformación y carga de datos (ETL), su respectiva documentación y visualización de reportes para ser evaluados por medio pruebas de concepto.”, se concluye que:

- Con respecto a las fuentes de datos,
 - Dentro del proceso se hace uso de cinco fuentes de datos principales, las cuales pueden llegar a ser consultadas a través de Big query o conexión por APIs.
 - De acuerdo con análisis Pareto, las fuentes de datos más utilizadas para realizar los reportes para “*Business Care*” y “*Boarding*” son: Salesforce, qualtrics y documentos de Google sheets.

- Con respecto a la estandarización del proceso,
 - Se obtuvo una documentación tanto del proceso como de los reportes generados para los departamentos de business care y Boarding.
 - Para el desarrollo de visualizaciones se establece un estándar para generar reportes, de esta forma se establece el cómo se realizarán los reportes dentro del departamento cuidando los colores, distribución y diseño.

- Con respecto a la funcionalidad del rediseño del proceso,
 - El proceso ETL generado cumple las necesidades de extraer los datos automáticamente, delegando muchas de las tareas al software y refrescarlos de forma diaria, lo cual favorece los tiempos de ejecución del proceso.
 - Por medio de los procesos ETLs desarrollados en las pruebas de concepto, es posible extraer la información desde la fuente de datos. Gracias a esto, se valida el proceso propuesto.
 - La decisión de Equifax para utilizar Python como herramienta principal para realizar los ETL abre la puerta a nuevas oportunidades como la minería de datos o “*Machine Learning*” para realizar futuras mejoras a los reportes.

6.3. Conclusiones: objetivo específico 3

Con respecto al objetivo específico 3: “Desarrollar un plan de implementación para la automatización del proceso de creación de reportes y su respectivo análisis de costo-beneficio para la valoración.”, se concluye lo siguiente:

- Con respecto al plan de implementación,
 - No existen fechas concretas para la puesta en marcha del plan de implementación generado, por lo cual se procede a realizar un cronograma basado en periodos de tiempo.
 - Se genera un plan de implementación el cual puede llegar a ser aplicado en el momento que la organización desee implementar el nuevo proceso.
 - La organización cuenta con los recursos materiales y de personal para la ejecución del plan propuesta.
- Con respecto al análisis de costo- beneficio.
 - De acuerdo con los resultados del costo vs beneficio, implementar el nuevo modelo genera más beneficios (financieros y no financieros) que el proceso existente actualmente en el departamento

Por último, en el siguiente capítulo se dan recomendaciones a la empresa en función de las conclusiones anteriores.

Capítulo 7. Recomendaciones

En este capítulo, se especifican las recomendaciones obtenidas a partir del trabajo realizado al equipo de Servicios de Datos dentro de Equifax.

- Se sugiere que el departamento implemente la propuesta presentada para obtener todos los beneficios definidos, y así generar un estándar dentro de la organización y disminuir el costo de licencias.
- Se recomienda que el departamento de Servicios de Datos establezca procedimientos definidos para la recepción de nuevas solicitudes.
- Se recomienda la implementación del proceso ETL automatizado, para el resto de los clientes y reportes con los que cuenta el departamento de datos.
- Se aconseja brindar capacitaciones a los encargados de realizar los reportes, con el fin de maximizar los beneficios de Python y demás herramientas de visualización de datos.
- Dados los cambios constantes en las aplicaciones utilizadas dentro de Equifax, se sugiere tener debidamente documentados todos los otros procesos y reportes que realiza el equipo. Así, se pueden buscar soluciones que favorezcan los procesos de automatización en vez de realizar procesos manuales.
- Se recomienda indagar en las herramientas dentro de la organización, con el fin de contar con las mejores herramientas posibles y apalancarse de software libres, los cuales no ocupan licencias y pueden llegar a ser explotados de mejor forma
- Se aconseja realizar revisión de costos si el plan propuesto es implementado a más de seis meses desde la fecha en la cual se realizó el cálculo, para contar con los cursos más actualizados con la realidad organizacional.

Capítulo 8. Referencias bibliográficas

- Aguilera Díaz, A. (2017). El costo-beneficio como herramienta de decisión en la inversión en actividades científicas. *Cofin Habana*, 11(2), 322-343. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2073-60612017000200022&lng=es&tlng=es.
- Acuña, L. (2018). Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes. Caso: Empresa Áltica. [Trabajo Final de Graduación, Instituto Tecnológico de Costa Rica]. https://tecdigital.tec.ac.cr/dotlrn/escuela/escuela.ATI/file-storage/view/documentos-trabajo-final-degraduaci-n%2Fproyectos-finales-de-graduaci-n-p-blicos%2F2018%2FBI_Automatizar_reportes_LuisAcu%C3%B1a_II-2018
- AWS. (s. f.). *¿Qué es un API?* Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/api/#:%7E:text=API%20significa%20%E2%80%9Cinterfaz%20de%20programaci%C3%B3n,de%20servicio%20entre%20dos%20aplicaciones>.
- Burns, E. (2020). *Decision-making process*. TechTarget. <https://www.techtarget.com/searchbusinessanalytics/definition/decision-making-process>
- Carrión, J. (2017). *Diferencia entre dato información y conocimiento*. <http://iibi.unam.mx/voutssasmt/documentos/dato%20informacion%20conocimiento.pdf>
- David Hillson [Project Management Institute (PMI)]. (9 de Junio de 2017). *PMBOK® Guide: What is it?* [Archivo de Video]. YouTube. <https://www.youtube.com/watch?v=MCF7g-BoBj4>
- Delgado, R. H. E., Paitán, Ñ. H. J., Vilela, P. J. y Dueñas, V. M. R. (2018). *Metodología de la investigación. Cuantitativa, cualitativa y redacción de la tesis (5ª ed.)*. Ediciones de la U.
- Gartner. (s. f.). *Definition of Business Process Management (BPM) – Gartner Information Technology Glossary*. [https://www.gartner.com/en/information-technology/glossary/business-process-managementbpm#:~:text=Business%20process%20management%20\(BPM\)%20is,support%20of%20a%20business%20strategy](https://www.gartner.com/en/information-technology/glossary/business-process-managementbpm#:~:text=Business%20process%20management%20(BPM)%20is,support%20of%20a%20business%20strategy).
- Equifax. (9 de febrero de 2022). *Equifax Delivers Record Revenue and Eighth Consecutive Quarter of Double-Digit Growth*. Equifax Inc. <https://investor.equifax.com/news-events/press-releases/detail/1214/equifaxdelivers-record-revenue-and-eighth-consecutive>
- Equifax. (s. f.). *Who we are*. <https://www.equifax.com/about-equifax/who-we-are/>
- Equifax. (s. f.). *Gobierno Corporativo*. <https://www.equifax.co.cr/acerca-de-equifax/gobierno-corporativo>
- Gartner. (s. f.). *Definition of Business Process Management (BPM) – Gartner Information Technology Glossary*.

[https://www.gartner.com/en/information-technology/glossary/business-process-managementbpm#:~:text=Business%20process%20management%20\(BPM\)%20is,support%20of%20a%20business%20strategy.](https://www.gartner.com/en/information-technology/glossary/business-process-managementbpm#:~:text=Business%20process%20management%20(BPM)%20is,support%20of%20a%20business%20strategy.)

Google. (s. f.). *¿Qué es la integración de datos?* Google Cloud. Google Cloud. <https://cloud.google.com/learn/what-is-data-integration?hl=es-419>

Google. (s. f.). *BigQuery – Ayuda de Google Cloud Platform Console*. Google Cloud Platform. <https://support.google.com/cloud/answer/9113366?hl=es>

Halicki, R., Mirabal, J., Price, E. y Buck, A. (julio de 2022). *Realización de una prueba de concepto o un piloto-Azure Architecture Center*. Microsoft Docs. <https://docs.microsoft.com/es-es/azure/architecture/serverless-quest/poc-pilot>

Hernández Sampieri, R. y Mendoza, P. (2018). *Metodología de investigación*.

IBM Cloud Education. (5 de julio de 2022). *ETL (Extract, Transform, Load)*. IBM. <https://www.ibm.com/cloud/learn/etl#toc-what-is-etl-xeCDpL69>

Kruse, K. (8 de marzo de 2016). *The 80/20 Rule and How It Can Change Your Life*. Forbes. <https://www.forbes.com/sites/kevinkruse/2016/03/07/80-20-rule/?sh=4ac2da703814>

Mata, C. (2022, 24 agosto). *Segunda Sesión TFG Charla Marco Metodológico* [Vídeo]. YouTube. https://www.youtube.com/watch?v=57u_Y_76gRg

MacNeil, C. (2022). *Análisis de costo-beneficio: 5 pasos para tomar mejores decisiones*. Asana. <https://asana.com/es/resources/cost-benefit-analysis>

Madison, D. (2005). *Process Mapping, Process Improvement and Process Management: A Practical Guide to Enhancing Work Flow and Information Flow*. Paton Professional.

Martins, J. (agosto de 2022). *Qué es un KPI, para qué sirve y cómo utilizarlo en tu proyecto*. Asana. <https://asana.com/es/resources/key-performance-indicator-kpi>

Microsoft. (2022). *Qué son las herramientas de inteligencia empresarial (BI)*. Microsoft Azure. <https://azure.microsoft.com/es-es/overview/what-are-business-intelligence-tools/>

Microsoft. (2022). *What is business intelligence?* <https://powerbi.microsoft.com/en-us/what-is-business-intelligence/>

Monje, C. (2011). *Metodología de la investigación cuantitativa y cualitativa*. Universidad Sur colombiana, Facultad de Ciencias Sociales y Humanas.

Naranjo, L. (25 de agosto de 2021). *Marco Metodológico: Enfoque Alternativo & Ciencia de Diseño* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=9hXYRZUxGUw>

Navarro, J. (2019). *Propuesta de Implementación de una Solución de Inteligencia de Negocios que apoye el Análisis de Información y la Toma de Decisiones para Unity Service Delivery and Electronic Payment* [Trabajo Final de Graduación, Instituto Tecnológico de Costa Rica]. <https://tecdigital.tec.ac.cr/dotlrn/escuela/escuela.ATI/file>

storage/view/documentos-trabajo-final-de-graduaci-n%2Fproyectos-finales-de-graduaci-n-pblicos%2F2019%2FPropuesta_Solucion_Inteligencia_Negocios_DanielNavarro_II-2019.pdf

Page, S. (2015). *The Power of Business Process Improvement: 10 Simple Steps to Increase Effectiveness, Efficiency, and Adaptability*. AMACOM.

Pérez, D. (2018). *Desarrollo de herramienta de inteligencia de negocios para el área de ventas Caso: Equifax USA*. [Trabajo Final de Graduación, Instituto Tecnológico de Costa Rica]. <https://tecdigital.tec.ac.cr/dotlrn/escuela/escuela.ATI/file-storage/view/documentos-trabajo-final-degraduaci-n%2Fproyectos-finales-de-graduaci-n-pblicos%2F2018%2FBI%20para%20%20C3%A1rea%20de%20ventasDouglas%20P%20C3%A9rez-I-2018>

OMG. (enero de 2011). *About the Business Process Model and Notation Specification Version 2.0*. Object Management Group. <https://www.omg.org/spec/BPMN/2.0/>

Oracle. (s. f.). *¿Qué es una base de datos?* <https://www.oracle.com/mx/database/what-is-database/>

Project Management Institute. (2022). *Guía de los fundamentos para la dirección de proyectos (guía del pmbok(r))* (7ª ed.). Project Management Institute.

Qlik. (s. f.). *Business Intelligence Reporting: A Complete Guide*. <https://www.qlik.com/us/business-intelligence/business-intelligence-reporting#:~:text=Reporting%20Best%20Practices-,What%20Is%20BI%20Reporting%3F,improve%20decisions%20and%20business%20performance.>

Real Academia Española. (s. f.). Herramienta. En *Diccionario de la lengua española*. <https://dle.rae.es/herramienta>

Real Academia Española. (s. f.). Implementar. En *Diccionario de la lengua española*. <https://dle.rae.es/implementar>

Real Academia Española. (s. f.). Reporte. En *Diccionario de la lengua española*. <https://dle.rae.es/reporte>

Tableau. (s. f.). *Toma de decisiones basadas en los datos: cómo tener éxito en la era digital*. <https://www.tableau.com/es-es/learn/articles/data-driven-decision-making#:~:text=%20BFQu%20es%20la%20toma%20de,las%20iniciativas%20de%20una%20organizaci%C3%B3n.>

Tenorio, M. (2022). *Introducción a Equifax* [Diapositivas]. Equifax.

Ulate, I. y Vargas, E. (2014). *Metodología para elaborar una tesis* (1ª ed.). Editorial Universidad Estatal a Distancia

Capítulo 9. Apéndices

9.1. Apéndice A. Plantilla de minutas

Minuta de Reunión			
Información general			
Núm. Reunión:		Fecha:	
Lugar/Medio:		Hora de inicio y finalización	
Objetivo:			
Participantes:	Presentes		
	Ausentes		
Temas tratados			
ID	Asunto	Comentarios	Acuerdos
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.2. Apéndice B. Plantilla para control de cambios

Hoja de Control de Cambios			
Datos Generales del Cambio			
Nº Cambio			
Solicitante		Fecha de solicitud del cambio	
Responsable de la implementación		Fecha de realización del cambio	
Estado	<input type="checkbox"/> Aprobado <input type="checkbox"/> En Revisión <input type="checkbox"/> Rechazado		
Detalles del Cambio			
Categoría			
Descripción detallada			
Justificación			
Implicaciones de realizar el cambio			
Impacto			
Comentarios/ Observaciones			

9.3. Apéndice C. Plantilla de recopilación de datos en revisión documental

Revisión Documental			
Fecha:		Fuente:	

Enlace:	
Objetivo	
Detalles	
1.	

9.4. Apéndice D. Plantilla de recopilación de datos en la observación

Observación de proceso					
Fecha:		Hora inicio:		Hora final:	
Proceso observado:					
Participantes					
Nombre			Rol dentro del proceso		
Detalles					
Descripción del proceso					
Pasos:					

9.5. Apéndice E. Plantilla de entrevista

Detalles de entrevista					
Fecha:		Hora inicio:		Hora final:	
Objetivo:					
Participantes					
Nombre			Rol dentro del proceso		
Preguntas					
Pregunta generadora.					

