



Licenciatura en Administración de Tecnología de Información

Desarrollo de solución de inteligencia de negocios para interpretación,
análisis y evaluación de indicadores clave de desempeño relacionados
con el procesamiento de criptomonedas. Caso: Progressio Digital

Informe de Trabajo Final de Graduación para optar por el grado de
Licenciatura en Administración de Tecnología de Información

Elaborado por: José Andrés Barboza González

Cartago

I semestre, 2020



Esta obra está sujeta a la licencia

Reconocimiento-NoComercial-

SinObraDerivada 4.0

Internacional de Creative Commons.

Para ver una copia de esta licencia,

visite <http://creativecommons.org/licenses/by-nc-nd/4.0>

INSTITUTO TECNOLÓGICO DE COSTA RICA

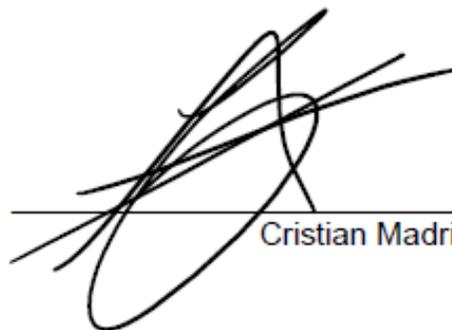
ÁREA ACADÉMICA DE ADMINISTRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN
GRADO ACADÉMICO LICENCIATURA

Los miembros del Tribunal Examinador del Área Académica de Administración de Tecnologías de Información recomendamos que el presente Informe Final del Proyecto de Graduación de la estudiante José Andrés Barboza González sea aceptado como requisito parcial para obtener el grado académico de Licenciatura en Administración de Tecnología de Información.

JACQUELINE
TATIANA SOLIS
CESPEDES (FIRMA)

Firmado digitalmente por
JACQUELINE TATIANA
SOLIS CESPEDES (FIRMA)
Fecha: 2020.08.03 21:19:10
-06'00'

Jacqueline Solís Céspedes
Profesora Asesora



Cristian Madrigal Arias
Lector

Digitally signed by NESTOR ALEJANDRO MORALES RODRIGUEZ (AUTENTICACION)
DN: SERIALNUMBER=CPF-03-0450-0858, SN=MORALES RODRIGUEZ, G=NESTOR
ALEJANDRO, C=CR, O=PERSONA FISICA, OU=CIUDADANO, CN=NESTOR
ALEJANDRO MORALES RODRIGUEZ (AUTENTICACION)
Reason: I am approving this document with my legally binding signature
Location:
Date: 2020-08-03 21:07:35

PDF

Néstor Morales Rodríguez
Lector

YARIMA TATIANA
SANDOVAL
SANCHEZ (FIRMA)

Firmado digitalmente por
YARIMA TATIANA SANDOVAL
SANCHEZ (FIRMA)
Fecha: 2020.08.04 15:06:25
-06'00'

Yarima Sandoval Sánchez
Coordinadora del Proyecto de Graduación de la Licenciatura en Administración de
Tecnología de Información

Resumen

El presente trabajo de graduación especifica el desarrollo de una solución de inteligencia de negocios dentro de la empresa Progressio Digital, específicamente para los sistemas de procesamiento de criptomonedas que esta ofrece.

Actualmente, la organización no cuenta con ningún proyecto similar o relacionado con soluciones de análisis de información; por otro lado, tampoco existen indicadores clave de desempeño adecuadamente identificados ni monitoreados, que permitan llevar control cuantiado sobre el funcionamiento de las soluciones.

A través del presente proyecto, se determinan estos identificadores clave de desempeño, para monitorearlos, a través de una aplicación de visualización de información, que consulta un almacén de datos y, a su vez, aloja un modelo relacional. Para llevar a cabo la elaboración de esta solución, se combinó un conjunto de varias herramientas y tecnologías, basándose en las buenas prácticas expuestas en la metodología de desarrollo de aplicación de inteligencia de negocios llamada Kimball.

El proceso de elaboración de esta solución inicia por el análisis de la situación actual de los sistemas de información de la organización, tomando en cuenta sus bases de datos y sus procesos. A partir de ahí, se procede a realizar la identificación y priorización de indicadores clave de desempeño con respecto a estos sistemas, además, se toma como base instrumentos y herramientas de recolección de información. Luego se procede a identificar los requerimientos y a desarrollar los diseños del almacén de datos y de los procesos ETL. Por último, se crean las visualizaciones respectivas a los indicadores de mayor prioridad.

Por medio del uso de esta solución, se pretende mejorar la visibilidad de la gerencia de la organización, con respecto al desempeño de los sistemas de información relacionados con el procesamiento de criptomonedas, de manera que tengan la capacidad de tomar decisiones sobre evidencia real y actualizada.

Palabras clave: Inteligencia de Negocios, Procesamiento de Criptomonedas, Almacén de datos

Abstract

This graduation paper specifies the development of a business intelligence solution within the company Progressio Digital, specifically for the cryptocurrency processing systems that they offer.

Currently, the organization does not have any similar project or related to information analysis solutions; on the other hand, there are no identified or monitored key performance indicators to maintain quantified control over the operation of the solutions.

Through this project, these Key Performance Indicators are identified, to monitored through an information visualization application, that retrieves the information from a datamart with a dimensional model. To carry out the process of developing this solution, a set of various tools and technologies were combined, protocols and good practices set out in the business intelligence application development methodology called Kimball.

The process of developing this solution begins by analyzing the current situation of the organization's information systems, considering its databases and processes. From there, proceed to identify and prioritize key performance indicators, and use information collection tools as a basis. Then proceed to identify requirements and develop data warehouse designs and ETL processes. Finally, visualizations are created for the highest priority indicators.

The use of this solution, is intended to improve the visibility of the organization's management, regarding the performance of the information systems related to the processing of cryptocurrencies, so that they have the ability to make decisions on real and updated evidence .

Keywords: Business intelligence, Cryptocurrency processing, Data warehouse

Dedicatoria

*A mi papá y a mi mamá; por todo su apoyo y esfuerzo durante mi vida,
ustedes son la mayor motivación que tengo
para alcanzar mis metas.*

*A mis hermanos, César y Antony; por ser mi mayor ejemplo de disciplina
y trabajo duro en los ámbitos académicos y profesionales.*

*A mi novia, Nati; por siempre estar ahí y apoyarme incondicionalmente,
sin importar las circunstancias,
gracias por todo el amor y apoyo que me ha brindado a través de los años.*

*Espero que este sea solo otro de los triunfos
que vamos a alcanzar juntos.*

Agradecimientos

A Fede, por abrirme las puertas de su organización hace más de dos años y por brindarme constantemente los instrumentos para crecer como profesional y como persona.

*A Edgar, y al resto del equipo de trabajo de Progressio Digital;
por todo su apoyo y consejos durante el desarrollo de este proyecto.*

*A mi tutora, Jaqueline; por su trabajo, dedicación y esfuerzo durante este período;
su guía, motivación y apoyo fueron clave
para alcanzar esta meta.*

*A Bryan Pérez, David Mata, Ronald Bolaños, Daniela Solís, David Corella, Daniel Corella, Magno
y a todos los amigos cercanos que dejó el TEC en mi vida
y que, de alguna forma,
me acompañaron a lo largo de este proceso.*

Carta de revisión filológica

Cartago, 28 de junio de 2020

A quien interese:

Yo, Gisela Alfaro Chaves, cédula de identidad 2-0701-0506 profesional en Filología Española y en Enseñanza del Castellano y la Literatura, perteneciente al Colegio de Licenciados y Profesores en Letras, Filosofía, Ciencias y Artes; leí y corregí el Proyecto Final de Graduación:

“Desarrollo de solución de inteligencia de negocios para interpretación, análisis y evaluación de indicadores clave de desempeño relacionados al procesamiento de criptomonedas. Caso: Progressio Digital”

Documento realizado por el estudiante: José Andrés Barboza González, Cédula de identidad: 207600470. Con el fin de optar por el grado de Licenciatura en Administración de Tecnología de la Información, del Instituto Tecnológico de Costa Rica, Campus Central Cartago.

Por este motivo, se revisaron y corrigieron aspectos como la construcción de párrafos, organización discursiva, vicios del lenguaje trasladados del campo oral al escrito, ortografía, puntuación, barbarismos, coherencia, cohesión y otros elementos relacionados con el campo filológico.

Realizadas las correcciones, doy fe de que el documento está listo para ser presentado.

Se suscribe de ustedes cordialmente,



Gisela Alfaro Chaves, céd. 207010506.

Carné 67138

Tabla de contenido

Resumen	iv
Abstract	v
Dedicatoria.....	I
Agradecimientos	II
Carta de revisión filológica	III
Tabla de contenido	IV
Índice de figuras.....	X
Índice de tablas.....	XII
1. CAPÍTULO I: INTRODUCCIÓN	15
1.1. Descripción General	15
1.2. Antecedentes.....	16
1.2.1 Descripción de la organización	16
1.2.1.1. Misión	16
1.2.1.2. Visión.....	16
1.2.1.3. Sobre la organización	16
1.2.1.4. Propuesta de valor.....	17
1.2.2. Equipo de trabajo	18
1.3. Proyectos similares	20
1.3.1. Desarrollo de herramienta de inteligencia de negocios para el área de ventas. Caso: Equifax USA	20
1.3.2. Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes. Caso: Empresa Áltica.....	20
1.4. Planteamiento del problema	21
1.4.1. Situación problemática	21
Problemáticas a nivel interno.....	21
Problemáticas a nivel externo.....	22
1.4.2. Beneficios esperados del proyecto.....	23
Beneficios directos	23
Beneficios indirectos	24
1.5. Objetivos del proyecto	25
1.5.1. Objetivo general.....	25

1.5.2. Objetivos específicos	25
1.6. Alcance del proyecto	26
1.7. Entregables del proyecto	28
1.7.1. Entregables de producto	28
1.7.1.1. Análisis de la situación actual	28
1.7.1.2. Indicadores clave de desempeño	28
1.7.1.3. Diseño e implementación del Datamart	28
1.7.1.4. Diseño e implementación de procesos ETL.....	28
1.7.1.5. Dashboards de visualización	29
1.7.2. Entregables académicos.....	29
1.7.2.1. Minutas.....	29
1.7.2.2. Cronograma.....	29
1.7.2.3. Gestión de cambios	29
1.8. Limitaciones del proyecto	30
1.9. Supuestos del proyecto	30
2. CAPÍTULO II: MARCO CONCEPTUAL	32
1.10. Análisis de requerimientos	32
1.10.1. Características de un requerimiento.....	32
1.11. <i>Blockchain</i>	34
1.11.1. Computación Peer-to-peer	36
1.11.2. Criptomonedas.....	36
1.11.2.1. Criptografía.....	36
1.11.2.2. Direcciones.....	37
1.11.2.3. Bitcoin (BTC)	37
1.11.2.4. Bitcoin cash (BCH)	38
1.11.2.5. Litecoin (LTC)	38
1.11.2.6. Ethereum (ETH)	38
1.11.2.7. Transacciones de criptomonedas	39
1.11.2.7.1. Confirmación de transacciones de criptomonedas.....	39
1.11.2.7.2. Comisión de las transacciones de criptomonedas	39
1.12. Inteligencia de negocios.....	41
1.12.1. Datos, información y conocimiento.....	41
1.12.2. Definición de inteligencia de negocios.....	41