9.6. Apéndice F. Plantilla de requerimientos

ID del requerimiento	
Tipo de requerimiento	
Descripción	
Objetivo	
Prioridad	

9.7. Apéndice G. Plantilla de la matriz *Two-by-Two*

Alto		
Impacto		
	Bajo	Valor

9.8. Apéndice H. Plantilla de hoja de tiempos

Hoja de tiempo en el proceso					
Paso #	Tiempo de procesamiento	Tiempo de espera	Tiempo de revisión	Tiempo de retrabajo	Tiempo total.

9.9. Apéndice I. Resultados de la observación–Extracción de datos del reporte mensual

Observación de proceso					
Fecha:	01/09/2022	Hora inicio:	8:00 am	Hora final:	9:10 am
Proceso observado:	Elaboración de reportes mensuales				
Participantes					
Nombre			Rol dentro del proceso		
Diego Alfaro			Desarrollador de inteligencia de negocios		
Detalles					
Descripción de la actividad	Por medio de este proceso, se extraen los datos necesarios para realizar reportes mensuales				
Pasos:	<p>Datos de Salesforce</p> <ol style="list-style-type: none"> 1. Para la versión antigua: <ol style="list-style-type: none"> a. Se accede a la herramienta de Salesforce. b. Se buscan los reportes creados en Salesforce los cuales son 4. c. Se descarga en un archivo de tipo CSV. 2. Para la versión moderna: <ol style="list-style-type: none"> a. Se accede a la herramienta de Salesforce. b. Se buscan los reportes creados en Salesforce los cuales son 6. c. Se descarga en un archivo de tipo CSV. <p>Datos de AODocs</p> <ol style="list-style-type: none"> 1. Se accede a la herramienta Bulk Updater. 2. Se ejecuta el reporte, el cual notifica una vez haya extraído los datos. 3. Los datos son enviados al correo por medio de una hoja de cálculos de Google. 4. Se descargan los datos en un archivo CSV. <p>Datos de Looker y hojas de cálculos Google</p> <ol style="list-style-type: none"> 1. Acceder a Gmail 2. Buscar los correos correspondientes 3. Abrir documento (Estos son archivos de Google sheets) 4. Se descargan los datos en un archivo CSV 				
Comentarios	<ul style="list-style-type: none"> - Se utilizan las dos versiones de Salesforce ya que algunos casos aun utilizan la versión antigua y no se han migrado. - En la versión moderna de Salesforce, hay problemas en la forma que los agentes seleccionan el cliente dentro de la herramienta, por lo cual hay que realizar la búsqueda de datos a dos niveles por limitaciones de la herramienta. - Con respecto a los datos de AODocs, no existe un tiempo definido de duración en la extracción ya que se depende de la herramienta. De acuerdo con Diego tiene un tiempo promedio de 10 - 15 minutos. Al realizar esta observación el sistema demora 11 minutos. - Los datos de Looker son enviados semanalmente al correo del encargado. - Existe información la cual es enviada por correo electrónico con el acceso a la hoja correspondiente. 				

9.10. Apéndice J. Resultados de la observación – Extracción de datos del reporte diario y semanal

Observación de proceso					
Fecha:	01/09/2022	Hora inicio:	8:00 am	Hora final:	8:50 am
Proceso observado:	Elaboración de reportes semanales y diarios				
Participantes					
Nombre			Rol dentro del proceso		
Diego Alfaro			Desarrollador de inteligencia de negocios		
Detalles					
Descripción de la actividad	Por medio de este proceso, se extraen los datos necesarios para realizar reportes de las distintas fuentes.				
Pasos:	<p>Datos de Salesforce</p> <ol style="list-style-type: none"> 3. Para la versión antigua: <ol style="list-style-type: none"> a. Se accede a la herramienta de Salesforce. b. Se buscan los reportes creados en Salesforce los cuales son 2. c. Se descarga en un archivo de tipo CSV. 4. Para la versión moderna: <ol style="list-style-type: none"> a. Se accede a la herramienta de Salesforce. b. Se buscan los reportes creados en Salesforce los cuales son 4. c. Se descarga en un archivo de tipo CSV. <p>Datos de AODOcs</p> <ol style="list-style-type: none"> 5. Se accede a la herramienta Bulk Updater. 6. Se ejecuta el reporte, el cual notifica una vez haya extraído los datos. 7. Los datos son enviados al correo por medio de una hoja de cálculos de Google. 8. Se descargan los datos en un archivo CSV. <p>Datos de hojas de cálculos Google</p> <ol style="list-style-type: none"> 5. Acceder a Gmail 6. Buscar los correos correspondientes 7. Abrir documento (Estos son archivos de Google sheets) 8. Se descargan los datos en un archivo CSV 				
Comentarios	<ul style="list-style-type: none"> - Se utilizan las dos versiones de Salesforce ya que algunos casos aun utilizan la versión antigua y no se han migrado. - En la versión moderna de Salesforce, hay problemas en la forma que los agentes seleccionan el cliente dentro de la herramienta, por lo cual hay que realizar la búsqueda de datos a dos niveles por limitaciones de la herramienta. - Con respecto a los datos de AODOcs, no existe un tiempo definido de duración en la extracción ya que se depende de la herramienta. De acuerdo con Diego tiene un tiempo promedio de 10 - 15 minutos. Al realizar esta observación el sistema demora 12 minutos. - Existe información la cual es enviada por correo electrónico con el acceso a la hoja correspondiente. 				

9.11. Apéndice K. Resultados de la observación – Transformación de datos del reporte diario y semanal

Observación de proceso					
Fecha:	01/09/2022	Hora inicio:	9:15 am	Hora final:	9:23 am
Proceso observado:	Elaboración de reportes				
Participantes					
Nombre			Rol dentro del proceso		
Diego Alfaro			Desarrollador de inteligencia de negocios		
Detalles					
Descripción de la actividad	En esta actividad del proceso se modifican algunos formatos que traen los archivos				
Pasos:	<ol style="list-style-type: none"> 1. Se abren los archivos 2. Se realiza cambios de formato en las fechas 				
Comentarios	Esta actividad aplica para todos los reportes debido a que en herramientas como AODOcs y Genesis la fecha está en formato DD/MM/YYYY cuando debe de ser MM/DD/YYYY.				

9.12. Apéndice L. Resultados de la observación – Unión de datos para la creación de reportes

Observación de proceso					
Fecha:	01/09/2022	Hora inicio:	9:25 am	Hora final:	9:40 am
Proceso observado:	Elaboración de reportes				
Participantes					
Nombre			Rol dentro del proceso		
Diego Alfaro			Desarrollador de inteligencia de negocios		
Detalles					
Descripción de la actividad	En esta actividad se realiza la unión manual de documentos				
Pasos:	<ol style="list-style-type: none"> 1. Abrir los documentos. 2. Crear una nueva hoja en el documento principal 3. Pegar los datos del otro reporte 				
Comentarios	Esta actividad aplica para todos los reportes.				

9.13. Apéndice M. Resultados de la observación – Realización de cálculos y emparejar datos para la creación de reportes

Observación de proceso				
Fecha:	01/09/2022	Hora inicio:	9:45 am	Hora final: 10:20 am
Proceso observado:	Elaboración de reportes			
Participantes				
Nombre			Rol dentro del proceso	
Diego Alfaro			Desarrollador de inteligencia de negocios	
Detalles				
Descripción de la actividad	En esta actividad se emparejan los datos según sea necesario y se realizan los cálculos necesarios para cada reporte.			
Pasos:	<ol style="list-style-type: none"> 1. Se crea una hoja nueva donde se va a realizar los cálculos. 2. Utilizando la función “vlookup” se emparejan los datos de las distintas hojas en esta. 3. Se realizan los cálculos de vejez de casos utilizando las fechas. 4. Con ayuda de tablas pivot se extraen valores como: <ol style="list-style-type: none"> a. Casos totales por tipo b. Casos totales por agente 			
Comentarios	Esta actividad aplica para todos los reportes. La duración de esta puede varias dependiendo la cantidad de datos.			

9.14. Apéndice N. Resultados de la observación – Actualizar fuentes de datos de los reportes.

Observación de proceso			
Fecha:	01/09/2022	Hora inicio:	10:30 am
Hora final:	10:55 am		
Proceso observado:	Elaboración de reportes		
Participantes			
Nombre		Rol dentro del proceso	
Diego Alfaro		Desarrollador de inteligencia de negocios	
Detalles			
Descripción de la actividad	En esta actividad se realiza la actualización de las fuentes de datos que cargan los reportes ya creados		
Pasos:	<ol style="list-style-type: none"> 1. Se accede a la fuente de datos consumida por el reporte. Estas son: <ol style="list-style-type: none"> a. Hoja de cálculos de Google. <ol style="list-style-type: none"> i. Se copia los datos obtenidos en la hoja de Excel en la hoja de Google. ii. Se actualiza el reporte de forma manual. b. Base de datos PostgreSQL <ol style="list-style-type: none"> i. Se coloca el documento en formato CSV en una carpeta especifica en el servidor, con un nombre determinado. ii. Se ejecuta una tarea la cual carga los datos a la base de datos. 		
Comentarios	Esta actividad aplica para todos los reportes. La duración de esta puede varias dependiendo la cantidad de datos ya que deben de ser cargados a una hoja de Google especifica en algunos casos.		

9.15. Apéndice O. Resultados de entrevista #1 - Descubrimiento de la organización.

Detalles de entrevista					
Fecha:	02/06/2022	Hora inicio:	10:00 am	Hora final:	10:30 am
Objetivo:	Conocer sobre Equifax				
Participantes					
Nombre			Rol dentro del proceso		
Luis Vindas			Gerente del departamento de Servicios de Datos		
Preguntas			Respuestas		
¿Cómo se podría definir Equifax?			Equifax se puede definir como una organización transnacional que brinda servicios de análisis de datos		
¿Cómo se podría definir el departamento de Servicios de Datos?			Un equipo que brinda servicios a la parte operativa de Equifax, es decir, ayudar a los departamentos que generan ingresos a la organización por medio de servicios de análisis de datos.		
¿Cuáles clientes tienen?			Actualmente, todos los clientes provienen de operaciones. En su mayoría son de US y Canadá, pero se brinda soporte a los equipos de servicio al cliente de Costa Rica y finanzas internacionales.		
¿Qué servicios brindan a los clientes?			Se brindan servicios de: <ul style="list-style-type: none"> • Creación y gestión de reportes. • Gestión de bases de datos. 		
¿Cuáles han sido los mayores retos del departamento?			Debido al crecimiento que ha tenido el departamento uno de los mayores retos ha sido la estandarización y automatización de procesos		
¿Qué herramientas son utilizadas?			Actualmente el departamento cuenta con: <ul style="list-style-type: none"> • Tableau • Bamboo • Pentaho • Excel • GCP 		
¿Cuáles herramientas se planean para el equipo?			Debido a que se va a prohibir el uso de algunas herramientas como Bamboo y Pentaho, se plantea el uso de Python ya que es uno de los lenguajes más funcionales y adaptables para la gestión de datos.		
¿Cuál es la dirección en la cual va el departamento?			Primeramente, crecer como equipo para brindar servicio a una mayor cantidad de clientes. Esto de la mano de procesos automatizados que faciliten las actividades diarias a los clientes y a nuestro propio equipo.		

9.16. Apéndice P. Resultados de entrevista #2 - Descubrimiento del proceso.

Detalles de entrevista					
Fecha:	07/09/2022	Hora inicio:	2:00 pm	Hora final:	3:00 pm
Objetivo:	Conocer sobre el proceso de generación de reportes para los departamentos de Business Care y Boarding				
Participantes					
Nombre		Rol dentro del proceso			
Diego Alfaro		Desarrollador de Inteligencia de negocios			
Luis Vindas		Gerente del departamento de Servicios de Datos			
Preguntas		Respuesta			
¿Cuál es el rol de Business Care y Boarding dentro de Equifax?		El departamento de Business Care se encarga del servicio a los clientes, todo lo relacionado a los call centers. Por otro lado, el departamento de Boarding se encarga de la gestión de los clientes, es decir, acreditar a los clientes cuando quieran consumir o adquirir productos de Equifax. Estos realizan una serie de actividades las cuales son necesarias para validar al cliente y los motivos por los cuales utilizara los datos brindados por Equifax. Estos departamentos actualmente son uno solo ya que fueron unidos el año pasado y ahora son gestionados por la misma persona.			
¿Qué tipo de reportes se generan para estos equipos?		Para estos departamentos los reportes que se generan son resúmenes en tablas o “dashboards” los cuales contienen información de los tiquetes creados por los agentes. La mayoría de estos reportes contienen tiempos, ya sea tiempo que lleva abierto un tiquete o tiempo que dura un agente en resolverlo. También se hacen reportes que contienen resúmenes semanales, mensuales o trimestrales los cuales tienen a una vista macro las operaciones de cada uno de estos departamentos.			
¿Cuántos reportes son generados?		En total, sumando los diarios, semanales, mensuales y trimestrales, a este departamento se le realizan un total de dieciséis reportes. Nota: Diego enviara descripción y fuentes de estos.			
¿Quién son los puntos de contacto?		La persona encargada de estos departamentos es Kimberly Long, pero el principal punto de contacto es Kevin Durio.			
¿Cuáles fueron los requerimientos principales?		Con este departamento no se ha realizado un proceso de requerimientos, la mayoría de estos reportes son solicitados y explicados en una misma llamada.			
¿Existe documentación del proceso?		No, realmente esto ha sido uno de los fallos que se ha tenido desde el inicio. Nunca se ha tomado el tiempo de documentar el proceso.			
¿Cuáles son las fuentes de datos utilizadas?		Las principales fuentes de datos son las dos versiones de Salesforce existentes, pero también se obtienen datos de Genesis, AODOcs y mucha información es enviada en archivos de Excel por correo electrónico.			
¿Cómo es el proceso actual?		Actualmente el proceso es realizado de forma manual. Se accede a la fuente de información, se descarga utilizando uno de los reportes disponibles en la herramienta y son arreglados a nivel local utilizando Excel y posteriormente se genera el reporte.			
¿Cómo se reciben las solicitudes de nuevos reportes?		Por medio de una llamada Kevin o alguien del departamento menciona el deseo de un nuevo reporte y explican cómo debe de ser este.			
¿Cuáles opciones de mejora ven en el proceso de creación de reportes?		La automatización del proceso de extracción de datos. Enviar los reportes de forma automática Actualizar las fuentes de datos automáticamente.			
¿Qué herramientas existen dentro de la organización que pueden ser explotadas?		Muchas herramientas cuentan con APIs las cuales pueden ser consultadas, excepto Salesforce. Ya que se está en proceso de migración de información a BigQuery para acceder a los datos utilizando esta herramienta.			

por el equipo de Servicios de Datos?	
--------------------------------------	--

9.17. Apéndice Q. Resultados de entrevista #3 – Uso y requerimientos del proceso para el cliente.

Detalles de entrevista					
Fecha:	12/09/2022	Hora inicio:	8:00 am	Hora final:	8:30 am
Objetivo:	Conocer el funcionamiento y requerimientos del cliente				
Participantes					
Nombre	Rol dentro del proceso				
Kevin Durio	Gerente del departamento de Business Care				
Kimberly Long	Vicepresidenta del departamento de Business Care y Boarding				
Preguntas	Respuestas				
¿Cuáles son las expectativas de los reportes?	La expectativa es que estos sean correctos y certeros. Que sean actualizados de forma diaria. Que cumplan satisfagan las necesidades por las cuales fueron creadas.				
¿Cuáles son los usos que le dan a los reportes?	<ul style="list-style-type: none"> - Los reportes diarios son utilizados por los gerentes de cada uno de los equipos para verificar el día a día de las operaciones. - Los semanales son utilizados en reuniones entre gerentes para validar el trabajo realizado y progreso tanto de los agentes como del equipo. - Los mensuales son enviados en modo de presentación para resúmenes a gerentes de más alto nivel. - El reporte trimestral es necesario para identificar cuales clientes son los que deben de Re acreditar durante el siguiente trimestre y a su vez sirve como método de control de clientes. 				
¿Cuáles han sido los principales problemas que han existido con estos reportes?	<ul style="list-style-type: none"> - Con los reportes diarios, el problema que se ha presentado es con los datos de cantidades de tiquetes realizados por día. - Con los reportes semanales, se han presentado problemas con algunos casos, ya que, por política de la organización, si un caso lleva más de días abierto debe ser abandonado y en muchas ocasiones se abandonan cuando no deben. - El trimestral es el que más problemas presenta, esto debido a que es un reporte relativamente nuevo. Las fuentes de datos muchas veces no son certeras y causa problemas en el reporte final, este reporte es uno de los más importantes ya que no es solo de consumo interno sino externo también. Esto debido a que es auditado por el CFPB y debido a fallos puede generar multas a la organización. 				
¿Qué les gustaría mejorar de los reportes?	<p>Nos gustaría que el reporte diario tenga menos errores humanos y estuviera disponible todos los días a la misma hora.</p> <p>Con respecto a los demás, nos gustaría que su ejecución sea automática ya que muchas veces hay que solicitarlo al equipo, lo cual causa atrasos en el proceso de crear las presentaciones.</p> <p>Con respecto al reporte trimestral, nos gustaría que mejorara el proceso ya que este utiliza una herramienta la cual no es conocida por muchos (SAS) y los recursos que presentan este conocimiento suelen ser muy caros y utilizados por periodos cortos de tiempo.</p>				

9.18. Apéndice R. Resultados de entrevista #4 – Limitaciones y requerimientos tecnológicos del departamento

Detalles de entrevista

Fecha:	12/09/2022	Hora inicio:	1:00pm	Hora final:	2:00pm
Objetivo:	Conocer los requerimientos tecnológicos del departamento				
Participantes					
Nombre			Rol dentro del proceso		
Mateo Marín			Encargado de soporte y desarrollo del departamento de servicio de datos		
Preguntas					
¿Cuál ha sido su rol dentro de la creación de reportes?			Brindo soporte a al equipo en temas de gestión de bases de datos, cree el ETL que carga los reportes a la base de datos.		
¿A cuáles herramientas le brinda soporte?			Relacionados al proceso: <ul style="list-style-type: none"> • Pentaho • PostgreSQL • Bamboo • GCP 		
¿Cuáles herramientas se encuentran disponibles para utilizar?			Debido a que se realizó un cambio en las políticas de la organización en términos de licencias, se encuentran disponibles herramientas como Python, BigQuery, Airflow y PostgreSQL.		
¿Cuál piensa usted que debe ser el siguiente paso del proceso de creación de reportes?			La estandarización pienso que es el siguiente paso. Muchos de estos reportes no tienen ni una documentación por lo que no existe una forma de hacer ETL de forma sencilla. Por esta razón, muchos optaron por hacerlo manual.		
¿Qué herramientas recomienda para estos siguientes pasos?			Para realizar estos procesos, la herramienta más factible para hacer ETL sería Pentaho o tratar de negociar el uso de Visual Studio y el SSIS. He oído que una de las formas que Luis y Gerardo pretenden hacer este proceso es usando Python, pero en mi opinión es más factible utilizar una herramienta ya diseñada para esto.		

9.19. Apéndice S. Resultados de Matriz *Two-by-Two*

Alto	Impacto y valor para el cliente	<ul style="list-style-type: none"> • Actualización de los datos en los reportes de forma automática. • Mejorar las visualizaciones utilizadas en vez de tablas. 	<ul style="list-style-type: none"> • Extracción automatizada de datos. • Calendarización del ETL 	
		<ul style="list-style-type: none"> • Documentar el proceso de creación de reportes 	<ul style="list-style-type: none"> • Brindar acceso a los reportes a todos los usuarios. • Uso de herramientas especiales para ETL. 	
Bajo	Costo y nivel complicación			Alto

9.20. Apéndice T. Resultados de la hoja de tiempos del proceso.

Hoja de tiempo en el proceso					
Paso #	Tiempo de procesamiento	Tiempo de espera	Tiempo de revisión	Tiempo de retrabajo	Tiempo total.
1.	15 min	18 min	2 min	0 min	35 min
2	10 min	3 min	3 min	10 min	26 min
3	13 min	3 min	2 min	0 min	18 min
4	10 min	8 min	0 min	10 min	28 min
5	5 min	5 min	0 min	0 min	10 min

9.21. Apéndice U. Resultado de análisis revisión documental #1

Revisión Documental			
Fecha:	18/09/2022	Fuente:	Process Mapping, Process Improvement, and Process Management
Objetivo	Mejora de procesos según Dan Madison		
Detalles			
<ol style="list-style-type: none"> 1. Los resultados óptimos de mejora de procesos ocurren cuando hay una participación sustancial de los empleados. El principio de implicación de los trabajadores se basa en dos premisas: <ol style="list-style-type: none"> a. Los empleados conocen mejor el trabajo y, si se les da la oportunidad, contribuirán con grandes ideas para mejorar b. Los empleados involucrados en el desarrollo de recomendaciones de cambio apoyarán, no resistirán, su implementación. 2. Un esfuerzo de mejora de procesos comienza con la búsqueda de procesos rotos. 			

9.22. Apéndice V. Resultado de análisis revisión documental #2

Revisión Documental			
Fecha:	19/09/2022	Fuente:	Business Process Improvement
Objetivo	Mejora de procesos según Susan Page		
Detalles			
<ol style="list-style-type: none"> 1. En términos de mejora de procesos estos 3 conceptos son de suma importancia: <ol style="list-style-type: none"> a. Eficacia b. Eficiencia c. Adaptabilidad 2. Es necesario identificar las actividades y participantes en cada uno de los pasos del proceso. 3. Se debe desarrollar un plan para conocer las herramientas y los materiales que necesita. 4. Es importante definir controles internos, métricas del desarrollo y tener herramientas que respalden el proceso. 			

9.23. Apéndice W. Resultado de análisis revisión documental #3

Revisión Documental			
Fecha:	20/09/2022	Fuente:	AWS
Enlace:	https://aws.amazon.com/blogs/big-data/top-8-best-practices-for-high-performance-etl-processing-using-amazon-redshift/		
Objetivo	Mejores prácticas de ETL		
Detalles			
<ol style="list-style-type: none"> 1. Copiar datos de múltiples archivos de tamaño uniforme. 2. Utilice la gestión de la carga de trabajo para mejorar los tiempos de ejecución de ETL. 3. Realice el mantenimiento de la tabla regularmente. 4. Realice múltiples pasos en una sola transacción. 5. Cargar datos de forma masiva. 6. Supervise la salud diaria de ETL mediante consultas de diagnóstico. 			

9.24. Apéndice X. Resultado de análisis revisión documental #4

Revisión Documental			
Fecha:	20/09/2022	Fuente:	IBM
Enlace:	https://www.ibm.com/docs/en/cognos-analytics/10.2.2?topic=preparation-etl-scenarios-best-practices		
Objetivo	Mejores prácticas de ETL		
Detalles			
<ol style="list-style-type: none"> 1. La carga de datos en tablas organizadas en columnas es muy similar a la carga de datos en tablas organizadas en filas, con algunas excepciones. 2. Carga de datos y dimensiones atómicos 3. Usar procedimientos almacenados para cargar tablas agregadas 4. Un procedimiento almacenado realiza tanto la carga inicial como la carga incremental para una tabla agregada. Si la tabla agregada está vacía, se realiza una carga inicial. Si hay datos en la tabla agregada, se realiza una carga incremental. 			

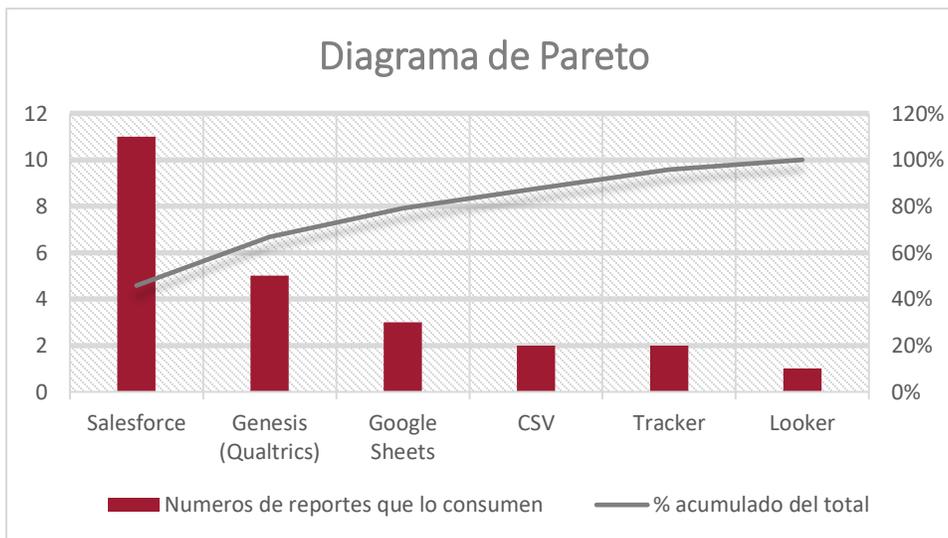
9.25. Apéndice Y. Resultado de análisis revisión documental #5

Revisión Documental	
Fecha:	20/09/2022
Fuente:	The Analytics Setup Guidebook
Objetivo	Metodología Kimball para la integración de datos
Detalles	
<ol style="list-style-type: none"> 1. El esquema en estrella es una forma particular de organizar datos con fines analíticos. Costa de dos tipos de tablas: <ol style="list-style-type: none"> a. Una tabla de hechos, que actúa como tabla principal para el esquema. Una tabla de hechos contiene las medidas, métricas o "hechos" principales de un proceso empresarial. b. Muchas tablas de dimensiones asociadas con la tabla de hechos. Cada tabla de dimensiones contiene 'dimensiones', es decir, atributos descriptivos de la tabla de hechos. 2. La respuesta de Kimball a eso es el proceso de cuatro pasos para el modelado de datos dimensionales. Estos cuatro pasos son los siguientes: <ol style="list-style-type: none"> a. Elija un proceso de negocio para modelar. El enfoque de Kimball comienza con un proceso comercial, ya que, en última instancia, los usuarios comerciales querrían hacer preguntas sobre los procesos. b. Decidir la granularidad. La granularidad significa el nivel de datos para almacenar como la tabla de hechos principal. Debe ser el nivel más atómico posible, es decir, un nivel de datos que no se pueda dividir más. c. Elija las dimensiones que se aplican a cada fila de la tabla de hechos. Esto suele ser bastante fácil de responder si ha 'recogido el grano' correctamente. d. Identifique los hechos numéricos que llenarán cada fila de la tabla de hechos. Los datos numéricos en la tabla de hechos quedan fuera de la pregunta "¿qué estamos respondiendo?". 3. Mapeo de relaciones: un mapeo de relaciones es una relación externa entre dos modelos de datos. 4. Modelos creados sobre otros modelos: en esta etapa, simplemente hemos demostrado que los modelos de datos le permiten anotar y enriquecer las tablas existentes en su almacén de datos. 	

9.26. Apéndice Z. Resultado de análisis revisión documental #6

Revisión Documental	
Fecha:	Fuente: Tableau
Enlace:	https://www.tableau.com/learn/articles/data-visualization-tips
Objetivo	Mejores prácticas para el diseño de dashboards
Detalles	
<ol style="list-style-type: none"> 1. Elija los cuadros y gráficos correctos para la visualización. <ol style="list-style-type: none"> a. Los gráficos de barras son efectivos para comparar categorías dentro de una sola medida y una de las visualizaciones de datos más comunes. Son especialmente eficaces cuando tiene datos que se pueden dividir en varias categorías. b. Los gráficos de viñetas muestran el progreso con respecto a un objetivo mediante la comparación de medidas y se diseñaron para reemplazar los indicadores, medidores y termómetros del tablero. c. El gráfico de líneas conecta varios puntos de datos distintos, presentándolos como una evolución continua. El resultado es una forma sencilla y directa de visualizar los cambios en un valor en relación con otro. d. Los histogramas y los diagramas de caja muestran dónde se agrupan sus datos y pueden comparar categorías. e. Los mapas son una obviedad para visualizar preguntas específicas de la ubicación o para ayudar en la exploración geográfica. f. Los gráficos circulares son útiles para agregar detalles a otras visualizaciones, pero no son tan efectivos por sí solos. De hecho, cada año, en el Día Pi (14 de marzo), el científico de datos Kaiser Fung y su colega encuentran un gráfico circular y lo transforman en una versión más adecuada, como forma de protestar contra el uso excesivo flagrante. 2. Usa patrones predecibles para los diseños 3. Cuente historias de datos rápidamente con señales de color claras 4. Incorpora pistas contextuales con formas y diseños. 5. Usa estratégicamente el tamaño para visualizar valores 	

9.27. Apéndice AA. Diagrama de Pareto



9.28. Apéndice BB. Retorno de Inversión

ROI =	$\frac{((\text{Ingresos} - \text{Inversión}) / \text{Inversión}) * 100}{}$
Datos =	Ingresos = \$24.450
	Inversión = 10,728,424.4
	Tipo de cambio = 631.1
	Periodo= 2 años
ROI =	81.8%

9.29. Apéndice CC. Documentación del proceso



Reportes de *“Business Care”* y *“Boarding”*

Documentación



Introducción

El presente documento cuenta con la documentación de los reportes generados para los departamentos de “Business Care” y “Boarding”. En este se describen los reportes realizados, las fuentes de datos y accesos requeridos para la ejecución de estos, herramientas necesarias para su realización, una descripción de la ejecución del proceso automatizado, el estándar para generar reportes y unas pruebas de concepto de las fuentes que son comúnmente utilizadas. Cabe destacar que para esta versión, existen cierta información la cual no se muestra por motivos de seguridad de la organización.