1.12.3. Enfoques para implementar inteligencia de negocios.....	43
1.12.4. Enfoque Top-Down (Inmon)	43
1.12.5. Enfoque Bottom-Up (Kimball).....	43
1.12.6. Elementos tecnológicos de la inteligencia de negocios	46
1.12.6.1. Almacén de datos	46
1.12.6.2. Modelo OLTP	46
1.12.6.3. Modelo dimensional.....	46
1.12.6.3.1. Tablas de hechos	47
1.12.6.3.2. Tablas de dimensiones	47
1.13. Proceso de extracción, transformación y carga de datos (<i>ETL</i>).....	47
3. CAPÍTULO III: MARCO METODOLÓGICO.....	50
1.14. Tipo de investigación.....	50
1.15. Alcance de la investigación	51
1.16. Diseño de la investigación.....	52
1.17. Fuentes de información	54
1.17.1. Fuentes de información primarias	54
1.17.2. Fuentes de información secundarias	54
1.18. Sujetos de información	55
1.19. Instrumentos de recolección de datos	57
1.19.1. Entrevista.....	57
1.19.2. Revisión documental.....	58
1.19.3. Observación.....	58
1.19.4. Tabla comparativa de herramientas	59
1.19.5. Grupos focales.....	59
1.20. Variables de la investigación	60
1.21. Procedimiento metodológico	64
1.21.1. Análisis de la situación actual.....	64
1.21.2. Propuesta de indicadores claves de desempeño	65
1.21.3. Diseño de estructura del <i>datamart</i>	65
1.21.4. Diseño de procesos <i>ETL</i>	65
1.21.5. Elaboración de <i>dashboards</i> para la visualización de indicadores de desempeño.....	66
1.21.6. Relación entre el procedimiento metodológico y ciclo de vida de Kimball	67
1.22. Matriz metodológica	69

4.	CAPÍTULO IV: ANÁLISIS DE RESULTADOS.....	74
1.23.	Fase I: Análisis de la situación actual.....	74
1.23.1.	Hallazgos sobre la situación actual de los sistemas de información de la organización y sus procesos.....	75
1.24.	Fase II: Propuesta de indicadores clave de desempeño.....	76
1.25.	Fase III: Diseño del <i>datamart</i>	89
1.25.1.	Identificación de requerimientos relacionados con el diseño del <i>datamart</i>	89
1.26.	Fase IV: Diseño de los procesos <i>ETL</i>	91
1.26.1.	Identificación de requerimientos relacionados con el diseño de los procesos <i>ETL</i>	91
1.26.2.	Comparativa de herramientas para desarrollo de procesos ETL.....	92
1.27.	Fase V: Elaboración de <i>dashboards</i> para la visualización de los indicadores clave de desempeño.....	99
1.27.1.	Identificación de requerimientos relacionados a la elaboración de los <i>dashboards</i>	99
1.27.2.	Comparativa de herramientas para elaboración de los <i>dashboards</i>	101
5.	CAPÍTULO V: PROPUESTA DE SOLUCIÓN.....	108
1.28.	Fase III: Diseño del <i>datamart</i>	108
1.28.1.	Diseño de la arquitectura técnica.....	108
1.28.1.1.	Especificaciones técnicas.....	109
1.28.1.2.	Gestión de accesos.....	110
1.28.2.	Diseño del modelo dimensional.....	113
1.28.2.1.	Diseño de alto nivel.....	115
1.28.3.	Diseño físico.....	117
1.28.3.1.	Modelado dimensional.....	117
1.28.3.2.	Staging area o área de trabajo.....	120
1.28.4.	Implementación.....	121
1.29.	Fase IV: Diseño de los procesos <i>ETL</i>	124
1.29.1.	Diseño de la arquitectura técnica.....	124
1.29.1.1.	Ambiente de ejecución de los Jobs.....	124
1.29.1.2.	Manejo de versiones.....	125
1.29.2.	Diseño de procesos <i>ETL</i>	126
1.29.2.1.	Diseño de los Jobs de carga del Staging Area.....	127
1.29.2.2.	Diseño de los Jobs de carga del almacén de datos.....	129
1.29.3.	Implementación de los procesos <i>ETL</i>	133
1.29.3.1.	Liberación y ejecución de Jobs.....	133

1.29.3.2.	Carga inicial de los datos.....	137
1.29.3.3.	Comprobación de resultados de los procesos ETL.....	138
1.29.3.4.	Automatización de la ejecución de los Jobs.....	139
1.30.	Fase V: Elaboración de <i>dashboards</i> para la visualización de los indicadores clave de desempeño.....	141
1.30.1.	Diseño de la arquitectura técnica.....	141
1.30.2.	Desarrollo de la aplicación de visualización de datos.....	142
1.30.2.1.	Carga de información.....	142
1.30.2.1.	Creación de visualizaciones.....	148
1.30.3.	Implementación.....	153
6.	CAPÍTULO VI: CONCLUSIONES.....	158
7.	CAPÍTULO VII: RECOMENDACIONES.....	163
	Apéndices.....	165
	Apéndice I. Plantilla de minutas para las reuniones del proyecto.....	165
	Apéndice II. Plantilla para la gestión de cambios.....	166
	Apéndice III. Plantilla para entrevistas semiestructuradas.....	167
	Apéndice IV. Entrevista abierta.....	168
	Apéndice V. Revisión documental.....	169
	Apéndice VI. Observación.....	170
	Apéndice VII. Tabla comparativa de herramientas.....	171
	Apéndice VIII. Grupo Focal.....	172
	Apéndice IX. Análisis de la situación actual de los sistemas de la organización.....	173
	Apéndice X. Revisión documental de las especificaciones técnicas e implementación de los APIs del sistema de procesamiento de criptomonedas.....	175
	Apéndice XI. Revisión documental de las especificaciones técnicas e implementación de los APIS del monedero electrónico.....	176
	Apéndice XII. Revisión documental de los diseños lógicos de las bases de datos del sistema de procesamiento de criptomonedas.....	177
	Apéndice XIII. Revisión documental de los diseños lógicos de las bases de datos del monedero electrónico.....	178
	Apéndice XIV. Revisión documental consulta para revisión de transacciones del sistema de procesamiento de criptomonedas.....	179
	Apéndice XV. Revisión documental ejecución de retiros planificados del monedero electrónico.....	180
	Apéndice XVI. Revisión documental sobre consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.....	181

Apéndice XVII. Observación del flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico con desarrollador de <i>software</i>	182
Apéndice XVIII. Grupo focal para la revisión y priorización de los <i>KPI</i> propuestos con gerente general, administrador de proyectos y desarrollador de <i>software</i>	183
Apéndice XIX. Grupo focal para la definición de requerimientos del <i>datamart</i> con administrador de proyectos, gerente general e ingenieros <i>DevOps</i>	184
Apéndice XX. Grupo focal con administrador de proyectos e ingenieros <i>DevOps</i> para la definición de requerimientos relacionados al desarrollo de los procesos <i>ETL</i>	185
Apéndice XXI. Grupo focal con administrador de proyectos y gerente general para para la definición de requerimientos relacionados las visualizaciones.	187
Apéndice XXII. Bitácora de reuniones y comunicaciones con la organización.....	189
Apéndice XXIII. Minuta de reunión 02 con la organización	192
Apéndice XXIII. Minuta de reunión 03 con la organización	193
Apéndice XXIV. Minuta de reunión 04 con la organización.....	194
Apéndice XXV. Minuta de reunión 05 con la organización.....	195
Apéndice XXVI. Minuta de reunión 06 con la organización.....	196
ANEXOS.....	197
Anexo I.....	197
Anexo II.....	198
Anexo III.....	199
GLOSARIO	200
BIBLIOGRAFÍA.....	201

Índice de figuras

Figura 1. Equipo de trabajo para el proyecto	18
Figura 2. Mapa conceptual sobre la problemática	23
Figura 3. Ejemplo de la estructura de los bloques en Blockchain	35
Figura 4. Representación de una red Peer-to-Peer	36
Figura 5. Ciclo de vida del Enfoque Kimball	44
Figura 6. Almacén de datos con proceso ETL	48
Figura 7. Procedimiento metodológico	64
Figura 8. Niveles de prioridad para los indicadores clave de desempeño	85
Figura 9. Detalles de instalación de la instancia del datamart	109
Figura 10. Detalles de la configuración de la instancia RDS del <i>datamart</i>	110
Figura 11. Configuración de accesos de la conexión a la instancia <i>RDS</i>	110
Figura 12. Intento fallido de conexión a las bases de datos del datamart.....	111
Figura 13. Intento exitoso de conexión a las bases de datos del <i>datamart</i>	112
Figura 14. Modelo estrella de alto nivel sobre transacciones del sistema de procesamiento de criptomonedas	115
Figura 15. Modelo estrella de alto nivel para las transacciones del sistema de monedero electrónico de criptomonedas	116
Figura 16. Diseño físico del modelo estrella de transacciones del procesador de criptomonedas ...	118
Figura 17. Diseño físico del modelo estrella de transacciones del monedero de criptomonedas	119
Figura 18. Tablas del área de trabajo.....	120
Figura 19. Definición de la tabla de hechos de transacciones del procesador de criptomonedas....	121
Figura 20. Definición de la tabla de hechos de transacciones del monedero de criptomonedas.....	122
Figura 21. Definición de llaves foráneas entre la tabla de hechos de transacción del procesador de criptomonedas y sus dimensiones	123
Figura 22. Definición de llaves foráneas entre la tabla de hechos de transacción del monedero de criptomonedas y sus dimensiones	123
Figura 23. Características de la instancia configurada por los ingenieros devops	124
Figura 24. Conexión fallida al no contar con conexión a la VPN.....	125
Figura 25. Conexión exitosa al servidor con la VPN activa.....	125
Figura 26. Repositorio de los Jobs ETL dentro de <i>Bitbucket</i>	126
Figura 27. Conexiones a las bases de datos configuradas para extraer o cargar los datos.....	127
Figura 28. Job encargado de cargar los datos correspondientes a la tabla de transacción para el sistema de procesamiento de criptomonedas	128
Figura 29. Detalle del componente <i>sbTxMap</i>	128
Figura 30. Carga de una de las dimensiones del almacén de datos.....	129
Figura 31. Carga de los datos de la tabla de estatus directa	129
Figura 32. Job para la carga de la tabla de hechos de transacción correspondiente al sistema de monedero de criptomonedas.....	130
Figura 33. Mapeo de los datos de la tabla fuente a la tabla de origen	131
Figura 34. <i>Job</i> diseñado para la carga de información de la tabla de hechos de las transacciones del procesador de criptomonedas.....	132
Figura 35. Mapeo entre el origen de los datos y la tabla destino	133