Reportes

En esta sección del documento se encuentran descritos los reportes generados para estos departamentos. A continuación, se enumeran los reportes y se detalla su uso, fuentes y periodicidad de ejecución. Para consultar estos reportes dirijase a tableau server > CRSSC.

1) VarVab

a) Descripción

Este reporte muestra la información de los casos creados por cada uno de los revendedores de información de Equifax. Estos generan casos de tipo: Acreditación, mantenimiento, nuevos productos, nuevos clientes.

b) Periodicidad

Debido a que estos casos deben ser concluidos lo más rápido posible por diferentes contratos de nivel de servicio, es necesario brindar un seguimiento diario a los supervisores.

c) Fuentes

La información requerida para la ejecución de este reporte se encuentra en Salesforce (CRM). Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery.

2) Productividad (Acreditación, provisionamiento)

a) Descripción

Este reporte muestra los casos completados por los agentes de forma diaria. Dentro del reporte se muestran los casos completados por: causa, tipo y equipo.

b) Periodicidad

El reporte debe ser actualizado de forma diaria para captar el estado de todos los casos asignados al final de cada día laboral.

5) Casos NCC

a) Descripción

Este reporte corresponde a uno de los clientes revendedores más grandes de la organización, en este se muestran los casos abiertos, el estado del caso, el agente encargado, compañía asociada, agente asociado y tiempo que lleva abierto el tiquete.

b) Periodicidad

Debido a que este es uno de los revendedores más grandes de la organización, se realiza de forma diaria para atender cada tiquete que estos genera.

c) Fuentes

La información requerida para la ejecución de este reporte se encuentra en Salesforce (CRM). Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery.

6) Calificación de los agentes

a) Descripción

Este reporte muestra la nota promedio semanal asignada a los agentes por gerente. Este reporte se basa en cálculos utilizando: cantidad de casos resueltos, tiempo promedio en llamada, tiempo promedio en contestar y calificación del cliente.

b) Periodicidad

Este reporte se genera de forma semanal, debido a que los agentes tienen derecho hacer reclamos semanales con respecto algún caso en concreto.

c) Fuentes

La fuente de datos de este reporte proviene de Salesforce y Qualtrics. Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery. Con respecto a la información de Qualtrics, es extraída automáticamente de forma diaria directamente desde el API y almacenada en la base de datos PostgreSQL.

7) Casos de mantenimiento

a) Descripción

Este reporte muestra todos los casos de mantenimiento de cuentas realizados mensuales. Muestra el tiempo de ejecución promedio, cantidad de tiempo enviado al estatus "Remediation" y tiempo que llevan abierto los casos.

b) Periodicidad

Este reporte se genera de forma mensual, debido a que es consumido por los gerentes para la revisión de resultados mensuales.

c) Fuentes

La información requerida para la ejecución de este reporte se encuentra en Salesforce (CRM). Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery.

8) Nuevos clientes

a) Descripción

En este reporte se muestran todos los clientes nuevos que han sido acreditados durante el mes sin importar el tipo (Directo o revendedor).

b) Periodicidad

Este reporte se genera de forma mensual, debido a que es consumido por los gerentes para la revisión de resultados mensuales.

c) Fuentes

La información requerida para la ejecución de este reporte se encuentra en Salesforce (CRM). Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery.

9) Operaciones mensuales (comercial)

a) Descripción

Este reporte muestra todos los casos que han sido procesados, cerrados o abandonados. También se muestra el promedio del tiempo de respuesta de los agentes y el nivel de satisfacción de los clientes. En el caso del comercial, se muestra el volumen de llamadas recibidas, la calidad asignada a la llamada, tiempo de espera del cliente, tiempo promedio de duración en la llamada y promedio de la satisfacción del cliente.

b) Periodicidad

Este reporte se genera de forma mensual, debido a que es consumido por los gerentes para la revisión de resultados mensuales.

c) Fuentes

La fuente de datos de este reporte proviene de Salesforce y Qualtrics. Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery. Con respecto a la información de Qualtrics, es extraída automáticamente de forma diaria directamente desde el API y almacenada en la base de datos PostgreSQL. Por otro lado, la información de tracker ya se encuentra cargada en la misma base de datos.

10) ACRO

a) Descripción

En este reporte se muestran las consultas generadas por clientes al mes, cuantas fueron resultas, tiempo promedio de resolución, volumen por tipo, porcentaje que fue atendido por el bot y el porcentaje que fueron atendidos por un agente.

b) Periodicidad

Este reporte se genera de forma mensual, debido a que es consumido por los gerentes para la revisión de resultados mensuales.

c) Fuentes

La fuente de datos de este reporte proviene de Salesforce y Qualtrics. Para acceder a esta es necesario contar accesos al proyecto de EDH en GCP, debido a que la información se encuentra en BigQuery. Además, los datos de las consultas realizadas por bots es enviado por un mail, este es extraído y cargado a la base de datos de forma automática.

11) Reacreditación

a) Descripción

Este reporte se utiliza para saber cuáles clientes deben ser Re acreditados durante el siguiente trimestre. Cada uno de los clientes son asignados a un nivel (T1-T4) y si estos son directos o indirectos.

b) Periodicidad

Este reporte consume información generada por los dos trimestres anteriores y se realiza de forma trimestral.

c) Fuentes

Este reporte datos de cuatro fuentes distintas:

- AODocs: se encuentra cargada en la base de datos gracias a la extracción diaria.
- Looker: es enviada mensualmente y almacenada en la base de datos PostgreSQL de forma automática.
- Indirectos: es enviada trimestralmente y almacenada en la base de datos PostgreSQL de forma automática.

Herramientas

Para la ejecución del proceso de reportes se utilizan varias herramientas las cuales permiten la extracción, transformación y almacenamiento de datos. La mayoría de las tareas se realizan por medio de Python gracias a su flexibilidad y fácil comprensión. Por otro lado, se utiliza una base de datos PostgreSQL para almacenar los datos en una base de datos la cual es gestionada por el equipo, de esta forma el uso de esta se torna sencillo.

- Python

Para la instalación de Python es necesario contar con los permisos de administrador. Para esto diríjase al TechBar para solicitarlos. Una vez cuente con los permisos, acceda al proyecto GCP "DataServices_Tracker" en el cual se encuentra el instalable en la carpeta de "Apps". Una vez instalado Python, las librerías necesarias dentro del proceso son:

Librería	Descripción	Instalación
Csv	Permite crear, leer y escribir sobre archivos de tipo csv.	pip install python-csv
Json	Permite realizar operaciones de lectura y escritura en formato json.	pip install jsonlib
Dateutil	Permite realizar distintas transformaciones en valores de fechas.	pip install python-dateutil
Numpy	Permite trabajar con arreglos dentro de Python para la manipulación de matrices y realizar operaciones algebraicas.	pip install numpy
Google api	Permite hacer uso de las herramientas de Google a través de Python.	pip install --upgrade google-api-python-client google-auth-httpplib2 google-auth-oauthlib
Pandas	Permite realizar análisis de datos a través de Python por medio de métodos especializados.	pip install panda
Psycopg2	Permite realizar conexiones con bases de datos PostgreSQL.	pip install psycopg2
regex	Permite realizar análisis por medio de expresiones regulares (Regex).	pip install regex


```
# Fetches db info from Prod.properties file and returns a connection string ready to be used
def get_db_connection_string(db_name):
    # Load DB properties file
    properties_db_file = open('credentials/Prod.properties')
    db_connection = 'dbname=[dbname] user=[user] password=[password] host=[host] port=[port]'

    host_name = 'DB.' + db_name + '.hostname'
    port = 'DB.' + db_name + '.port'
    name = 'DB.' + db_name + '.Name'
    user = 'DB.' + db_name + '.User'
    password = 'DB.' + db_name + '.Password'

    # Insert DB info into connection string
    for line in properties_db_file:
        line = line.strip().split('=')
        if len(line) < 2:
            continue
        if line[0] == host_name:
            db_connection = db_connection.replace('[host]', line[1])
        elif line[0] == port:
            db_connection = db_connection.replace('[port]', line[1])
        elif line[0] == name:
            db_connection = db_connection.replace('[dbname]', line[1])
        elif line[0] == user:
            db_connection = db_connection.replace('[user]', line[1])
        elif line[0] == password:
            db_connection = db_connection.replace('[password]', decrypt_password(line[1]))

    properties_db_file.close()

    # Verify connection string
    if '[' in db_connection:
        sys.exit('Invalid DB properties, check Prod.properties file')

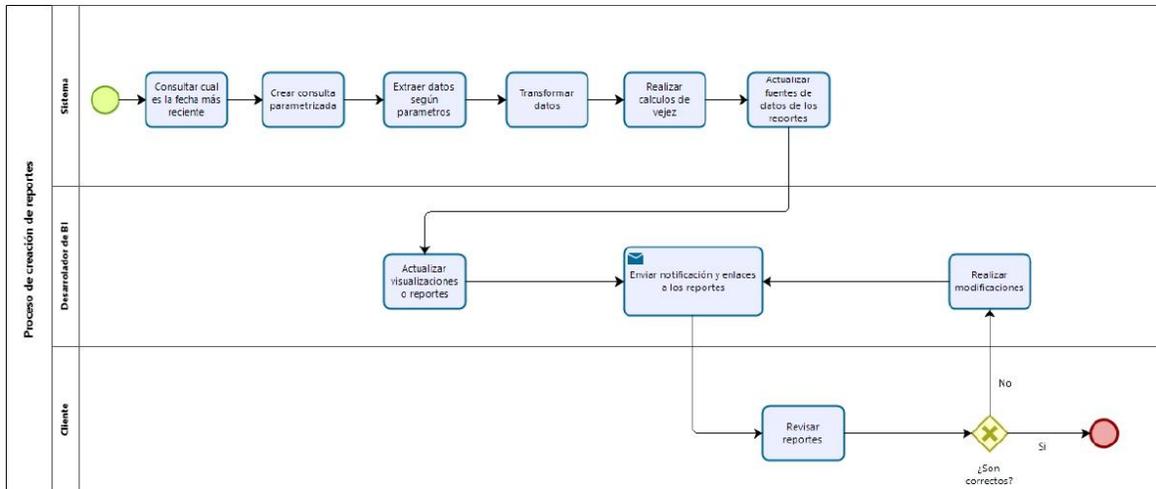
    return db_connection
```

3. Genere la conexión.

```
# Create db connection
db_connection = oct_utils.get_db_connection_string('OPS_Metrics_PSQL')
connection = psycopg2.connect(db_connection)
cursor = connection.cursor()
```

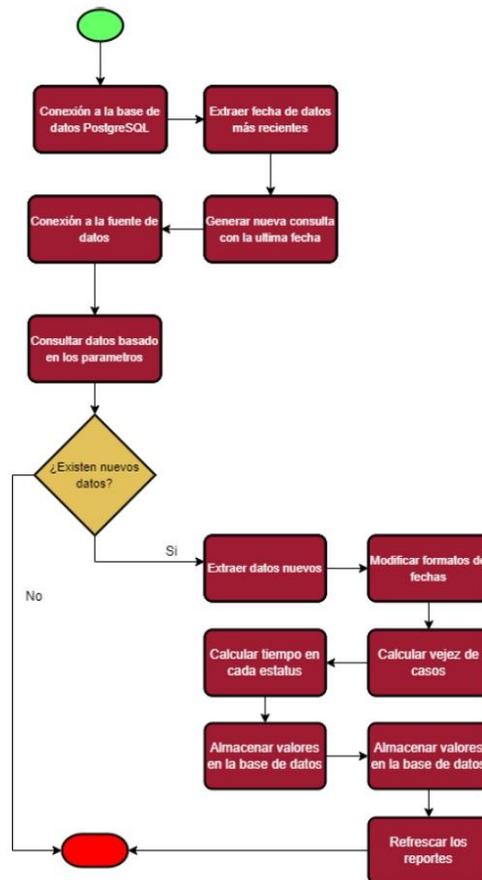
Funcionamiento

En la figura que se encuentra a continuación se ejemplifica como es el proceso de realizar reportes. En este se aprecia como la mayoría de las tareas deben ser calendarizadas y realizadas por el sistema sin la necesidad de una interacción humana. De esta forma se disminuye el error por manipulación de datos.



Por otro lado, en la figura a continuación se presenta el diagrama de flujo de la automatización del proceso. En este vemos como el sistema realiza una primera consulta a la base de datos para extraer la última fecha que encuentre, esto con el fin de realizar la consulta a las distintas fuentes basado en este valor. De esta forma las consultas realizadas por medio de "BigQuery" consumen una menor cantidad de procesamiento, disminuyendo el costo. En el caso del API de qualtrics, utilizamos este valor para no sobrescribir valores ya almacenados.

Una vez se extraen los valores, es importante realizar las transformaciones de datos necesarias y los cálculos, para esto se utiliza python para la estandarización de fechas, valores numéricos, eliminación de espacios en blancos y para la validación de IDs por medio del regex. Ejecutada esta tarea de transformación y limpieza se realizan los cálculos necesarios, como calcular vejez de casos, vejez por fases y por tipos, para finalmente cargarlos a la base de datos de esta forma se realizan las visualizaciones con valores ya calculados favoreciendo el desempeño.



Estándar de visualizaciones

Con respecto a la visualización de los datos, se utiliza un estándar para ajustarse a los estándares organizacionales. Como se aprecia en la figura a continuación se insta a la creación de reportes o “dashboards” de una forma sencilla. En la cual se muestran indicadores importantes para el cliente y un numero limitado de visualizaciones. De esta forma se obliga al usuario a optimizar los datos seleccionados y realizar gráficos representativos y que agreguen valor al cliente. En el estándar encontramos un dashboard de dos hojas, pero este puede llegar a ser de solo una si así lo desea el cliente.



Color para Texto, Graficos, figuras, etc.

Primary color palette

Equifax Red is the primary color that represents the brand identity of Equifax and the Equifax logo. It is paired with solid black, white, and a family of grays to make up the primary palette.

Always use Equifax Red at 100% opacity. Never screen or tint, create a gradient or transparency, or use any other effect that takes away from its true color integrity.

<p>Equifax Red</p> <p>PMS 201 CMYK 24 / 99 / 78 / 18 RGB 158 / 27 / 50 HEX #9e1d32</p>	<p>Black</p> <p>PMS Black CMYK 0 / 0 / 0 / 100 RGB 0 / 0 / 0 HEX #000000</p>	<p>White</p> <p>CMYK 0 / 0 / 0 / 0 RGB 255 / 255 / 255 HEX #ffffff</p>
<p>Gray</p> <p>PMS 431 CMYK 67 / 52 / 44 / 17 RGB 91 / 103 / 113 HEX #5a6771</p>	<p>Medium Gray</p> <p>PMS 429 CMYK 38 / 27 / 23 / 0 RGB 163 / 170 / 173 HEX #a3aaa1</p>	<p>Light Gray</p> <p>PMS Black 106 CMYK 8 / 6 / 6 / 0 RGB 231 / 231 / 231 HEX #e6e7e7</p>

Secondary color palette

The secondary color palette is a complement to the primary color palette. It is made up of five accent colors used to enhance a design layout.

Secondary colors should always:

- Appear at 100 percent whenever possible, but may be used as a screen in certain circumstances
- Be used with Equifax Red, never independently
- Be only used as accents, never overpowering the primary colors
- Be used sparingly and with purpose; if they fail to add value or enhance the reader's understanding of information, avoid adding color for the sake of adding it

Secondary colors should not:

- Be used to represent business units, industries, or product lines
- Be used randomly; always limit the use of secondary colors

NOTE: For the most optimal reader visibility, avoid using yellow for type and limit the amount of white type on orange.

<p>Orange</p> <p>PMS 152 CMYK 6 / 67 / 100 / 1 RGB 231 / 114 / 4 HEX #e72c94</p>	<p>Green</p> <p>PMS 303 CMYK 73 / 24 / 100 / 9 RGB 93 / 69 / 132 / 42 HEX #588d2a</p>	<p>Blue</p> <p>PMS 633 CMYK 90 / 47 / 25 / 3 RGB 0 / 114 / 152 HEX #007298</p>
<p>Purple</p> <p>PMS 525 CMYK 69 / 95 / 22 / 13 RGB 101 / 47 / 188 HEX #6a279c</p>	<p>Yellow</p> <p>PMS 3406 CMYK 6 / 22 / 100 / 0 RGB 241 / 195 / 25 HEX #f1e319</p>	

Lo importante de este estándar es seguir las siguientes recomendaciones:

- **Granularidad:** Las visualizaciones, por lo general, se realizan de acuerdo con el movimiento natural del ojo. Este movimiento es de arriba hacia abajo y de izquierda a derecha. Debido a esto, se prioriza las visualizaciones de mayor importancia al lado superior izquierdo para posteriormente ir colando las de menor relevancia.
- **Estructura:** Se establecieron cuatro principales secciones dentro del *dashboard* las cuales son:
 - **Información:** En esta sección se coloca el nombre del reporte con una breve descripción de este. Acompañado de información de la organización y fechas.
 - **Filtros:** En esta sección se colocan los filtros que permitan una mejor interacción con los datos al usuario. Debido a tamaño del *dashboard* se estableció un máximo de cinco filtros.
 - **Indicadores:** En esta sección se colocan los indicadores de mayor impacto para el *dashboard*. En general, se colocan valores numéricos los cuales muestran totales, porcentajes, promedios, etc.
 - **Visualizaciones:** En esta sección se colocan los gráficos que conforman el *dashboard*, esto siguiendo la misma lógica anteriormente mencionada.
- **Paleta de colores:** Con respecto a la paleta de colores, se definió utilizar la establecida en el libro de marca de la organización. Este con el fin de ajustarse a las políticas de la organización.

▪ Gráficos:

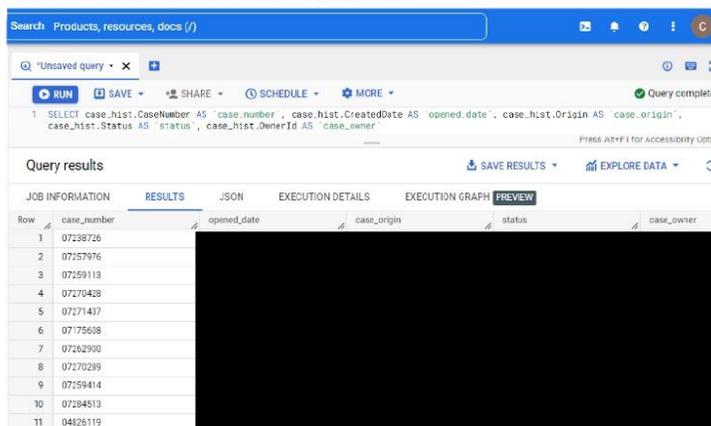
- Gráficos de barras: son efectivos para comparar categorías dentro de una sola medida y una de las visualizaciones de datos más comunes. Son especialmente eficaces cuando tiene datos que se pueden dividir en varias categorías.
- Gráficos de viñetas: muestran el progreso con respecto a un objetivo mediante la comparación de medidas y se diseñaron para reemplazar los indicadores, medidores y termómetros del tablero.
- Gráfico de líneas: conectan varios puntos de datos distintos, presentándose como una evolución continua. El resultado es una forma sencilla y directa de visualizar los cambios en un valor en relación con otro.
- Los histogramas y los diagramas de caja muestran dónde se agrupan sus datos y pueden comparar categorías.
- Los mapas son una obviedad para visualizar preguntas específicas de la ubicación o para ayudar en la exploración geográfica.
- Los gráficos circulares son útiles para agregar detalles a otras visualizaciones, pero no son tan efectivos por sí solos. De hecho, cada año, en el Día Pi (14 de marzo), el científico de datos Kaiser Fung y su colega encuentran un gráfico circular y lo transforman en una versión más adecuada, como forma de protestar contra el uso excesivo flagrante.

Ejemplo



Pruebas de concepto

- **Extracción por medio de BigQuery**



```
1 SELECT case_hist.CaseNumber AS `case_number`, case_hist.CreatedDate AS `opened_date`, case_hist.Origin AS `case_origin`, case_hist.Status AS `status`, case_hist.OwnerId AS `case_owner`
```

Row	case_number	opened_date	case_origin	status	case_owner
1	07238726				
2	07257976				
3	07259113				
4	07270428				
5	07271417				
6	07175608				
7	07262900				
8	07270289				
9	07259414				
10	07284513				
11	04826119				

- **Conexión a BQ:** Se importan las librerías necesarias para realizar la conexión. Por otro lado, se establecen valores necesarios para la misma como:

Scope: Estos son los sitios los cuales son desplegados en caso de no contar el usuario ya autenticado. Una vez se ejecuta el código la primera vez y se autentifica el usuario, se genera el token, lo cual permite acceder a la información sin la necesidad de realizar este proceso nuevamente.

Project_Bigquery: Este es el nombre del proyecto dentro de GCP.

Queries: Estos son archivos ubicados en la misma locación del código. En estos se encuentra la consulta que se estará ejecutando dentro de Bigquery. Al tenerlos separados, podemos realizar modificaciones en el código de la consulta de forma sencilla.

```
import base64
import csv
import json
import re
import time
from datetime import datetime, timedelta
from email.mime.image import MIMEImage
from email.mime.multipart import MIMEMultipart
from email.mime.text import MIMEText
from math import ceil
from os import path, listdir, remove, rename

import matplotlib.pyplot as plt
import pytz
from dateutil.parser import parse
from google.auth.transport.requests import Request
from google.cloud.bigquery import Client
from google.oauth2.credentials import Credentials
from google.auth.oauthlib.flow import InstalledAppFlow
from googleapiclient.discovery import build
from googleapiclient.errors import HttpError
from numpy import busday_count
from openpyxl.utils.cell import get_column_letter

SCOPES = list = ['https://www.googleapis.com/auth/bigquery', 'https://www.googleapis.com/auth/spreadsheets',
                'https://www.googleapis.com/auth/gmail.send']
PROJECT_BIGQUERY = str = 'corpsvc-fint-edh1-prd-de9f'

FILE_CLIENT_SECRET = str = 'client_secret.json'
FILE_TOKEN = str = 'token.json'

QUERY_PRODUCTIVITY_EMAIL_REPORT = str = 'Productivity email report.sql'
QUERY_PENDING_CASES_ASSIGNED = str = 'Pending Cases - Assigned.sql'
QUERY_QUEUE_CASES_CREATED_BY_DATE = str = 'Queue Cases Created by Day.sql'
```

- **Autenticación:** Por medio de los valores establecidos anteriormente, se utiliza la librería de Google para realizar la autenticación por medio del Token y el Scope definido. Esto abre un sitio web de Google el cual ayuda a generar el client secret.

```
def authenticate_user(scopes: list[str], credential_directory: str, client_secret_file: str,
                    token_file: str) -> Credentials | None:
    """Authenticate as an end user.

    Authenticate with Google services as an end user and store the token/credentials in the specified directory.

    Args:
        scopes: The permissions that need to be requested.
        credential_directory: Directory for client secret and token.
        client_secret_file: File containing the client secret.
        token_file: File containing the authorized user's token.

    Returns:
        Credentials for accessing Google services and the authorized user or None is authentication fails.
    """
    creds: Credentials | None = None

    if path.exists(path.join(credential_directory, token_file)):
        creds = Credentials.from_authorized_user_file(path.join(credential_directory, token_file), scopes)
    # If there are no (valid) credentials available, let the user log in.
    if not creds or not creds.valid:
        if creds and creds.expired and creds.refresh_token:
            creds.refresh(Request())
        else:
            flow = InstalledAppFlow.from_client_secrets_file(path.join(credential_directory, client_secret_file),
                                                            scopes)
            creds = flow.run_local_server(port=0)
        # Save the credentials for the next run
        with open(path.join(credential_directory, token_file), 'w', encoding='utf-8') as token:
            token.write(creds.to_json())

    return creds
```

- **Ejecución de la consulta:** Una vez autenticado el usuario, se abren los archivos y se ejecuta la consulta que esta dentro de este. Los resultados de la consulta son extraídos.

```
def fetch_report_bq(bigquery_client: Client, file_directory: str, file_name: str, query_string: str) -> bool:
    """Update reports used as inputs for other processes.

    Runs the query and saves the results to a csv file with the name specified, selected fields must have an alias
    assigned ex: table.field as 'field' so the csv can be extracted.

    Args:
        bigquery_client: Authenticated BigQuery client for running queries.
        file_directory: Directory for csv file.
        file_name: Name of the csv file where the results are stored.
        query_string: Query for fetching report data.

    Returns:
        True if the report updates successfully.
        False if the file already exists.
    """
    if path.exists(path.join(file_directory, file_name)):
        return False

    # Get headers from query
    headers: list[str] = [h.replace("'", '') for h in HEADER_REGEX.findall(query_string)]

    query = bigquery_client.query(query_string)

    # Write query results to csv
    with open(path.join(file_directory, file_name), 'w', newline='', encoding='utf-8') as csvfile:
        writer = csv.writer(csvfile)
        writer.writerow(headers)

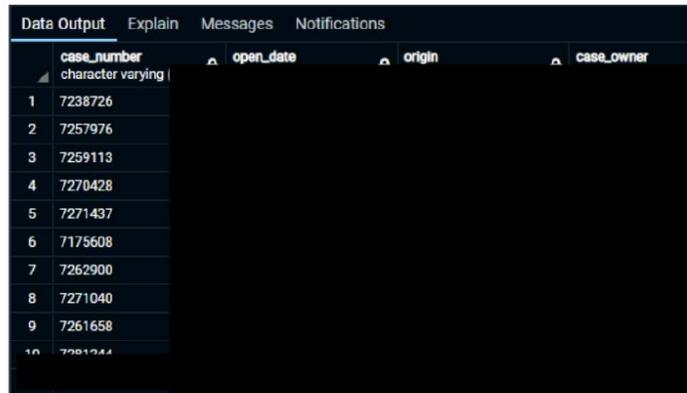
        for row in query.result():
            row_data = []

            for h in headers:
                row_data.append(row[h])

            writer.writerow(row_data)

    return True
```

Una vez ejecutados las consultas, son almacenadas en la base de datos PostgreSQL y la información queda lista para ser consultada.



	case_number	open_date	origin	case_owner
	character varying			
1	7238726			
2	7257976			
3	7259113			
4	7270428			
5	7271437			
6	7175608			
7	7262900			
8	7271040			
9	7261658			
10	7261244			

• **Extracción por medio de API (Qualtrics)**

StartDate	EndDate	Status	IPAddress	Progress	Duration	Finished	RecordedData	Response	Recipient	Recipient	Recipient	ExternalID	LocationJ	LocationL	Distributi	UserLang	Q
3/29/2018 20:53	3/29/2018 20:54	51	100	35	1	#####	R_9	LA	SA	SA	SA	SA	SA	SA	SA	SA	EN
3/29/2018 21:03	3/29/2018 21:04	51	100	56	1	#####	R_9	SA	SA	SA	SA	SA	SA	SA	SA	SA	EN
3/30/2018 12:23	3/30/2018 12:24	51	100	51	1	#####	R_5	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
3/30/2018 12:45	3/30/2018 12:46	42	100	48	1	#####	R_1	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/2/2018 18:33	4/2/2018 18:37	51	100	187	1	#####	R_9	DC	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:17	4/3/2018 13:19	42	100	77	1	#####	R_2	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:19	4/3/2018 13:19	51	100	41	1	#####	R_3	ER	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:21	4/3/2018 13:22	42	100	79	1	#####	R_3	LL	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:22	4/3/2018 13:23	42	100	106	1	#####	R_2	11	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:28	4/3/2018 13:29	42	100	47	1	#####	R_2	M	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:30	4/3/2018 13:32	42	100	108	1	#####	R_2	CP	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:31	4/3/2018 13:33	42	100	137	1	#####	R_1	MA	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:35	4/3/2018 13:37	51	100	129	1	#####	R_1	PP	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:38	4/3/2018 13:39	51	100	47	1	#####	R_1	IA	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:42	4/3/2018 13:42	51	100	49	1	#####	R_2	PH	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:39	4/3/2018 13:47	42	100	458	1	#####	R_1	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 13:57	4/3/2018 13:58	48	100	33	1	#####	R_2	MA	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 14:11	4/3/2018 14:13	41	100	104	1	#####	R_9	E	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 14:35	4/3/2018 14:35	47	100	38	1	#####	R_1	LO	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 14:51	4/3/2018 14:53	46	100	70	1	#####	R_1	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 14:49	4/3/2018 14:53	47	100	255	1	#####	R_0	18	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 15:24	4/3/2018 15:27	48	100	201	1	#####	R_2	H	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 15:30	4/3/2018 15:30	51	100	50	1	#####	R_2	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 15:41	4/3/2018 15:45	41	100	218	1	#####	R_1	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 15:50	4/3/2018 15:51	41	100	43	1	#####	R_1	NY	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 17:58	4/3/2018 17:58	42	100	53	1	#####	R_1	SM	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 18:17	4/3/2018 18:19	51	100	130	1	#####	R_1	JN	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/3/2018 18:31	4/3/2018 18:33	44	100	118	1	#####	R_1	SA	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/4/2018 14:41	4/4/2018 14:44	41	100	139	1	#####	R_1	CU	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/4/2018 14:56	4/4/2018 14:57	41	100	65	1	#####	R_2	LI	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/5/2018 18:16	4/5/2018 18:17	51	100	74	1	#####	R_3	CH	SA	SA	SA	SA	SA	SA	SA	SA	EN
4/5/2018 18:18	4/5/2018 18:19	42	100	32	1	#####	R_2	19	SA	SA	SA	SA	SA	SA	SA	SA	EN