Figura 36. Opción de construcción de Jobs en Talend para su exportación	134
Figura 37. Ventana para la exportación de los Jobs	134
Figura 38. Proceso para subir la nueva versión del job construido a través de la consola de una terminal de <i>git</i>	135
Figura 39. Interfaz de <i>Jenkins</i>	136
Figura 40. Directorio de los archivos dentro del servidor	136
Figura 41. Cambios realizados en la consulta de carga de datos del <i>job</i>	137
Figura 42. Comparativa de los datos para la dimensión de categoría de cliente del sistema procesador de criptomonedas	138
Figura 43. Cambio en la consulta para actualizar e ingresar las tablas de hechos transacciones ...	140
Figura 44. Modalidad de inserción de los <i>Jobs</i> configurada en <i>Talend</i>	140
Figura 45. Ubicaciones disponibles para el servicio de <i>Quicksight</i>	141
Figura 46. Notificación de registro del servicio de <i>Quicksight</i>	142
Figura 47. La configuración de una nueva fuente de información MySQL.....	143
Figura 48. Carga inicial de información en el set de datos de <i>Quicksight</i>	144
Figura 49. Fuentes de información importadas para el desarrollo de las visualizaciones	144
Figura 50. Ventana de importación de nuevas tablas de datos	145
Figura 51. Configuración de un <i>join</i> entre la dimensión entre dos tablas importas en <i>Quicksight</i> ...	146
Figura 52. Fuentes de datos relacionadas con el modelo estrella de las transacciones del monedero de criptomonedas	147
Figura 53. Visualización para el indicador IND-21: Cantidad promedio pagada por comisión de mineros.....	148
Figura 54. Cantidad de retiros por categoría de cliente y cantidad de transacciones por categoría de cliente	149
Figura 55. Monto de dinero de depósitos por categoría de cliente y monto de dinero de retiros por categoría de cliente.....	150
Figura 56. Filtros de moneda, de tipo de transacción y de año.....	151
Figura 57. Indicador IND-28: Cantidad de transacciones de compra dinero cripto	151
Figura 58. Indicador IND-27: Monto de las transacciones de compra de dinero cripto procesadas .	152
Figura 59. Indicador IND-29: monto promedio de las transacciones de compra de dinero cripto.....	153
Figura 60. Refrescamiento de los datos de manera calendarizada	154
Figura 61. Ejemplo de configuración de envío por correo electrónico de las visualizaciones.....	155
Figura 62. Invitaciones al correo electrónico	156
Figura 63. Vista de los dashboards para los usuarios con permisos de lectura solamente.....	156

Índice de tablas

Tabla 1. Roles de los miembros del equipo	19
Tabla 2. Características de los requerimientos	33
Tabla 3. Comparativa entre criptomonedas BTC, BCH, LTC y ETH.	38
Tabla 4. Alcances de una investigación.....	51
Tabla 5. Descripción de los diseños de investigación	52
Tabla 6. Sujetos de información	55
Tabla 7. Cuadro de Variables	60
Tabla 8. Relación entre el proceso metodológico y el ciclo de vida de Kimball.	68
Tabla 9. Matriz metodológica del proyecto.....	69
Tabla 10. Listado de indicadores clave de desempeño para los sistemas de procesamiento de criptomonedas.	77
Tabla 11. Priorización de los indicadores clave de desempeño	86
Tabla 12. Requerimientos relacionados con el diseño del datamart	89
Tabla 13. Requerimientos relacionados con el diseño de procesos ETL	91
Tabla 14. Criterios de selección de herramientas para desarrollo de ETL	93
Tabla 15. Justificación de la escogencia de las herramientas de ETL analizadas.....	94
Tabla 16. Tabla comparativa de herramientas para el desarrollo de procesos ETL.....	95
Tabla 17. Comparativa de herramientas de desarrollo de ETL con su respectivo peso	98
Tabla 18. Requerimientos relacionados con la elaboración de dashboards.....	99
Tabla 19. Criterios de escogencia de herramientas de dashboards y su peso.....	101
Tabla 20. Justificación de escogencia de herramientas de visualización de datos a evaluar	102
Tabla 21. Tabla comparativa de herramientas para el desarrollo de dashboards.	103
Tabla 22. Comparativa de herramientas de desarrollo de dashboards con su respectivo peso	106
Tabla 23. Indicadores clave de desempeño con mayor nivel de prioridad	113
Tabla 24. Tablas de hechos y sus métricas	114
Tabla 25. Resultado de la ejecución de las pruebas en las tabla de origen y en la tabla de hechos de las transacciones del procesador de criptomonedas.....	139
Tabla 26. Cantidades de filas para la tabla de origen de las transacciones del procesador de criptomonedas	139
Tabla 27. Plantilla para la gestión de minutas de reunión	165
Tabla 28. Plantilla para la gestión de solicitudes de cambio	166
Tabla 29. Plantilla para entrevistas semiestructuradas	167
Tabla 30. Entrevista abierta	168
Tabla 31. Revisión documental.....	169
Tabla 32. Observación.....	170
Tabla 33. Tabla comparativa de herramientas.....	171
Tabla 34. Grupo Focal.....	172
Tabla 35. Entrevista sobre situación actual de los sistemas de la organización.....	173
Tabla 36. Detalle de la revisión documental del detalle del APIs del sistema de procesamiento de criptomonedas	175
Tabla 37. Revisión documental de las especificaciones técnicas e implementación de los APIS del monedero electrónico	176

Tabla 38. <i>Revisión documental de los diseños lógicos de las bases de datos del sistema de procesamiento de criptomonedas.</i>	177
Tabla 39. <i>Revisión documental de los diseños lógicos de las bases de datos del monedero electrónico.</i>	178
Tabla 40. <i>Consulta para revisión de transacciones del sistema de procesamiento de criptomonedas.</i>	179
Tabla 41. <i>Ejecución de retiros planificados del monedero electrónico</i>	180
Tabla 42. <i>Revisión documental sobre consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.</i>	181
Tabla 43. <i>Observación del flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico.</i>	182
Tabla 44. <i>Grupo focal para la revisión y priorización de los KPI propuestos</i>	183
Tabla 45. <i>Grupo focal para la definición de requerimientos del datamart.</i>	184
Tabla 46. <i>Grupo focal para la definición de requerimientos relacionados al desarrollo de los procesos ETL</i>	185
Tabla 47. <i>Grupo focal para para la definición de requerimientos de relacionados las visualizaciones</i>	187

CAPÍTULO I: INTRODUCCIÓN

1. CAPÍTULO I: INTRODUCCIÓN

1.1. Descripción General

Progressio Digital es una empresa que optó por la innovación, al brindar al mercado una solución de procesamiento de criptomonedas en el *blockchain*. Por medio de los sistemas de la organización se gestiona una importante cantidad de dinero de forma segura y eficiente; sin embargo, hasta el momento no se ha decidido por realizar ningún tipo de análisis a los datos resultantes de estos procesos.

Los datos almacenados como resultado de las operaciones de los sistemas de la organización se manejan de manera independiente; esto debido a que no se cuenta con una solución que permita estandarizar. Además, no existe un mecanismo dentro de la organización con la capacidad de identificar, efectivamente, el tráfico no convencional en las redes de *blockchain*, esto provoca una inversión de tiempo, recursos humanos y monetarios para el procesamiento forzado de las transacciones.

Por otro lado, se tiene certeza de que el servicio ofrecido a los clientes cumple con las expectativas de rendimiento; sin embargo, no se sabe cuál es el desempeño a nivel cuantitativo que se les está ofreciendo a estos, ni se cuenta con métricas que valoren el funcionamiento de los sistemas y que permitan la definición y seguimiento de estrategias organizacionales.

La anterior, es una de las principales razones por las que no se han definido explícitamente acuerdos de nivel de servicio para medir el cumplimiento de los tiempos de respuesta y ejecución. La definición de este tipo de acuerdos entre Progressio Digital y quienes consumen sus sistemas, resulta relevante para aumentar la confianza entre estos, así como para atraer nuevos prospectos. Se emplean datos reales, actualizados y relevantes sobre el funcionamiento y los desempeños presentes; de igual forma, estos permitirían negociar los niveles de uso por parte de todos estos clientes, con respecto al volumen de procesamiento.

En resumen, lo que se busca con el desarrollo de la solución, es aplicar estrategias de inteligencia de negocios para apoyar, internamente, a la empresa en el proceso de toma de decisiones a nivel operacional y gerencial, de una manera oportuna e integral. Por tanto, se pretende mejorar la experiencia de los clientes de la empresa al contar con la capacidad de brindarles un servicio transparente, entendible y útil para su propio proceso de toma de decisiones; esto último, por medio de la generación de reportes personalizados y automatizados, donde se visualice información relevante con respecto al manejo de su dinero y del desempeño de los sistemas de Progressio Digital.

1.2. Antecedentes

En esta sección se lleva a cabo una contextualización sobre la organización en la cual se realizará el Trabajo Final de Graduación.

1.2.1 Descripción de la organización

En esta sección se describen aspectos relacionados con la empresa Progressio Digital, se detallan aspectos generales como misión, visión, valores e historia.

1.2.1.1. *Misión*

La siguiente es la misión de Progressio Digital (2020):

“Ofrecer soluciones para el procesamiento de criptomonedas que sean fáciles de utilizar, seguras y asequibles para el mercado” (párr. 1)

1.2.1.2. *Visión*

La siguiente es la visión de Progressio Digital (2020):

“Ser uno de los principales promotores del mundo sobre el uso de las criptomonedas y de otras soluciones basadas en el *blockchain* para facilitar el diario quehacer de personas y organizaciones.” (párr.2)

1.2.1.3. *Sobre la organización*

Progressio Digital es una empresa que inició formalmente sus operaciones en enero de 2018, se dedica principalmente a brindar servicios de procesamiento de criptomonedas. Federico López-Calleja Salazar, gerente general de Progressio Digital, fundó la organización luego de identificar, dentro de la empresa donde en ese entonces se encontraba laborando, la ausencia de una herramienta, en el mercado de procesamiento de criptomonedas, que fuese fácil de implementar y adaptar a la realidad organizacional, así como asequible a nivel económico.

En ese momento se desconocía de alguna persona en el país que estuviese trabajando formalmente en soluciones de *blockchain*; además, se contaba con conocimiento acerca de la competencia internacional, debido al trabajo en la implementación de esas soluciones dentro de la organización.

En un principio, la herramienta que se desarrolló para solventar la necesidad identificada soportaba el procesamiento de Bitcoin, la criptomoneda más popular, utilizada mundialmente para llevar a cabo transacciones. Posteriormente, fueron implementadas otras dos: Bitcoin Cash y Litecoin, monedas que resultaron de variables del protocolo Bitcoin, evento llamado técnicamente como Fork. Por último, bajo la necesidad y persistencia del cliente, se procedió a implementar la criptomoneda conocida como Ethereum. Según un análisis de capitalización del mercado de las criptomonedas, las monedas soportadas por Progressio Digital actualmente, se encuentran entre los primeros seis lugares con mayor relevancia. (Top 100 Cryptocurrencies by Market Capitalization, 2019)

Cabe destacar que actualmente en Progressio Digital se trabajan los siguientes proyectos principalmente relacionados con el procesamiento de criptomonedas:

- **Procesador de criptomonedas:** es un producto ofrecido principalmente por medio de una interfaz de programación de aplicaciones, o *API*, por el cual es posible realizar transacciones a través de la red de *blockchain*, tanto de retiro como de depósito, de todas las monedas mencionadas con anterioridad. Esta solución cuenta con su propio sitio administrativo de consulta de información.
- **Sistema de compra y envío de criptomonedas:** este sistema permite la compra de criptomonedas por medio de cargos en tarjetas de crédito, así como un envío seguro, eficiente y controlado de dinero a través de *blockchain*. Al realizar una transacción de compra de criptomoneda, la transacción de la tarjeta de crédito, llevada a cabo a través de un procesador tercerizado, se ve reflejada por una transacción en el *blockchain*. Este sistema consume los servicios del sistema de monedero para el control de cuentas y balances de dinero, así como para la gestión de transacciones criptográficas.

1.2.1.4. Propuesta de valor

La organización se basa en los siguientes valores:

- Enfoque en las necesidades del cliente.
- Soporte directo y eficiente de los requerimientos del cliente.
- Seguridad y transparencia.
- Trato personalizado y enfocado en obtener resultados positivos para los clientes.

1.2.2. Equipo de trabajo

En esta sección se presenta una descripción del equipo de trabajo involucrado en el desarrollo del proyecto. La **Figura 1**, modela este equipo, así como la interrelación entre cada uno de los miembros:

Figura 1. Equipo de trabajo para el proyecto



Figura 1. Equipo de trabajo para el proyecto. Fuente: elaboración propia. (2020).

Seguidamente, en la **Tabla 1**, se presenta la descripción detallada del rol de cada uno de los miembros del equipo que están relacionados con el desarrollo del proyecto, tanto dentro de la empresa como en el presente trabajo.

Tabla 1. Roles de los miembros del equipo

Miembro	Rol en la empresa	Rol en el proyecto
Director general	Es el gerente y, como tal, el principal contacto con los clientes de la empresa. Cuenta con un panorama general de todos los proyectos y de cómo estos se complementan, también del tipo de información que es almacenada en las bases de datos y de su relevancia para el funcionamiento de todos los sistemas. Patrocinador del proyecto.	Supervisor dentro de la empresa para el Trabajo Final de Graduación, es el encargado de aprobar, rechazar o debatir las decisiones a tomar con respecto al proyecto.
Administrador de proyectos	Es la persona que se encuentra directamente a cargo de la dirección de todos los proyectos dentro de la empresa.	Fuente de información relevante para llevar a cabo el proyecto. De igual forma, cuenta con conocimientos amplios en el área de inteligencia de negocios, de modo que posee la capacidad para orientar el rumbo del proceso de ejecución del proyecto planteado.
Desarrolladores de software	Dentro de la empresa, existen cuatro desarrolladores de <i>software</i> que han aportado su trabajo en la creación de las soluciones actualmente ofrecidas. Estos desarrolladores participaron, de una manera u otra, en los procesos de planeación, ejecución, pruebas y soporte de estas soluciones.	Son una fuente de información relevante para llevar a cabo el proyecto de inteligencia de negocios que permita el análisis de los datos relacionados con los sistemas que han desarrollado y a los que se les da mantenimiento y soporte.
Estudiante	Desarrollador de <i>software</i> .	Desarrollador del proyecto.

Nota. Fuente: elaboración propia. (2020).

1.3. Proyectos similares

En esta sección se describen proyectos con una naturaleza similar al Trabajo Final de Graduación y que se han llevado a cabo dentro o fuera de la organización.

Hasta el momento, en Progressio Digital nunca se ha llevado a cabo un proyecto relacionado con el área de inteligencia de negocios; sino que, los que han sido desarrollados anteriormente corresponden a sistemas de *software* enfocados, mayormente, en manejo de transacciones de *blockchain*, conexión con *API* e interfaces gráficas; es decir, no existe en la organización ninguna solución que se base en el tema tratado en este Trabajo Final de Graduación.

A continuación, se presentan proyectos similares que se ejecutaron fuera de la organización, estos se obtuvieron luego de la investigación sobre el tema, asimismo, es importante aclarar que serán utilizados como insumos para llevar a cabo la solución de inteligencia de negocios.

1.3.1. Desarrollo de herramienta de inteligencia de negocios para el área de ventas. Caso: Equifax USA

Equifax es una empresa global de soluciones de información que recopila datos de consumidores, empresas y archivos de empleadores en todo el mundo. Con base en el proyecto, se brindó solución a un problema relacionado con el departamento de ventas y la consolidación de información, a través del uso de herramientas de inteligencia de negocios y el análisis respectivo de esta información.

1.3.2. Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes. Caso: Empresa Ática.

Este proyecto brindó una solución a las necesidades y problemática que se presentaban en la empresa Ática Socios Financieros S.A para la generación de reportes. Primeramente, se llevó a cabo un análisis del entorno organizacional con la definición de roles, frente a la información que se registraba y que también era requerida; además se incluyó la definición de los sistemas de información con la estructura de datos que los componen.

Con los resultados de dicho análisis se desarrollan dos propuestas. La primera, es una propuesta que ataca todas las causas del problema con un alcance integrador a largo plazo; la segunda, busca resolver necesidades actuales con un enfoque *ad hoc* y con un alcance a corto plazo. (Acuña, 2018)

1.4. Planteamiento del problema

En esta sección se describe la problemática encontrada dentro de la organización. Esta situación genera una necesidad de solventar o brindar una solución de inteligencia de negocios en la organización.

1.4.1. Situación problemática

Dada la naturaleza digital de los sistemas que son ofrecidos por Progressio Digital, esta organización debe encargarse de almacenar una gran cantidad de datos que permitan la continuidad del funcionamiento del negocio; sin embargo, estos datos se emplean para fines meramente operativos.

Progressio Digital ofrece a sus clientes un sitio de gestión administrativa que permite la generación de reportes y hojas de cálculo con información solamente del sistema de procesamiento de criptomonedas; sin embargo, los datos son presentados en pantalla de una forma tabular, es decir, de manera idéntica a como son almacenados originalmente en la fuente de información. Esto es debido a que la herramienta es utilizada por la empresa cliente dentro de los departamentos de contaduría y servicio al cliente, con el fin de llevar un proceso contable y consultar estos datos específicos, como los relacionados con una transacción en particular.

Por otro lado, Progressio Digital se ha respaldado en el conocimiento técnico con que cuenta su gerente general y otros colaboradores, cuando es necesario llevar a cabo consultas, con mayor nivel de complejidad, que no se pueden obtener directamente del sistema administrativo y son requeridas por los clientes o para la toma de decisiones en un momento determinado.

Lo anterior conduce a la problemática por la cual se pretende llevar a cabo el presente proyecto, es decir: la **nula visibilidad sobre el desempeño de los sistemas actuales de la empresa**. A partir de esta situación expuesta, se derivan otras problemáticas que afectan directamente la operación del negocio e inclusive la relación con el cliente, algunas de estas problemáticas son las siguientes:

Problemáticas a nivel interno

- Dificultad para la toma de decisiones gerenciales y para la definición de proyectos y actividades que se adapten a las necesidades reales de la empresa y a sus estrategias organizacionales.

- Las consultas necesarias para obtener información en específico se dificultan, y se traduce a una inversión de recursos y tiempo, elementos importantes de emplear en otras tareas.
- Existe información que no se tiene disponible en el momento que es requerida, y que podría ser de utilidad para la toma de decisiones.
- Los sistemas de la empresa están ligados internamente para llevar a cabo operaciones del negocio; sin embargo, no existe forma de visualizar el comportamiento que se da entre estos sistemas; es decir, no existe una integración de los datos a nivel histórico que muestre cómo los sistemas se relacionaron a lo largo de periodos particulares de tiempo.

Problemáticas a nivel externo

- Existe un nivel de incertidumbre sobre la manera en que serán procesadas las transacciones, cuando los sistemas están directamente ligados al *blockchain*, una red que cuenta con un comportamiento errático. Como resultado de una falta de visibilidad sobre el comportamiento de la red, es necesario tomar medidas reactivas de manera acelerada para atacar la aparición de un problema de mayor impacto, traduciéndose esta problemática en una inversión de tiempo, esfuerzo y, en algunos casos, de recursos económicos.
- El personal, de análisis del negocio, por parte del cliente, debe llevar a cabo estudios y análisis complicados sobre los datos generados en el sitio administrativo, o como resultado de las transacciones realizadas, para obtener patrones con base en los cuáles tomar decisiones que permitan proceder con el uso del sistema.
- El cliente debe consultar otras fuentes para obtener información, como históricos de los tipos de cambios de las monedas, cuando es posible presentarla visualmente y de manera directa, como parte de una mejora en el servicio.
- A nivel gerencial, el cliente no tiene una clara visualización sobre el desempeño con que los sistemas de Progressio Digital atienden sus necesidades; si se deseara presentar el beneficio obtenido por el uso del sistema a la entidad gerencial del cliente, ya sea por medio de reportes o gráficos entendibles, sería necesario que estos fueran generados de manera manual bajo un proceso de selección y carga de datos.

En la **Figura 2**, se muestra un mapa conceptual con la información resumida sobre la problemática expuesta anteriormente; donde se exponen las problemáticas a nivel interno y externo que se derivan del problema principal.