- **Conexión al API:** Se importan las librerías necesarias para realizar la conexión. Por otro lado, se establecen valores necesarios para la misma como:

Properties_api: En este documento se encuentran los valores necesarios para enviar la solicitud por medio del API. Debido al funcionamiento del API, es necesario realizar 3 consultas distintas. En la primera se envía la solicitud con el id de la encuesta, la segunda consulta es para validar el progreso en la consulta debido a que el API inicia un proceso para generar un csv con los datos. Una vez el estado es completo, es necesario enviar una solicitud de descarga la cual envía el documento (imagen anterior) en una carpeta comprimida. Cabe destacar que para la ejecución del código se utilizó el documento en formato json.

```
import os
import sys
from time import sleep

import psycopp2.extras
import requests

FILE_PATH = os.path.realpath(__file__).replace(os.path.basename(__file__), '')
sys.path.append(FILE_PATH + '/../_Shared/python_modules')

import oct_utils

properties_api = oct_utils.get_properties(FILE_PATH + 'Properties/api_properties.json')

# Insert survey ID into urls
export_url_start = properties_api['export_url_start'].replace('[survey_id]', properties_api['survey_id'])
export_url_progress = properties_api['export_url_progress'].replace('[survey_id]', properties_api['survey_id'])
export_url_file = properties_api['export_url_file'].replace('[survey_id]', properties_api['survey_id'])

# Create headers and add decrypted API token
headers = {'X-APT-TOKEN': oct_utils.decrypt_password(properties_api['api_token'])}
```

- **Consultas:** Al realizar tener las configuraciones listas, se ejecuta el código por partes. Primeramente, se realiza se realiza una consulta a la base de datos para extraer la fecha del último dato agregado. Esto con el fin de no duplicar valores y cargar lo más nuevo. Una vez teniendo este valor, se genera el cuerpo de la consulta que se realizara por medio del API.

```
# Fetch start date from db
cursor.execute("SELECT CONCAT(to_json(MAX(recorded_date) + interval '1 second')#>> '{}', 'Z') FROM "
               "metricimetrackerschema.qualtrics_b2b_support_survey;")
start_date = cursor.fetchone()[0]
connection.commit()

if start_date == 'Z':
    start_date = '1970-01-01T01:00:00Z'

# Prepare request body
body = {'format': 'json', 'compress': False, 'startDate': start_date,
       "surveyMetadataIds": ["startDate", "endDate", "status", "progress", "duration", "finished", "recordedDate",
                             "_recordId"]}

# Send export start request
response = requests.post(export_url_start, headers=headers, json=body)
response = response.json()

# Insert progress ID into url
export_url_progress = export_url_progress.replace('[progress_id]', response['result']['progressId'])

# Check export progress until status is completed
while True:
    response = requests.get(export_url_progress, headers=headers)
    response = response.json()

    if response['result']['status'] == 'complete':
        break

    sleep(5)

# Insert file ID into url
export_url_file = export_url_file.replace('[file_id]', response['result']['fileId'])
```

```
# Get result from api
response = requests.get(export_url_file, headers=headers)
response = response.json()

# Read response data and prepare it for inserting into db
survey_data = []
```

- **Carga de datos:** Una vez se tienen los datos en formato json, se extraen los valores necesarios para la realización de los reportes. Estos son extraídos por etiquetas, y posteriormente se realiza la carga de datos por medio de la función de la librería psycopg que permite ejecutar consultas por lotes.

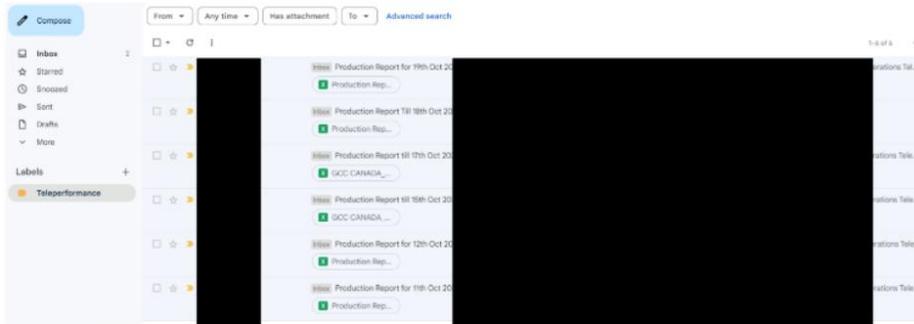
```
for survey in response['responses']:
    values = survey['values']
    labels = survey['labels']

    if values.get('QID59_TEXT'):
        comment = values.get('QID59_TEXT')
    else:
        comment = values.get('QID122463778_TEXT')

    survey_data.append((
        values.get('startDate'), # start_date timestamp
        values.get('endDate'), # end_date timestamp with
        values.get('progress'), # progress smallint
```

```
# Insert new surveys into db
query = 'INSERT INTO metricimetrackerschema.qualtrics_b2b_support_survey(start_date, end_date, progress, duration, ' \
        'finished, recorded_date, response_id, osat, hireagent, ease, howmany, sc0, q_total_duration, bu, bu_segment, ' \
        'call_center_segment, category, company_name, contact_reason, contact_type, country, csr_name, ' \
        'csr_supervisor, customer_segment, industry, manager, primary_region, ticket number, type, vertical, ' \
        'deployment_group, description, efx_id, csi_score, csi_percentage, emotions_sentiment, comment, osat_label, ' \
        'hireagent_label, ease_label, howmany_label, emotions) VALUES (%s, %s, %s, %s, %s, %s, %s, %s, %s, %s, ' \
        '%s, %s, ' \
        '%s, %s, %s);'
psycopg2.extras.execute_batch(cursor, query, survey_data)
connection.commit()
connection.close()
```

- **Extracción de Gmail**



- **Conexión a servicios Google:** Se importan las librerías necesarias para realizar la conexión. Por otro lado, se establecen valores necesarios para la misma como:
Scope: Estos son los sitios los cuales son desplegados en caso de no contar el usuario ya autenticado. Una vez se ejecuta el código la primera vez y se autentifica el usuario, se genera el token, lo cual permite acceder a la información sin la necesidad de realizar este proceso nuevamente.

```
import base64
import csv
import json
import re
import time
from datetime import datetime, timedelta
from email.mime.image import MIMEImage
from email.mime.multipart import MIMEMultipart
from email.mime.text import MIMEText
from math import ceil
from os import path, listdir, remove, rename

import matplotlib.pyplot as plt
import pytz
from dateutil.parser import parse
from google.auth.transport.requests import Request
from google.cloud.bigquery import Client
from google.oauth2.credentials import Credentials
from google_auth_oauthlib.flow import InstalledAppFlow
from googleapiclient.discovery import build
from googleapiclient.errors import HttpError
from numpy import busday_count
from openpyxl.utils.cell import get_column_letter

SCOPES: list = ['https://www.googleapis.com/auth/bigquery', 'https://www.googleapis.com/auth/spreadsheets',
                'https://www.googleapis.com/auth/gmail.send']
PROJECT_BIGQUERY: str = 'corpsvc-fint-edh1-prd-de9f'
SPREADSHEET_ID: str = ''
```

- **Autenticación:** Por medio de los valores establecidos anteriormente, se utiliza la librería de Google para realizar la autenticación por medio del Token y el Scope definido. Esto abre un sitio web de Google el cual ayuda a generar el client secret.

```
def authenticate_user(scopes: list[str], credential_directory: str, client_secret_file: str,
                    token_file: str) -> Credentials | None:
    """Authenticate as an end user.

    Authenticate with Google services as an end user and store the token/credentials in the specified directory.

    Args:
        scopes: The permissions that need to be requested.
        credential_directory: Directory for client secret and token.
        client_secret_file: File containing the client secret.
        token_file: File containing the authorized user's token.

    Returns:
        Credentials for accessing Google services and the authorized user or None is authentication fails.
    """
    creds: Credentials | None = None

    if path.exists(path.join(credential_directory, token_file)):
        creds = credentials.from_authorized_user_file(path.join(credential_directory, token_file), scopes)
    # If there are no (valid) credentials available, let the user log in.
    if not creds or not creds.valid:
        if creds and creds.expired and creds.refresh_token:
            creds.refresh(Request())
        else:
            flow = InstalledAppFlow.from_client_secrets_file(path.join(credential_directory, client_secret_file),
                                                            scopes)
            creds = flow.run_local_server(port=0)
        # Save the credentials for the next run
        with open(path.join(credential_directory, token_file), 'w', encoding='utf-8') as token:
            token.write(creds.to_json())

    return creds
```

- **Extracción de Gmail:** Una vez autenticado el usuario se procede a la extracción de los datos del correo electrónico. Es importante mencionar que estos están previamente configurados con una etiqueta para diferenciarlos de los demás. En el caso de este PoC se utilizaron los valores de “Teleperformance”. Una vez se obtienen los adjuntos se extraen y con ayuda de pandas se extrae la información para ser manipulada con ayuda de Python.

```
try:
    # create gmail api client
    service = build('gmail', 'v1', credentials=creds)

    # Load labels created on the user account
    labels = service.users().labels().list(userId='me').execute().get('labels', [])

    # Extract label ID from specific label
    label_id = labelIdExtraction(labels, 'teleperformance')

    # List all emails by label ID
    emails_list = service.users().threads().list(userId='me', labelIds=[label_id]).execute().get('threads', [])

    # Get data from last email received
    last_email_data = service.users().threads().get(userId='me', id=emails_list[0]['id'], format='full').execute()

    # Get the ID of the attachment received
    attachment_id = last_email_data['messages'][0]['payload'][1][1]['body']['attachmentId']

    # Get the attachment content
    attachment = service.users().messages().attachments().get(userId='me', messageId=emails_list[0]['id'], id=attachment_id).execute()

    # Using urlsafe_b64decode to retrieve data from the attachment
    file_data = base64.urlsafe_b64decode(attachment['data'].encode('UTF-8'))

    # Read Excel File
    raw_data = pd.read_excel(file_data)

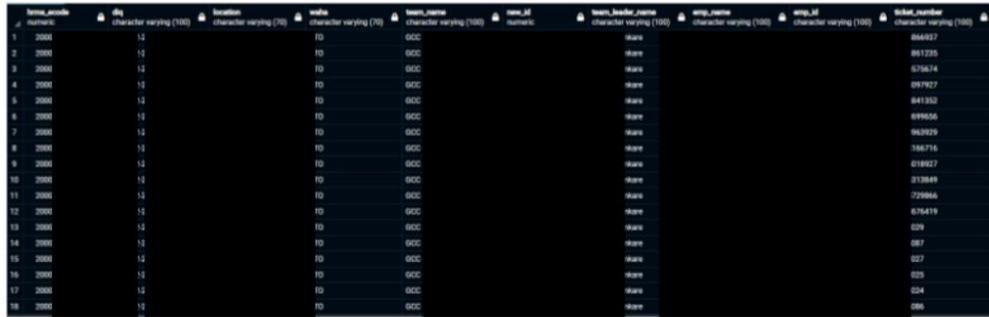
    # Transform NaT format to None: this because it breaks the flow when try to load data into the database
    raw_data['DIQ'] = (raw_data['DIQ'].astype(str).replace({'NaT': None}))
    raw_data['StartTime'] = (raw_data['StartTime'].astype(str).replace({'NaT': None}))
    raw_data['EndTime'] = (raw_data['EndTime'].astype(str).replace({'NaT': None}))
    raw_data['Date'] = (raw_data['Date'].astype(str).replace({'NaT': None}))
    raw_data['Tenure Date'] = (raw_data['Tenure Date'].astype(str).replace({'NaT': None}))

    # Convert Excel file to list using pandas
    raw_data = raw_data.values.tolist()

    # Slice list to get the 26 columns needed
    for row in raw_data:
        raw_data[raw_data.index(row)] = row[0:26]
    return raw_data

except HttpError as err:
    print(err)
```

- **Carga de datos:** Para cargar los datos, se utiliza la función de la librería psycog que permite ejecutar consultas por lotes. Como resultado se obtienen todos los datos que enviaron por correo electrónico en la base de datos.



fecha_solicitud	dia	location	status	email_name	email_id	email_header_name	email_name	email_id	ticket_number
20000	10	10	OSCC	ibero					844037
20000	10	10	OSCC	ibero					881225
20000	10	10	OSCC	ibero					676674
20000	10	10	OSCC	ibero					887927
20000	10	10	OSCC	ibero					841302
20000	10	10	OSCC	ibero					899056
20000	10	10	OSCC	ibero					863029
20000	10	10	OSCC	ibero					146276
20000	10	10	OSCC	ibero					818027
20000	10	10	OSCC	ibero					313049
20000	10	10	OSCC	ibero					729866
20000	10	10	OSCC	ibero					676419
20000	10	10	OSCC	ibero					029
20000	10	10	OSCC	ibero					087
20000	10	10	OSCC	ibero					027
20000	10	10	OSCC	ibero					023
20000	10	10	OSCC	ibero					024
20000	10	10	OSCC	ibero					080

9.30. Apéndice DD. Plan de implementación



Plan de implementación

Implementación del nuevo modelo.