Figura 2. Mapa conceptual sobre la problemática

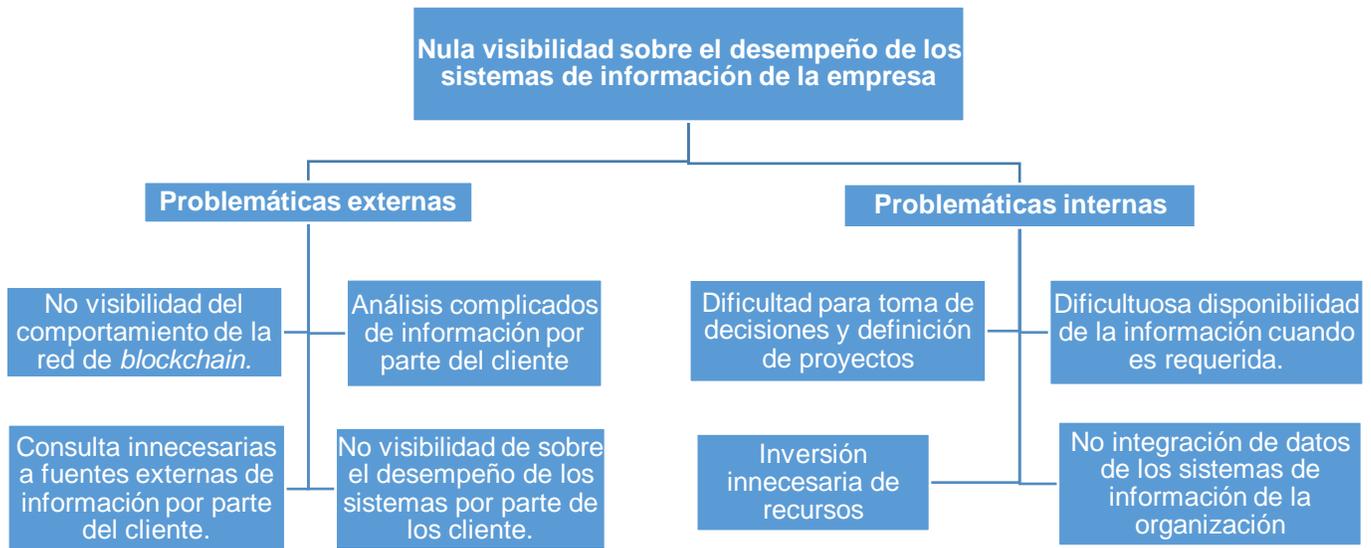


Figura 2. Mapa conceptual sobre la problemática. Fuente: elaboración propia. (2020).

1.4.2. Beneficios esperados del proyecto

Al identificar explícitamente cuáles son algunas de las problemáticas actuales con que cuenta la empresa, se pretende desarrollar una solución de inteligencia de negocios que permita analizar e interpretar datos de manera automatizada, se podrían alcanzar los siguientes beneficios:

Beneficios directos

- Integración y consolidación automatizada de información, proveniente de los sistemas de la empresa, que permita obtener visibilidad sobre el rendimiento ofrecido y las oportunidades de mejora.
- Identificación de proyectos y actividades que se alineen con las estrategias y metas empresariales de organización.
- Brindar un acceso más sencillo.
- Transparencia de cara con el cliente con respecto al desempeño de los sistemas; así como apoyo en su proceso de toma de decisiones, al brindarles acceso a información fácil de interpretar.
- La definición de métricas para el control y monitoreo de las soluciones, que permitan consolidar estrategias a nivel gerencial.

Beneficios indirectos

- Identificación de oportunidades de mejora en el desempeño de los sistemas con base en el análisis de la información.
- Desarrollo de una cultura organizacional de toma de decisiones con base en el análisis de datos reales.
- Identificación temprana de problemas como los presentados en la red de *blockchain*, que permita tomar medidas proactivas y evitar con ello la inversión innecesaria de recursos, dinero y tiempo.

1.5. Objetivos del proyecto

A continuación, son definidos tanto el objetivo general que se pretende alcanzar con la elaboración del Trabajo Final de Graduación, como sus correspondientes objetivos específicos.

1.5.1. Objetivo general

Desarrollar una solución de inteligencia de negocios que facilite la interpretación, el análisis y la evaluación de indicadores clave de desempeño relacionados con el procesamiento de criptomonedas para el favorecimiento de la toma de decisiones gerenciales en la empresa Progressio Digital, durante el primer semestre del 2020.

1.5.2. Objetivos específicos

- I. Analizar los sistemas de información de la empresa, sus procesos y sus correspondientes bases de datos para el planteamiento de la situación actual que oriente el desarrollo de la solución de inteligencia de negocios.
- II. Proponer indicadores clave de desempeño sobre los sistemas de la organización para que estos sirvan como apoyo en procesos de control, monitoreo y toma de decisiones en la organización.
- III. Diseñar el modelo del *datamart*, así como los procesos de extracción, transformación y carga de datos (*ETL*) para soportar la carga, estandarización y actualización de la información proveniente de los sistemas de la empresa.
- IV. Elaborar *dashboards* con vistas gráficas de indicadores de desempeño para el favorecimiento del proceso de análisis de la información y la toma de decisiones gerenciales, con respecto al procesamiento de criptomonedas.

1.6. Alcance del proyecto

En esta sección se describe en detalle el alcance que se abarcará con la realización del Trabajo Final de Graduación.

El actual proyecto consiste en el desarrollo de una solución de inteligencia de negocios que permita llevar a cabo la interpretación, el análisis y la evaluación de los datos relacionados con el procesamiento de criptomonedas dentro de la empresa Progressio Digital.

En primera instancia, para la realización del proyecto, se llevará a cabo un análisis de la situación actual de la empresa, donde se incluya los procesos actuales de análisis de información; así como el estado de los sistemas de información y sus correspondientes bases de datos.

Por otro lado, para el desarrollo de la solución de inteligencia de negocios propuesta, se trabajará bajo la metodología llamada Kimball, la cual cuenta con un enfoque *bottom-up* donde se preparan y crean conjuntos de datos iniciales que es posible desarrollar de manera incremental, y eventualmente unir en una sola fuente consolidada de información. (Sansu, 2012).

Esta metodología Kimball discutida con anterioridad, define un ciclo de vida para la implementación de soluciones de inteligencia de negocios, sobre el cual se trabajó para la construcción del proceso metodológico del desarrollo del proyecto. Las fases definidas abarcan el ciclo de vida hasta la etapa de implementación de la solución; asimismo, otras etapas posteriores como el crecimiento y el mantenimiento de la solución, quedan fuera del alcance para su posterior implementación en proyectos futuros de la compañía.

Para llevar a cabo lo anterior, se integrará los datos provenientes de los sistemas de información que maneja la empresa relacionados con el procesamiento de criptomonedas. Con la finalidad de entender a fondo los datos que van a ser analizados por medio de la solución de inteligencia de negocios, se explicará la funcionalidad de estos sistemas a continuación:

- **El procesador de criptomonedas:** es el principal sistema de la empresa, este es ofrecido por medio de un *API* a los clientes, dentro de este se generan transacciones, tanto de tipo depósito como de retiro, por sumas de dinero elevadas para cuatro principales monedas: *Bitcoin*, *Bitcoin Cash*, *Litecoin* y *Ethereum*. Este sistema cuenta con su propio sitio administrativo para el cliente, además, constituye la principal fuente de datos que sustentará la solución de inteligencia de negocios.

- **Monedero electrónico de criptomonedas:** esta solución permite la compra de criptomonedas con tarjetas de crédito, así como un envío directo de dinero por medio de la red de *blockchain*. Para realizar lo anterior, el sistema se encuentra relacionado directamente con el monedero de criptomonedas, el cuál gestiona los balances y cuentas desde las cuáles se llevan a cabo las transacciones; de igual forma, este consume los servicios de procesadores externos de tarjetas de crédito que generan transacciones bancarias.

Con el proyecto se pretende alcanzar la estandarización de la información de los diferentes sistemas por medio de procesos de extracción, transformación y carga de datos (*ETL*), estos procesos serán llevados a cabo por medio de la herramienta de unificación de datos brindada por la empresa Talend, llamada *Open Data Studio for Data Integration*.

De igual forma, se definirán indicadores de desempeño, así como reportes y tableros de información que serán entregados a la gerencia general y a algunos miembros del equipo de desarrollo de empresa, para el apoyo de sus operaciones, tal y como se mencionó anteriormente. Para llevar a cabo estas visualizaciones de información se utilizará la herramienta *Amazon Quicksight*. Además, es necesario contemplar la generación automatizada de estos reportes y su distribución por medio de correo electrónico a clientes y demás interesados de cada sistema.

En resumen, las fases que se llevarán a cabo para la realización del proyecto son las siguientes:

- Análisis de la situación actual.
- Propuesta de indicadores claves de desempeño.
- Diseño de estructura del *Datamart*.
- Diseño de procesos *ETL*.
- Elaboración de *dashboards* para la visualización de indicadores de desempeño.

1.7. Entregables del proyecto

En esta sección se describirá la gestión de los entregables del proyecto, los cuales, se dividirán en dos categorías: entregables académicos y entregables de producto.

1.7.1. Entregables de producto

Con el fin de llevar un manejo eficiente sobre la gestión del proyecto y el debido control de todas las etapas de este, se hará uso de los siguientes documentos:

1.7.1.1. *Análisis de la situación actual*

Identificación y definición de los sistemas de la empresa relacionados con el procesamiento de criptomonedas, así como modelado de sus correspondientes bases de datos y definición de sus principales operaciones.

1.7.1.2. *Indicadores clave de desempeño*

Indicadores explícitamente definidos para el control de operaciones del negocio relacionadas con el procesamiento de criptomonedas, para estos indicadores se definirá un nivel de priorización y se identificará a los de mayor urgencia para ser abarcados en el resto de la solución propuesta.

1.7.1.3. *Diseño e implementación del Datamart*

Se entregará e implementará el modelo dimensional con base en los requerimientos del negocio. Este modelo es el encargado de soportar la carga de datos por parte de los procesos ETL; además, es la fuente para la carga de datos que se presentará en los *dashboards*.

1.7.1.4. *Diseño e implementación de procesos ETL*

Para la elaboración del proyecto es necesario el diseño de procesos ETL, para la manipulación de los datos, los diseños resultantes de esto se refieren a un entregable del proyecto. Para la gestión de estos procesos, también será necesaria la definición de un proceso de control de versiones, así como un proceso de automatización de la carga de información. Por otro lado, se entregará un análisis de las alternativas con respecto a herramientas para la gestión de estos procesos.

1.7.1.5. Dashboards de visualización

Se diseñará *dashboards* de visualización de la información manipulada a través de los procesos *ETL*, considerando el tipo de visualización adecuada para los datos de análisis. Además, estos incluyen la automatización en la carga de la información y la compartición por correo electrónico a los interesados.