Introducción

El presente documento presenta el plan de implementación del proceso automatizado para la elaboración de reportes a los departamentos “*Business care*” y “*Boarding*”. En esta propuesta se pretende realizar una serie de mejoras al proceso actual. Algunas de estas mejoras serían:

- Extracción de datos automatizada

La extracción de datos se realiza de forma automatizada diariamente. Esto quiere decir que, se obtienen los datos más nuevos por medio de conexiones directas a las distintas fuentes de datos. Un aspecto importante para mencionar es que la extracción de datos puede llegar a ser calendarizada con ayuda de alguna otra herramienta de como Bamboo o Airflow.

- Carga de datos automatizada

La carga de datos a las fuentes de los reportes se realiza de forma automática de forma diaria. Esto removiendo los pasos manuales en los cuales se debían almacenar los documentos en folder específicos para que fueran cargados por otra persona. Gracias a este proceso

- Conexión en vivo del reporte a la base de datos

Aprovechando las actualizaciones en el proceso de extracción y carga, se obtiene el beneficio de realizar conexiones en vivo desde los reportes a la base de datos. Con esto se ahorra el paso de tener que abrir cada uno de los reportes y actualizar las fuentes de datos. Ahora se tiene una conexión directa por lo cual, cualquier dato nuevo en la base de datos se ve reflejado en los reportes.

El documento busca ser guía de las gestiones necesarias para el proceso de implementación, por esta razón, dentro de este se establece el alcance, la gestión del tiempo, gestión de recursos, gestión de riesgos, gestión de interesados y la gestión de la comunicación. Esto con el fin de concluir el proyecto de la forma más efectiva y eficiente.

Alcance

El presente proyecto busca realizar una implementación de un nuevo proceso para la creación de reportes. Este proceso consiste en aplicar un ETL dentro del proceso actual, eliminando las tareas manuales realizadas por los encargados de realizar dichos reportes. El proceso de implementación se divide en las siguientes fases:

- Fase 0: Recopilación: En esta fase se realizará una recopilación de nuevos requerimientos y validación de los existentes. también, se realiza una recopilación de KPIs, cálculos y funciones necesarias que deben ser aplicadas en el código.
- Fase 1: Extracción: En esta fase se realizará el análisis de las distintas fuentes de datos, para validar los datos que deben ser extraídos de cada una de estas. Una vez identificados, se procede a la realización del código Python utilizando distintos medios y herramientas para extraer la información cruda.
- Fase 2: Transformación: En esta fase se procede a realizar las funciones necesarias para realizar todos los cálculos necesarios. Estos cálculos pueden ser: transformar tipos de datos como fechas, valores numéricos, etc. También, se realiza los campos calculados de los reportes, tales como vejez por caso, tiquete, tipo y causa.
- Fase 3: Carga: En esta fase se decide como cargar los datos de las distintas fuentes, estos pueden ser cargados por lotes si la cantidad es grande o bien hacer uso de funciones iterativas las cuales hagan inserciones individuales.
- Fase 4: Automatización: En esta fase se realiza el proceso de automatización, calendarizando la ejecución de los códigos para que estos sean ejecutados por el sistema o el servidor de forma automático a cierta hora del día.

Gestión del tiempo

El tiempo de ejecución del proyecto se encuentra establecido según experiencias previas dentro del equipo. Este cronograma se encuentra establecido por tareas y el tiempo estimado de ejecución de cada una de estas.

NÚMERO EDIT	TÍTULO DE LA TAREA	RESPONSABLE DE LA TAREA	DURACIÓN (horas)	% COMPLETADO DE LA TAREA	SEMANA 1		SEMANA 2		SEMANA 3		SEMANA 4		SEMANA 5		SEMANA 6	
					L	M	X	J	V	L	M	X	J	V	L	M
1 Fase #0: Recopilación																
1.1	Validación de requerimientos	Desarrollador de BI	2	0 %												
1.2	Recopilación de KPIs	Desarrollador de BI	6	0 %												
1.3	Recopilación de campos calculados	Desarrollador de BI	6	0 %												
1.4	Validación de KPIs y campos	Desarrollador de BI	2	0 %												
1.5	Solicitar accesos	Gerente del departamento	2	0 %												
2 Fase #1: Extracción																
2.1	Analizar fuentes de datos	Desarrollador de BI	16	0 %												
2.2	Realizar conexión con las fuentes	Desarrollador de BI	24	0 %												
2.3	Revisar datos disponibles	Desarrollador de BI	8	0 %												
3 Fase #2: Transformación																
3.1	Calcular vejez	Desarrollador de BI	8	0 %												
3.2	Estandarizar tipos de fechas	Desarrollador de BI	4	0 %												
4 Fase #3: Carga																
4.1	Diagramación de tablas	Desarrollador de BI	30	0 %												
4.2	Creación de tablas	Desarrollador de BI Encargado de soporte	16	0 %												
4.3	Cargar datos	Desarrollador de BI	10	0 %												
5 Fase #4: Automatización																
5.1	Automatizar la extracción de cada una de las fuentes de datos	Desarrollador de BI Encargado de soporte	24	0 %												
5.2	Calendarizar las actividades	Desarrollador de BI Encargado de soporte	4	0 %												
6 Fase #5: Pruebas																
6.1	Pruebas del nuevo proceso	Desarrollador de BI	24	0 %												

Gestión de los recursos

Para realizar este proyecto será necesario contar con ciertos recursos dedicados al mismo. Esto con fin de terminarlo en el tiempo establecido en el cronograma. Cabe destacar que los colaboradores ya se encuentran dentro de la organización por lo cual el tiempo de entrenamiento se excluye.

Recurso	Cantidad	Descripción
Desarrollador de BI	1	Es la persona encargada de realizar la implementación.
Desarrollador de BI	1	Esta persona es la que tiene conocimiento en el rol.
Softwares	6	Los softwares necesarios para realizar la implementación son: <ul style="list-style-type: none">• Python• Excel• BigQuery• Postman• GCP• PostgreSQL
Hardware	2	Equipo físico necesario para la implementación (Laptop)

Gestión de riesgos

Los principales riesgos del proyecto a implementar se encuentran identificados.

ID	Descripción
1	Problemas con las licencias
2	Problemas de hardware
3	Aumento de costos operacionales
4	Problemas con la resistencia al cambio
5	Retrasos en instalaciones de software
6	Mala configuración de softwares
7	Poca comunicación
8	Cambios en políticas de la organización
9	Problemas con los accesos

Por otro lado, es importante complementar los riesgos previamente identificados, para esto, se detallan los criterios de probabilidad e impacto que serán utilizados para clasificar cada uno.

- **Probabilidad:** Se refiere al grado que tiene el riesgo para materializarse durante la ejecución de las iniciativas.

Nivel	Probabilidad	Descripción
1	Baja	Ocurre en solo ciertas ocasiones
2	Media	No tiene ninguna condición especial, pero es posible que ocurra
3	Alta	Tiene una alta ocurrencia

- **Impacto:** Indica el grado de influencia positiva o negativa que posee un riesgo al materializarse.

Nivel	Impacto	Descripción
1	Bajo	El riesgo puede tener un impacto bajo o nulo
2	Medio	Si se presenta, debe ser atendido de rápidamente, no tiene un alto nivel de urgencia
3	Alto	Presenta un impacto muy grave dentro de las operaciones de la organización

Con base a los criterios establecidos anteriormente, se elabora una matriz de calor para los riesgos del proyecto, los cuales se encuentran evaluados en la tabla a continuación.

	Bajo	Medio	Alto
Alto	3	6	9
Medio	2	4	6
Bajo	1	2	3

Una vez definidos estos criterios se realizó una evaluación de los riesgos identificados anteriormente.

ID	Nivel de probabilidad	Nivel de impacto	Resultado
1	Baja	Medio	
2	Baja	Alto	
3	Baja	Bajo	
4	Baja	Alto	
5	Medio	Alto	
6	Baja	Alto	
7	Baja	Medio	
8	Baja	Alto	
9	Media	Alta	

Gestión de interesados

En esta sección se encuentran los principales interesados en la ejecución del proyecto. Cada uno de estos cumple un rol importante dentro del proceso y son importante para la toma de decisiones dentro del mismo.

Interesados	Expectativa	Importancia
Director del departamento del Servicio de Datos	Colaboración con temas de herramientas, licencias y presupuesto.	Baja
Gerente del departamento del Servicio de Datos	Colaboración con accesos, información, conexiones, y conocimiento.	Alta
Desarrollador de BI del departamento de Servicio de Datos	Brindar la información completa del proceso y cumplir un rol de control de calidad.	Alta
Encargado de soporte del equipo	Brindar soporte a las herramientas utilizadas.	Media
Departamento de TI y seguridad	Brindar llaves para el uso de conexiones como BigQuery o el API de Google.	Media
Gerente del departamento de "Business Care"	Brindar conocimiento e información sobre los reportes generados y los indicadores esperados.	Alta

Gestión de comunicación

En esta sección se establecen los interesados a los cuales se les distribuirá información del proyecto. Esta sección es de gran importancia debido a que permitirá mantener al equipo completo informado de los pasos realizados y se pueda tomar decisiones en momentos claves.

Interesados	Medio de comunicación	Tipo de información	Periodicidad
Director del departamento del Servicio de Datos	Video llamada Correo electrónico	Reporte de avance	Semanal
Gerente del departamento del Servicio de Datos	Video llamada Correo electrónico	Reporte de avance Reporte de bloqueos	Martes y jueves
Departamento de TI y seguridad	Correo electrónico	Reporte de uso seguro de información	Al extraer los datos.
Gerente del departamento de "Business Care"	Correo electrónico	Reporte de avance Consultas	Semanal

9.31. Apéndice EE. Minutas

9.31.1. Minuta 1. Presentación formal estudiantes y tutor

Minuta de Reunión			
Información general			
Núm. Reunión	01	Fecha	29/07/2022
Lugar/Medio	Teams	Hora de inicio y finalización	4:30 pm – 5:00 pm
Objetivo	Presentación formal estudiantes y tutor		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Araya (Estudiante) Ricardo Oliver (Estudiante) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Presentación	Se realizó una presentación entre las partes.	
2	Plan de trabajo	Se definió el plan de trabajo para el desarrollo del proyecto, revisión de tareas y orden de realización	Cada estudiante pasar el anteproyecto a la estructura del proyecto y utilizar un editor en línea para compartirlo con el tutor. Este revisara semanalmente los avances y realizara comentarios
3	Medios de comunicación		Se abrieron canales como un grupo de chat por WhatsApp y si es necesario realizar reuniones por Teams o Meets.
Próxima reunión			
Temas por tratar		Fecha	Convocados
Establecer roles, alcance de proyecto y expectativas de la Organización.		5/08/2022	Agustín Francesa Luis Vindas Carlos Morales Ricardo Oliver

9.31.2. Minuta 2. Establecer roles, alcance y expectativas del proyecto.

Minuta de Reunión			
Información general			
Núm. Reunión	02	Fecha	05/08/2022

Lugar/Medio	Google Meets	Hora de inicio y finalización	5:00 pm – 6:00 pm
Objetivo	Establecer roles, alcance y expectativas del proyecto.		
Participantes	Presentes	Agustín Francesa (Tutor) Luis Vindas (Encargado de empresa) Ricardo Oliver (Estudiante) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Asignación de roles	Tanto el tutor como el encargado de la organización dieron a conocer sus roles y tareas a realizar a fines de la ejecución del proyecto	El encargado de empresa se compromete a colaborar en la ejecución del proyecto, facilitando información y contactos.
2	Establecer expectativas	Carlos explico los objetivo y alcance del proyecto con el fin de delimitar el proyecto. Luis menciono expectativas esperadas con la ejecución del proyecto.	
Próxima reunión			
Temas por tratar		Fecha	Convocados
Evaluación Tutor-Organización		2/9/2022	Agustín Francesa Luis Vindas Carlos Morales

9.31.3. Minuta 3. Realizar ajustes del proyecto en términos del alcance y objetivos del proyecto.

Minuta de Reunión			
Información general			
Núm. Reunión	03	Fecha	05/08/2022

Lugar/Medio	Google Meets	Hora de inicio y finalización	6:00 pm – 7:00 pm
Objetivo	Realizar ajustes del proyecto en términos del alcance y objetivos del proyecto.		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Reestablecer alcance	Al obtener una descripción del problema y del proyecto se reestableció el alcance del proyecto.	Se cambio el alcance del proyecto para que se realice el plan de implementación y no la implementación.
2	Realizar ajustes a los objetivos	Debido al cambio en el alcance del proyecto se realizaron ajustes en los objetivos.	Cambiar los objetivos y agregar nuevos.
Próxima reunión			
Temas por tratar		Fecha	Convocados
Revisión del capítulo 1 del proyecto.		10/08/2022	Agustín Francesa Carlos Morales

9.31.4. Minuta 4. Revisión del capítulo 1

Minuta de Reunión			
Información general			
Núm. Reunión	04	Fecha	10/08/2022
Lugar/Medio	Whatsapp	Hora de inicio y finalización	6:00 pm – 6:30 pm
Objetivo	Revisión del capítulo 1		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisión del alcance del proyecto.	Se realizo una revisión en el alcance del proyecto	Detallar más el que se va hacer y que no se va a realizar en el proyecto.
Próxima reunión			
Temas por tratar		Fecha	Convocados
Revisión de cambios		17/8/2022	Agustín Francesa Carlos Morales

9.31.5. Minuta 5. Revisión de cambios realizados en el capítulo 1 e idealizar el proceso metodológico que se utilizara.

Minuta de Reunión			
Información general			
Núm. Reunión	05	Fecha	17/08/2022
Lugar/Medio	Meets	Hora de inicio y finalización	3:00 pm – 3:30 pm
Objetivo	Revisión de cambios realizados en el capítulo 1 e idealizar el proceso metodológico que se utilizara.		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisar cambios realizados del capítulo 1.		
2	Definir estrategia para realizar el capítulo 3.	Se vio un posible camino a seguir para la ejecución del capítulo 3.	Investigar sobre metodologías de rediseño de procesos y otras metodologías ya probadas para realizar el procedimiento metodológico.
Próxima reunión			
Temas por tratar		Fecha	Convocados
			Agustín Francesa Carlos Morales

9.31.6. Minuta 6. Aclarar dudas con respecto al marco metodológico y sus partes.

Minuta de Reunión			
Información general			
Núm. Reunión	06	Fecha	26/08/2022
Lugar/Medio	Whatsapp	Hora de inicio y finalización	6:00 pm – 7:00 pm
Objetivo	Aclarar dudas con respecto al marco metodológico y sus partes.		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Estructura del capítulo 3.	Se aclaro ciertas dudas con respecto a las partes que lleva el capítulo 3	Eliminar la sección de “Población y selección de muestra”
Próxima reunión			
Temas por tratar		Fecha	Convocados
			Agustín Francesa Carlos Morales