1.7.2. Entregables académicos

En esta sección se deberán describir los entregables en cuanto al documento académico del proyecto. A continuación, se detallan:

1.7.2.1. Minutas

A lo largo del transcurso del proyecto, se realizarán minutas con información relevante acerca de las reuniones realizadas con las personas involucradas en el proyecto. Dentro de las reuniones se debe contemplar las realizadas en el ámbito académico y las ejecutadas dentro de la organización. Estas minutas utilizarán la plantilla definida en el 12.1. Apéndice I. Plantilla de minutas para las reuniones del proyecto.

1.7.2.2. Cronograma

El cronograma del proyecto tiene como finalidad estructurar las fechas o plazos de entrega de los distintos productos entregables del proyecto, de forma que se genere un control de estos y se tenga mejor entendimiento de la situación del proyecto con respecto a lo planeado. Este cronograma funciona como insumo para la gestión de cambios del proyecto, dado que constantemente se tendrá que comparar el estado actual con el definido en la planeación.

1.7.2.3. Gestión de cambios

Con la finalidad de mantener un control a lo largo del desarrollo del proyecto sobre los cambios que ocurran, se utilizará una plantilla para su gestión, en esta se detalla información relevante. Finalmente, es posible visualizar esta plantilla en 12.2. Apéndice II. Plantilla para la gestión de cambios.

1.8. Limitaciones del proyecto

A continuación, se definen las limitaciones que se deben contemplar al momento de llevar a cabo el proyecto.

- Disponibilidad o apoyo por parte de los miembros del equipo para abarcar la completitud de la solución final.
- Debido a la naturaleza financiera de los datos que van a analizarse para el proyecto, es necesario mantener confidencialidad, por ello, la información mostrada en la propuesta de solución fue manipulada.
- Dependencia de la disponibilidad de representantes del cliente debido a la situación mundial de pandemia, para validar la información por incluir en los tableros y los reportes de la solución final.
- El proyecto debe apegarse al principal proveedor de servicios en la nube de la organización para configuración de los servidores necesarios, *Amazon Web Services (AWS)*.
- No existe un registro histórico sobre el tiempo invertido por parte de los colaboradores o el gerente general en la ejecución de consultas para obtener información, ni es posible estimar una cantidad de horas semanales porque este corresponde a un dato que varía.

1.9. Supuestos del proyecto

Los siguientes puntos son aspectos tratados como supuestos para la realización del proyecto:

- La organización brindará todas las herramientas tecnológicas necesarias para llevar a cabo el Trabajo Final de Graduación, esto incluye, pero no se limita a software, *hardware*, servicios y licencias.
- La información necesaria por parte de la organización para la implementación del proyecto se entregará en el momento oportuno.
- Se realizarán reuniones periódicas, solicitadas con una o dos semanas de anticipación por el estudiante, en las que se llevará a cabo revisiones y validaciones de los avances y consultas de información o dudas específicas con respecto al proyecto.

CAPÍTULO II:

MARCO CONCEPTUAL

2. CAPÍTULO II: MARCO CONCEPTUAL

En el presente capítulo se definen conceptos y términos teóricos que apoyan el desarrollo de la solución de inteligencia de negocios para el análisis de datos relacionados con el procesamiento de criptomonedas.

Estos conceptos son relevantes sobre el entendimiento de términos como *Blockchain*, criptomonedas, así como de inteligencia de negocios, gestión y visualización de información.

En primer lugar, y con el fin de detallar características de la solución y necesidades de los usuarios finales o interesados de esta, se lleva a cabo un análisis de requerimientos basado en aspectos como los explicados a continuación.

1.10. Análisis de requerimientos

Un proyecto de inteligencia de negocios requiere de un estudio sobre las necesidades de cada interesado con respecto a la información que se requiere procesar. Estas necesidades o requerimientos tienen influencia sobre las decisiones tomadas a lo largo del ciclo de vida de una solución de inteligencia de negocios, sin un correcto análisis de estos se podría cometer errores en la definición del alcance del proyecto, en la escogencia de las herramientas e inclusive en el proceso de modelado y de visualización de la información. (Kimball, Ross, Thorntwaite, Mundy, & Becker, 2006)

1.10.1. Características de un requerimiento

Según Paulish y Berenbach (2009), para que un requerimiento sea considerado realmente una fuente de información funcional y de calidad, debe al menos cumplir con las características que son descritas en la **Tabla 2**.

A continuación, se evidencia la **Tabla 2**.

Tabla 2. Características de los requerimientos

Característica	Descripción
Factible	Un requerimiento es factible si su implementación en la solución es posible dentro de las limitaciones del proyecto.
Válido	Se dice que un requerimiento es válido si y solo si es algo que la solución debe cumplir.
No ambiguo	Un requisito no es ambiguo si solo cuenta con una clara interpretación.
Verificable	Un requisito es verificable si el producto o sistema terminado se puede probar para asegurar que cumple con el requerimiento.
Modificable	Un requisito es modificable si su estructura y estilo son tales que cualquier cambio se pueda hacer de manera fácil, completa y consistente mientras se conserva la estructura y el estilo.
Completo	Es cuando un requerimiento cuenta con todos los datos necesarios y hay suficiente información disponible para poder construirlo en la solución.
Consistente	Un requisito es consistente si no contradice o no entra en conflicto con una norma u otros requisitos de productos o proyectos.
Trazable	Un requisito es trazable si cualquier componente de la solución que implemente dicho requerimiento se puede identificar fácilmente.

Nota. Se detallan las características de los requerimientos. Adaptado de *Software & Systems Requirements Engineering: In Practice*, por D, Paulish y B, Berenbach, 2009.

Después de comprender más sobre el análisis de requerimientos para la satisfacción de las necesidades de los interesados, se continúa definiendo en detalle aspectos relacionados con *blockchain*, esto debido a que los *KPI* que se desea analizar con la solución de inteligencia de negocios están relacionados con sistemas de la empresa, que manejan información afín a esta temática.

1.11. *Blockchain*

Los volúmenes de transacciones están creciendo exponencialmente en todo el mundo y, con ello, se incrementa la complejidad, vulnerabilidad, ineficiencia y costo de los sistemas de transacciones existentes. El comercio se ha basado casi, exclusivamente, en instituciones financieras que actúan como terceros para procesar pagos electrónicos; por su parte, *Blockchain*, fue una propuesta para hacer posible que dos partes dispuestas a realizar transacciones directamente entre sí lo hagan sin la necesidad de este tipo de terceros. (Nakamoto, 2008)

Blockchain es una lista de registros transaccionales firmados criptográficamente y que comparten todos los participantes en una red (Gartner, 2020). Cada registro contiene una marca de tiempo y enlaces de referencia a transacciones anteriores. Con esta información, cualquier persona con derechos de acceso puede rastrear un evento transaccional, en cualquier momento de su historial, que pertenezca a un participante.

La palabra *Blockchain*, que significa “cadena de bloques”, viene de la manera en que se almacenan estos registros transaccionales: en bloques.

Los bloques están vinculados entre sí, formando una cadena. En la **Figura 3**, se muestra la estructura y la manera en la cual los bloques se conectan entre sí por medio de *hashes*, para formar una cadena. Cada bloque contiene su propio *hash*, que funciona como un identificador único, marcas de tiempo de transacciones válidas recientes, y el *hash* del bloque anterior. Este último es el que vincula los bloques entre sí y evita que se altere cualquier bloque o que se inserte un bloque entre dos bloques existentes.

Figura 3. Ejemplo de la estructura de los bloques en Blockchain

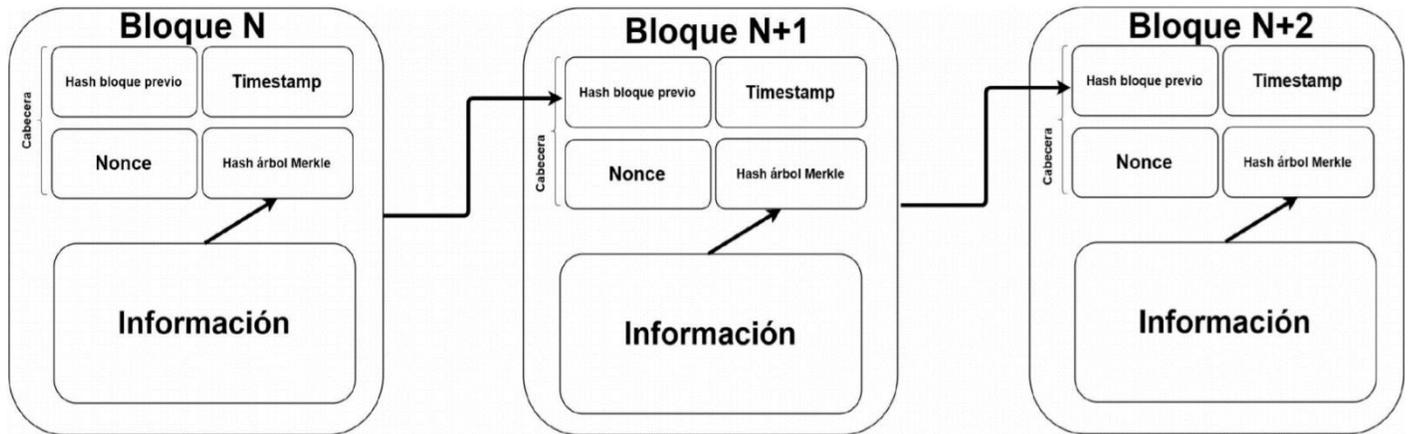


Figura 3. Ejemplo de la estructura de los bloques en Blockchain. Fuente: adaptado de La Blockchain: Fundamentos, por C. Dolader, J. Bel y J. Muñoz, 2017.

La red *Blockchain* es económica y eficiente, ya que elimina la duplicación de esfuerzos y la necesidad de intermediarios; además, las transacciones son seguras, autenticadas y verificables (Gupta, 2017).

Algunos beneficios concretos que *Blockchain* ofrece son:

- **Ahorro de tiempo:** los tiempos de transacción para interacciones complejas de múltiples partes se reducen de días a minutos. La liquidación de la transacción es más rápida, ya que no requiere verificación por parte de una autoridad central.
- **Ahorro de costos:** los intermediarios se reducen porque los participantes pueden intercambiar artículos de valor pagándose, directamente, comisiones bajas; en comparación a las pagadas por las instituciones financieras convencionales.
- **Seguridad:** las características de seguridad de *Blockchain* protegen contra la manipulación, el fraude y el delito cibernético. La privacidad se mantiene a través de técnicas criptográficas y técnicas de partición de datos.
- **Distribución y sostenibilidad:** *Blockchain* es una red distribuida, descentralizada, y se replica selectivamente y casi en tiempo real entre los participantes. Debido a que no es controlada por ninguna organización, la existencia de *Blockchain* no depende de ninguna entidad individual.
- **Transparencia y auditabilidad:** debido a que los participantes en una transacción tienen acceso a los mismos registros, pueden validar las transacciones y verificar identidades o propiedad sin la necesidad de intermediarios externos.