9.31.7. Minuta 7. Revisión del capítulo 3.

Minuta de Reunión			
Información general			
Núm. Reunión	07	Fecha	05/09/2022
Lugar/Medio	Meets	Hora de inicio y finalización	1:00 pm – 2:00 pm
Objetivo	Revisión del capítulo 3.		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante) Ricardo Oliver (Estudiante) Carlos Araya (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Aclarar dudas con aspecto del capítulo 3 de cada uno de los integrantes		Buscar una mejor metodología para realizar el análisis de costo beneficio
Próxima reunión			
Temas por tratar		Fecha	Convocados
			Agustín Francesa Carlos Morales

9.31.8. Minuta 8. 2da revisión del proyecto de graduación

Minuta de Reunión			
Información general			
Núm. Reunión	08	Fecha	28/09/2022
Lugar/Medio	Whatsapp	Hora de inicio y finalización	3:30 pm – 4:30 pm
Objetivo	2da revisión del proyecto de graduación		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante) Carlos Araya (Estudiante) Ricardo Oliver (Estudiante) Luis Vindas (Contraparte)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisar avance en el proyecto	Cada estudiante brindo detalles de cómo iban con cada proyecto.	
2	Análisis de herramientas	Se hablo con Luis la posibilidad de utilizar otras herramientas, pero se concluyó que es ideal enfocarse solo en Python.	
Próxima reunión			
Temas por tratar		Fecha	Convocados
			Agustín Francesa Carlos Morales

9.31.9. Minuta 9. Recolección de requerimientos para estándar

Minuta de Reunión			
Información general			
Núm. Reunión	09	Fecha	30/09/2022
Lugar/Medio	Meets	Hora de inicio y finalización	10:00 am – 11::30 am
Objetivo	Recolección de requerimientos para estándar		
Participantes	Presentes	Carlos Morales (Estudiante) Ricardo Oliver (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisar dashboard actuales	Se identifican las necesidades actuales.	
2	Definir cantidad de visualizaciones	Se revisan la cantidad de visualizaciones por dashboard.	Se define tener de 3 a 5 visualizaciones por tablero
3	Revisión de libro de marca de la organización	Revisar detalles establecidos por la organización en el libro de marca.	
Próxima reunión			
Temas por tratar		Fecha	Convocados
Realizar propuesta de estandarización		10/05/2022	Carlos Morales Ricardo Oliver

9.31.10. Minuta 10. Validación del modelo “as-is”

Minuta de Reunión			
Información general			
Núm. Reunión	10	Fecha	30/09/2022
Lugar/Medio	Meets	Hora de inicio y finalización	3:30 pm – 4:30 pm
Objetivo	Validación del modelo “as-is”		
Participantes	Presentes	Carlos Morales (Estudiante) Diego Alfaro (Encargado de realizar reportes) Luis Vindas (Contraparte)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisar la información recolectada en la fase 1 y el modelo “as-is”	Se mostro los resultados de la fase 1 para brindar contexto de los resultados y así validar el modelo “as-is”.	El modelo fue validado tanto por Diego, que es la persona que posee más conocimiento en el proceso y por Luis.
2	Revisión de tiempos dentro del proceso	Se revisaron los tiempos tomados dentro del proceso de observación.	Utilizar tiempos de observación para el tiempo de ejecución y el de espera, revisión y retrabajo, colocar los calculados por Diego.
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.11. Minuta 11. Realizar propuesta de estandarización

Minuta de Reunión			
Información general			
Núm. Reunión	11	Fecha	05/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	10:00 am – 11:30 am
Objetivo	Realizar propuesta de estandarización		
Participantes	Presentes	Carlos Morales (Estudiante) Ricardo Oliver (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Estructura de dashboard		Se define las formas en las que se debe mostrar cada visualización
2	Colores, fuentes y distribución del dashboard		Ajustarse a las políticas de la organización.
3	Uso recomendado de gráficos		Se brinda una guía para selección de gráficos.
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.12. Minuta 12. Validar propuesta de estandarización

Minuta de Reunión			
Información general			
Núm. Reunión	12	Fecha	07/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	2:00 pm – 2:30 pm
Objetivo	Validar propuesta de estandarización		
Participantes	Presentes	Carlos Morales (Estudiante) Ricardo Oliver (Estudiante) Luis Vindas (Contra parte) Diego Alfaro (Encargado de realizar reportes) Hazel Fernández (Encargado de realizar reportes)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Estructura de dashboard		Se define las formas en las que se debe mostrar cada visualización
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.13. Minuta 13. Validación de pruebas de concepto y consultas sobre proyectos similares

Minuta de Reunión			
Información general			
Núm. Reunión	13	Fecha	10/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	9:30 am – 10:30 am
Objetivo	Validación de pruebas de concepto y consultas sobre proyectos similares		
Participantes	Presentes	Carlos Morales (Estudiante) Diego Alfaro (Encargado de realizar reportes) Luis Vindas (Contraparte)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Revisión de pruebas de concepto	Se presentaron las pruebas de concepto al encargado de realizar los reportes. Con el fin de validar el nuevo proceso.	
2	Tiempos de ejecución en proyectos similares	Se consulto con Luis y Diego los periodos promedios de realizar este tipo de implementación conociendo el entorno de la organización	
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.14. Minuta 14. Atender dudas capítulos 4 y 5

Minuta de Reunión			
Información general			
Núm. Reunión	14	Fecha	10/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	6:15 pm – 7:00 pm
Objetivo	Atender dudas capítulos 4 y 5		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante) Carlos Araya (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Atender dudas capítulo 4 y 5	Se reviso el capítulo 4 análisis de resultados y se atendieron dudas del capítulo 5 Propuesta de solución	
2	Revisar figuras		Las figuras deben estar en horizontal para una mejor visión de estas
3	Validación de tiempo para terminar el proyecto	Se analizo partes faltantes del proyecto de cada uno de los estudiantes.	Se definió que cada estudiante debe de terminar su proyecto para enviarlo al filólogo con una semana de anticipación a la entrega final.
Próxima reunión			
Temas por tratar		Fecha	Convocados
			Agustín Francesa Carlos Morales

9.31.15. Minuta 15. Revisión de propuesta de solución

Minuta de Reunión			
Información general			
Núm. Reunión	15	Fecha	18/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	6:45pm – 7:30 pm
Objetivo	Revisión de propuesta de solución		
Participantes	Presentes	Agustín Francesa (Tutor) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Dudas capítulo 5	Se atendieron consultas correspondientes a los entregables del proyecto y el análisis costo beneficio.	Realizar suma de costos y validación de beneficios, para enumerarlos como resultados del análisis
2	Revisión de todo el proyecto		El estudiante debe revisar concordancia entre todos los capítulos del proyecto.
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.16. Minuta 16. Validación de beneficios de la propuesta

Minuta de Reunión			
Información general			
Núm. Reunión	16	Fecha	20/10/2022
Lugar/Medio	Meets	Hora de inicio y finalización	9:00 am – 9:30 am
Objetivo	Validación de beneficios de la propuesta		
Participantes	Presentes	Luis Vindas (Contraparte) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Presentación de propuesta	Se presenta la propuesta desarrollada a la contraparte de la organización.	
2	Beneficios esperados	Definir cuales son los comentarios esperados con la propuesta por parte de la organización.	
Próxima reunión			
Temas por tratar		Fecha	Convocados

9.31.17. Minuta 17. 3era revisión del proyecto de graduación

Minuta de Reunión			
Información general			
Núm. Reunión	17	Fecha	03/11/2022
Lugar/Medio	Meets	Hora de inicio y finalización	2:00 pm – 2:30 pm
Objetivo	3era revisión del proyecto de graduación		
Participantes	Presentes	Luis Vindas (Contraparte) Agustín Francesa Alfaro (Tutor) Ricardo Oliver (Estudiante) Carlos Araya (Estudiante) Carlos Morales (Estudiante)	
	Ausentes		
Temas tratados			
ID	Asunto	Comentario	Acuerdo
1	Presentación de la propuesta y conclusiones	Se presenta la propuesta desarrollada y las conclusiones a la contraparte de la organización.	
2	Cierre del proyecto		
Próxima reunión			
Temas por tratar		Fecha	Convocados

Capítulo 10. Anexos

10.1. Carta filológica

San José, Costa Rica, 3 de noviembre de 2022

Ingeniero Luis Javier Chavarría Sánchez
Director
Área Académica de Administración de Tecnologías de Información
Administración de Tecnología de información
Instituto Tecnológico de Costa Rica

Estimado señor:

Por este medio hago constar que he revisado el Trabajo Final de Graduación, titulado *Propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax*, elaborado como requisito para obtener el grado de Licenciatura en Administración de Tecnología de Información por el estudiante Carlos Humberto Morales Porras, cédula: 1-1678-0014.

Al respecto, indico que he revisado y corregido errores gramaticales, de puntuación y ortografía, construcción de párrafos, vicios del lenguaje, aplicación del formato de citación APA séptima edición, así como otros aspectos relacionados con el campo filológico, que se manifestaron en el documento escrito.

Desde ese punto de vista, considero que, con las correcciones realizadas en el documento, está listo para ser presentado como Trabajo Final de Graduación.

Atentamente,



Licda. Marjorie Martínez Castro
Cédula 5-0339-0587
Carné de colegiatura: 94974

10.2. Firma minutas tutor

Firma de Minutas por parte del profesor tutor

Por medio de este documento, se agrupan y firman las minutas de reuniones realizadas a lo largo del desarrollo del Trabajo Final de Graduación titulado “propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax”, y elaborado por el estudiante Carlos H. Morales Porras, carné 2017238921.

El profesor tutor, Agustín Francesa Alfaro, Valida su participación en las siguientes minutas:

- Minuta 1: Presentación formal estudiantes y tutor
- Minuta 2: Establecer roles, alcance y expectativas del proyecto.
- Minuta 3: Realizar ajustes del proyecto en términos del alcance y objetivos del proyecto.
- Minuta 4: Revisión del capítulo 1
- Minuta 5: Revisión de cambios realizados en el capítulo 1 e idealizar el proceso metodológico que se utilizara.
- Minuta 6: Aclarar dudas con respecto al marco metodológico y sus partes.
- Minuta 7: Revisión del capítulo 3.
- Minuta 8: 2da revisión del proyecto de graduación
- Minuta 14: Atender dudas capítulos 4 y 5
- Minuta 15: Revisión de propuesta de solución
- Minuta 17: 3era revisión del proyecto de graduación


JOSE AGUSTIN
FRANCESA
ALFARO (FIRMA)
2022.11.04
09:04:49-06'00'

Agustín Francesa Alfaro


Carlos Morales Porras

10.3. Firma minutas empresa

Firma de Minutas por parte de la organización

Por medio de este documento, se agrupan y firman las minutas de reuniones realizadas a lo largo del desarrollo del Trabajo Final de Graduación titulado “propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax”, y elaborado por el estudiante Carlos H. Morales Porras, carné 2017238921.

El encargado de empresa, Luis Eduardo Vindas Espinoza, valida su participación en las siguientes minutas e instrumentos aplicados:

- Minuta 2: Establecer roles, alcance y expectativas del proyecto.
- Minuta 8: 2da revisión del proyecto de graduación
- Minuta 10: Validación del modelo “as-is”
- Minuta 12: Validar propuesta de estandarización
- Minuta 13: Validación de pruebas de concepto y consultas sobre proyectos similares
- Minuta 16: Validación de beneficios de la propuesta
- Minuta 17: 3era revisión del proyecto de graduación
- Entrevista #1: Conocer sobre Equifax
- Entrevistas #2: Conocer sobre el proceso de generación de reportes para los departamentos de Business Care y Boarding



Luis Eduardo Vindas Espinoza



Carlos Morales Porras

10.4. Firma minutas con colaboradores

10.4.1. Desarrollador de BI

Firma de Minutas e instrumentos aplicados a la organización

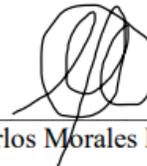
Por medio de este documento, se agrupan y firman las minutas de reuniones realizadas a lo largo del desarrollo del Trabajo Final de Graduación titulado “propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax”, y elaborado por el estudiante Carlos H. Morales Porras, carné 2017238921.

El desarrollador de inteligencia de negocios, Diego Alfaro, valida su participación en las siguientes minutas e instrumentos aplicados:

- Minuta 10: Validación del modelo “as-is”
- Minuta 12: Validar propuesta de estandarización
- Minuta 13: Validación de pruebas de concepto y consultas sobre proyectos similares
- Observación del proceso de creación de reportes
- Entrevistas #2: Conocer sobre el proceso de generación de reportes para los departamentos de Business Care y Boarding

DIEGO
ALFARO
QUIROS
(FIRMA)
Digitally signed
by DIEGO
ALFARO QUIROS
(FIRMA)
Date: 2022.10.31
22:44:10 -06'00'

Diego Alfaro Quirós



Carlos Morales Porras

10.4.2. Encargado de soporte

Firma de Minutas e instrumentos aplicados a la organización	
<p>Por medio de este documento, se agrupan y firman las minutas de reuniones realizadas a lo largo del desarrollo del Trabajo Final de Graduación titulado “propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax”, y elaborado por el estudiante Carlos H. Morales Porras, carné 2017238921.</p>	
<p>El encargado de brindar soporte al equipo de servicios de datos, Mateo Marín, valida su participación en las siguientes minutas e instrumentos aplicados:</p>	
<ul style="list-style-type: none">• Entrevistas #4: Conocer los requerimientos tecnológicos del departamento	
 Mateo Marín Lopera	 Carlos Morales Porras

10.4.3. Desarrollo de estándar

Firma de Minutas e instrumentos aplicados a la organización	
<p>Por medio de este documento, se agrupan y firman las minutas de reuniones realizadas a lo largo del desarrollo del Trabajo Final de Graduación titulado “propuesta de implementación de una solución para la automatización del proceso de generación de reportes de los departamentos de Business care y Boarding en Equifax”, y elaborado por el estudiante Carlos H. Morales Porras, carné 2017238921.</p>	
<p>El desarrollador de inteligencia de negocios, Ricardo Oliver Murillo, valida su participación en las siguientes minutas e instrumentos aplicados:</p>	
<ul style="list-style-type: none">• Minuta #9: Recolección de requerimientos para estándar• Minuta #11: Realizar propuesta de estandarización• Minuta #12: Validar propuesta de estandarización	
 Ricardo Oliver Murillo	 Carlos Morales Porras