1.11.1. Computación Peer-to-peer

Peer-to-peer se refiere a sistemas que funcionan como un colectivo organizado de nodos, permite que cada uno de estos interactúe directamente con los demás. En la **Figura 4** se muestra la interacción existente entre estos nodos en una red de este tipo.

Figura 4. Representación de una red Peer-to-Peer

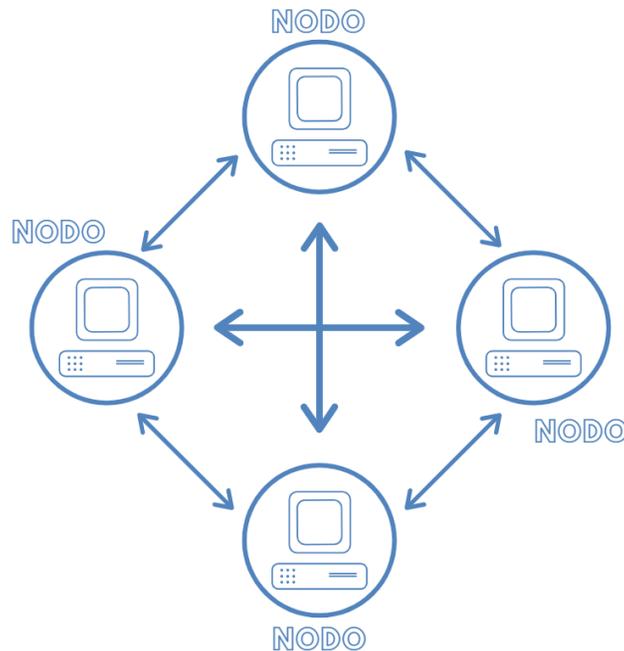


Figura 4. Representación de una red Peer-to-Peer. Fuente: elaboración propia. (2020).

1.11.2. Criptomonedas

A diferencia de las monedas tradicionales, que son emitidas por bancos centrales, las criptomonedas no tienen autoridad monetaria central. Estas no se imprimen como dólares o euros; sino que son "minadas", por ello no dependen de una autoridad monetaria central para monitorear, verificar, aprobar transacciones o administrar el suministro de dinero, sino que están habilitadas para una red informática *peer-to-peer* compuesta por máquinas. (Gupta, 2017)

1.11.2.1. Criptografía

En general, es necesario que exista una forma de controlar el suministro y hacer cumplir varias propiedades de seguridad para evitar ataques en cualquier moneda, sin importar si su naturaleza es

física o digital. En las monedas *fiat*, las organizaciones, como los bancos centrales, controlan la oferta monetaria e intentan evitar, por su cuenta, la falsificación del dinero físico; estas características de seguridad dificultan un atacante o intento de robo, pero no hacen que el dinero sea imposible de falsificar. Las criptomonedas también deben tener medidas de seguridad que eviten que las personas alteren el estado del sistema, para lo cual se implementaron técnicas como la criptografía. (Narayanan, Bonneau, Felten, Miller, & Goldfeder, 2016)

La criptografía proporciona un mecanismo para codificar, de forma segura, las reglas de un sistema de criptomonedas; como sugiere su nombre, las criptomonedas hacen un uso intensivo de la criptografía. Esta es la rama de las matemáticas que permite crear pruebas matemáticas que brinden altos niveles de seguridad.

Uno de los principales métodos criptográficos es el conocido como *hash*. Este es una función matemática que recibe una entrada de tipo *string* de cualquier tamaño y devuelve un resultado de un tamaño determinado. Este *hash* es un proceso eficiente en términos de tiempo y recurso computacional.

1.11.2.2. Direcciones

Una dirección de criptomoneda es un *hash* de una clave pública, similar a una dirección física o un correo electrónico. Sin embargo, una diferencia importante es que cada dirección solo debería usarse para una sola transacción, aunque esto es algo que no siempre sucede.

Para enviar o recibir cantidades de criptomonedas, es necesario compartir o intercambiar las direcciones entre las partes involucradas, estas usualmente se comparten en forma de texto o como un código QR.

Las criptomonedas actualmente soportadas por los sistemas de Progressio Digital son las siguientes:

1.11.2.3. Bitcoin (BTC)

Es la criptomoneda original y actualmente la más grande, por capitalización de mercado. Bitcoin se introdujo en octubre de 2008 con un documento publicado por una persona anónima con el alias de Satoshi Nakamoto. El documento se llama "*Bitcoin: un sistema de efectivo electrónico peer-to-peer*". (Bitcoin, 2020)

1.11.2.4. *Bitcoin cash (BCH)*

Bitcoin Cash es el resultado de una bifurcación o la división de la red bitcoin, que tuvo lugar el 1 de agosto de 2017. En pocas palabras, un grupo de miembros de la comunidad de *Bitcoin* quería aumentar el tamaño del bloque. El cambio resultante fue la división de la red en dos. (Bitcoin Cash, 2020)

1.11.2.5. *Litecoin (LTC)*

También es una moneda digital resultante de una bifurcación de la red de bitcoin, esta fue destinada a permitir pagos instantáneos y más. Litecoin presenta tiempos de confirmación más rápidos y una mayor eficiencia de almacenamiento. (Litecoin, 2020)

1.11.2.6. *Ethereum (ETH)*

Ethereum se lanzó en el 2015 y comparte muchas características del Bitcoin. Se puede enviar al instante, a cualquier persona y en cualquier parte del mundo. Además, es la mayor plataforma de *software* descentralizada abierta, que permite la creación de aplicaciones distribuidas (dApps) y contratos inteligentes. (Ethereum, 2020)

La **Tabla 3**, muestra un resumen comparativo entre algunas de las características de las criptomonedas ofrecidas por Progressio Digital.

Tabla 3. Comparativa entre criptomonedas BTC, BCH, LTC y ETH.

Nombre	Código ISO 4217	Año de lanzamiento	Límite de oferta de monedas	Máximo de decimales sugerido	Tiempo estimado de bloque
Bitcoin	BTC	2009	21 millones	8	10 minutos
Bitcoin Cash	BCH	2017	21 millones	8	10 minutos
Litecoin	LTC	2011	84 millones	8	1.5 minutos

Nombre	Código ISO 4217	Año de lanzamiento	Límite de oferta de monedas	Máximo de decimales sugerido	Tiempo estimado de bloque
Ethereum	ETH	2015	No definido	18	15 segundos

Nota. Fuente: elaboración propia. (2020).

1.11.2.7. Transacciones de criptomonedas

Las transacciones con criptomonedas cuentan con un identificador único que permite rastrearles en la red de *Blockchain*; además, especifica una cantidad de entradas y salidas. Por un lado, una entrada especifica una transacción anterior, por lo que contiene un *hash* de esa transacción y la entrada también contiene el índice de las salidas de la transacción anterior. Por otro lado, las salidas tienen un valor, y la suma de todos los valores de salida tiene que ser menor o igual que la suma de todos los valores de entrada. Si la suma de los valores de salida es menor que la suma de los valores de entrada, la diferencia es una tarifa de transacción que corresponde al minero que publica esta transacción. (Nakamoto, 2008)

1.11.2.7.1. Confirmación de transacciones de criptomonedas

Cuando se dice que una transacción es incluida en su respectivo *Blockchain*, se quiere decir que la transacción alcanzó una cantidad determinada de confirmaciones, considerada suficiente según las reglas de cada criptomoneda. Sin embargo, no existe un número máximo de confirmaciones, solamente se sabe que mientras más confirmaciones tenga una transacción, es más seguro que esta forme parte de una cadena consensuada.

1.11.2.7.2. Comisión de las transacciones de criptomonedas

La minería de criptomonedas es el proceso de resolución de un problema matemático mediante equipos informáticos. Básicamente, lo que hacen estos equipos informáticos es lanzar una serie de posibles soluciones, hasta que la solución planteada coincide con el valor hash del bloque. (Hileman & Michel, 2017)

Una comisión de transacción es la diferencia entre el valor total de las monedas que entran en una transacción menos el valor total de las monedas que salen. Esa comisión va al minero que procesó el bloque que incluye a esta transacción. La idea de una comisión es compensar a los mineros por los costos en los que incurren para procesar una transacción. Es posible definir cuánto pagar por una transacción; en general, si se paga un monto elevado como comisión, entonces es probable que la transacción sea procesada con rapidez.

Luego de haber comprendido sobre la naturaleza de la información que se desea manejar, se definen en detalle conceptos relacionados con la inteligencia de negocios, tema en que se basa la solución planteada en el presente documento.

1.12. Inteligencia de negocios

1.12.1. Datos, información y conocimiento

Antes de introducir la definición de inteligencia de negocios y sus aspectos, es importante conocer la diferencia entre los términos de datos, información y conocimiento, los cuales se describen seguidamente. Según Paniagua y Cadenas (2015):

- **Datos:** son el antecedente o testimonio para llegar al conocimiento exacto de algo o para deducir las consecuencias legítimas de un hecho.
- **Información:** es la representación de los datos en una forma comprensible asignándoles un significado con respecto a un contexto o temática.
- **Conocimiento:** el conocimiento es la información que ha sido comprendida y evaluada con base en la experiencia y el entendimiento intelectual del conocedor de un tema.

1.12.2. Definición de inteligencia de negocios

De acuerdo con Gartner (2020) la inteligencia de negocios (IN) o *Business Intelligence* (BI) es el proceso para diseñar, desarrollar e implementar procesos empresariales e integrar, respaldar y administrar las aplicaciones y plataformas tecnológicas relacionadas. Estas incluyen aplicaciones de negocios, necesidades de análisis e infraestructura de almacenamiento de datos.

Asimismo, una solución de inteligencia de negocios debería cumplir con al menos las siguientes cualidades (Schepps, 2008):

- **Precisión:** la precisión de la información es fundamental. Cuando se incorpora al proceso de toma de decisiones, este conocimiento debe guiar a la gerencia hacia el camino correcto. Si la información no es precisa, es probable que el resultado de estas decisiones no solo falle en ayudar a la organización, sino también es posible que ocasione mayores problemáticas.
- **Valor:** la inteligencia de negocios no solamente debe producir información correcta, sino que esta debe causar un impacto en las organizaciones. Es posible visualizar este impacto en áreas

como la reducción de costos, mejoras en la producción, apoyo a las operaciones diarias, entre otros.

- **Información a tiempo:** el resultado de un proceso de IN debe ser relevante, útil y valioso para un tomador de decisiones, además debe de estar actualizado. La información debe estar disponible cuando es requerida, también, reflejar la realidad buscada en un momento o por un rango de tiempo preciso.
- **Guía sobre conclusiones:** debe haber un curso factible y visible que aproveche la información mostrada. La solución de IN debe ser una base que guie las acciones futuras en la organización; por lo cual, tiene que ser posible pasar de las conclusiones encontradas a la implementación.

Asimismo, existen los siguientes tipos de aplicaciones de inteligencia de negocios:

- **Consultas directas:** las consultas realizadas por usuarios desde aplicaciones de herramientas de consulta de escritorio.
- **Informes estándar:** informes programados regularmente que en general se entregan a través del portal de inteligencia de negocios, y en formato de hoja de cálculo o PDF a un centro compartido de información.
- **Aplicaciones analíticas:** aplicaciones que contienen potentes algoritmos de análisis, además de consultas convencionales sobre bases de datos.
- **Dashboards y cuadros de mando:** interfaces de usuario que muestran indicadores clave de rendimiento (KPI) de forma textual y/o gráfica.
- **Minería de datos y modelos:** análisis exploratorio de conjuntos de datos, generalmente descargados del almacén de datos al *software* de minería de datos. La minería de datos también se utiliza para crear los modelos utilizados por algunas aplicaciones de IN analíticas y operativas.
- **Inteligencia de negocios operativa:** consultas en tiempo real o casi en tiempo real del estado operativo, a menudo acompañadas por interfaces de reescritura de transacciones (Kimball, Ross, Thorntwaite, Mundy, & Becker, 2006).

1.12.3. Enfoques para implementar inteligencia de negocios

Existen diversos enfoques o metodologías de implementación que sirven como guía para el desarrollo de soluciones de inteligencia de negocios. En general, cuando se describen este tipo de enfoques se les suele separar en dos grandes secciones: el enfoque *top-down*, expuesto por el autor Bill Inmon, y el enfoque *bottom-up*, cuyo autor principal es Ralph Kimball.

A continuación, se explican con mayor nivel de detalle ambos enfoques, tomando en consideración que, por la naturaleza de la organización consultada, el enfoque utilizado para la ejecución del presente proyecto es el propuesto por Ralph Kimball.

1.12.4. Enfoque Top-Down (Inmon)

Este enfoque es denominado Top-Down porque en él se propone la construcción de las soluciones de inteligencia de negocios desde el nivel superior dentro de una organización y hacia abajo, especificando cada vez más una misma solución generalizada con base en la cual se desarrolla lo demás.

Dentro de los principales aspectos a destacar de esta metodología para el desarrollo de soluciones de Inteligencia de Negocios, se encuentra la sugerencia del desarrollo de un ambiente entero de bases de datos a nivel organizacional, donde se incluya inclusive las bases de datos operacionales, y a partir del cual los departamentos o las divisiones del negocio pueden construir sus almacenes de datos. (Inmon, 2002)

1.12.5. Enfoque Bottom-Up (Kimball)

Según se explica en Kimball et al. (2006), “Kimball” es un enfoque para la creación de soluciones de inteligencia de negocios, este se orienta al desarrollo incremental o *Bottom-Up*. Asimismo, este tipo de enfoque señala que se debe construir una solución inicial desde abajo, para que esta funcione como una base sólida para la futura construcción de soluciones organizacionales de inteligencia de negocios, cada vez de mayor nivel y complejidad.

En la **Figura 5**, se muestra las actividades que componen el ciclo de vida del desarrollo de soluciones de inteligencia de negocios que el enfoque Kimball propone.

Figura 5. Ciclo de vida del Enfoque Kimball

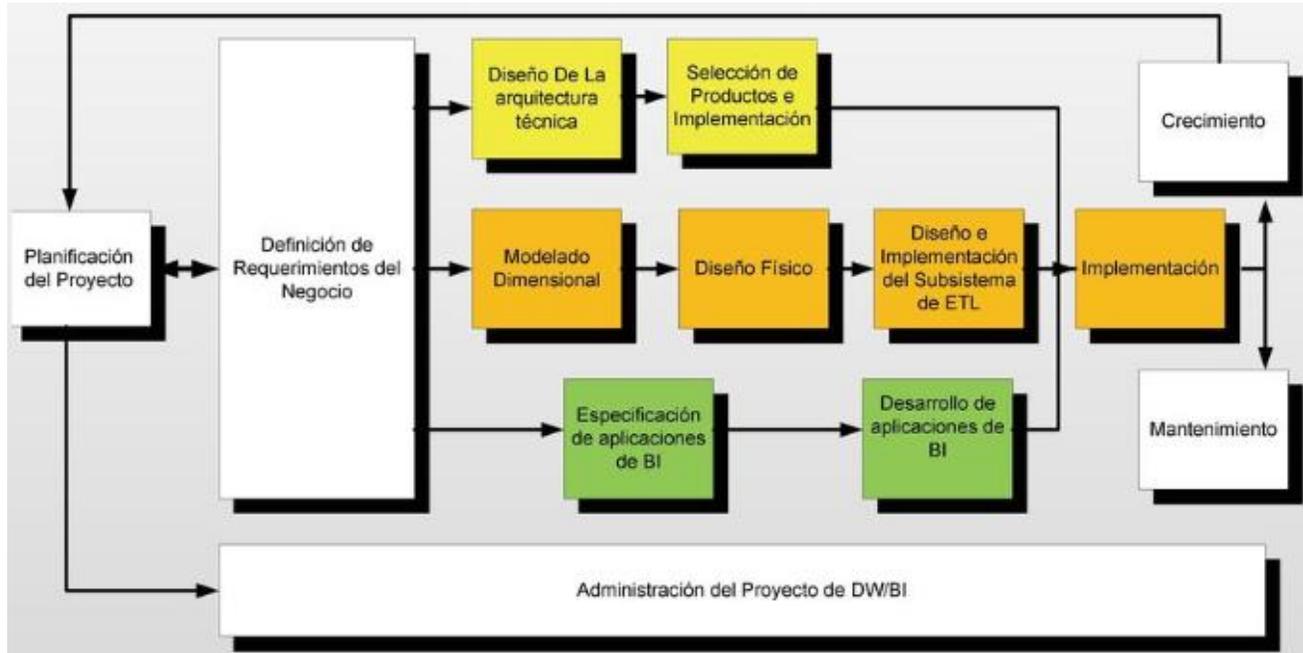


Figura 5. Ciclo de vida del Enfoque Kimball. Fuente: adaptado de The Kimball Group Reader: Relentlessly Practical Tools, por R, Kimball, M, Ross, W, Thorntwaite, T, Mundy y B, Becker, 2006.

Las siguientes son las actividades que componen el ciclo de vida del enfoque Kimball, para el desarrollo de soluciones de IN, estas siguen el orden presente en la Figura 5, en el cual, las que se encuentran de manera paralela en el flujo deben ser ejecutadas simultáneamente.

- **Planificación del Proyecto:** proporciona consejos y pautas para comenzar con la iniciativa para la solución de inteligencia de negocios, desde la evaluación y la preparación, hasta el alcance y la justificación.
- **Definición de requerimientos del negocio:** entendimiento de los usuarios del negocio y sus necesidades. Es necesario identificar los factores que guían al negocio para traducirlos en requerimientos a tomar en consideración en el diseño de la solución. Kimball brinda una guía para recolectar estos requerimientos de manera exitosa.
- **Diseño de la arquitectura técnica:** ambientes de IN necesitan la integración de varias tecnologías. Se debe tomar en cuenta tres factores principales: los requerimientos del negocio, el ambiente técnico actual en la organización y la dirección técnica a nivel estratégico.

- **Selección de productos e implementación:** al emplear el diseño de la arquitectura técnica, se genera una lista de necesidades a nivel de *hardware*, sistemas de gestión de bases de datos, herramientas de *ETL*, herramientas de reportería, entre otros.
- **Modelado dimensional:** se diseña este modelado para apoyar las necesidades analíticas del negocio. Se identifica la granularidad de la tabla de hechos, dimensiones y atributos asociados y hechos numéricos.
- **Diseño físico:** se enfoca en la definición de estructuras físicas, incluyendo configuración del ambiente de bases de datos y seguridad.
- **Diseño e implementación del subsistema ETL:** se basa en diseñar y desarrollar el sistema que soporte el proceso de extracción, transformación y carga de los datos. Es una etapa clave y compleja que se enfrenta en el desarrollo de la solución de IN.
- **Especificación de aplicaciones de BI:** identificación de las aplicaciones candidatas para la solución, así como la definición de interfaces de navegación adecuadas y de acuerdo con las necesidades del usuario final.
- **Desarrollo de aplicaciones de BI:** configuración de metadatos y herramientas de infraestructura de las aplicaciones, luego estas aplicaciones se construyen y validan.
- **Implementación:** una planificación exhaustiva es necesaria para asegurar una buena implementación. Se debe tomar en consideración el trabajo desarrollado en las etapas anteriores para ejecutar la implementación de la solución de IN.
- **Mantenimiento:** incluye tareas como monitoreo, ajustes de desempeño y creación de respaldos para el sistema.
- **Crecimiento:** la finalidad de esta fase es expandir y evolucionar el valor agregado al negocio. Es importante destacar que en esta fase se regresa al principio del ciclo y se reinicia para para construir nuevos avances sobre la base ya establecida.

La fase de planificación del proyecto descrita en el ciclo de vida de Kimball es principalmente abarcada en el CAPÍTULO I: INTRODUCCIÓN. Por otro lado, es relevante aclarar que el alcance del presente proyecto llega hasta la etapa de implementación, con el fin de dejar un producto base que beneficie el desarrollo de futuras soluciones de IN; las actividades de crecimiento y de mantenimiento expuestas en el ciclo se excluyen para su futura aplicación en otros proyectos dentro de la organización.

1.12.6. Elementos tecnológicos de la inteligencia de negocios

La inteligencia de negocios no es una solución basada en una única tecnología, sino que consiste en un conjunto de herramientas y técnicas que consolidan todo el ciclo de vida de la solución (KIMBALL, 2012). En las siguientes secciones se detalla sobre los elementos tecnológicos utilizados en el desarrollo del presente proyecto.

1.12.6.1. Almacén de datos

El propósito completo de una implementación de IN es convertir los datos operativos en conocimiento significativo. Esto significa que IN debe de estar conectado con los datos de una organización para que sea efectivo. La mayoría de las empresas producen datos en una variedad de departamentos diferentes. Los datos dispersos en todas estas áreas probablemente se guardan en una variedad de formatos y *hardware*. Con los datos provenientes de todas las partes del negocio, el desafío es ponerlos todos en un solo lugar y en un formato estándar. Los almacenes de datos o *datawarehouses* son una solución para cumplir con esto.

1.12.6.2. Modelo OLTP

OLTP es una herramienta de procesamiento de transacciones en línea, el enfoque principal de estos sistemas es registrar la actualización, inserción y eliminación de los datos actuales de una transacción. Las consultas *OLTP* son más simples y cortas, por lo tanto requieren menos tiempo en procesamiento y también menos espacio. (Microsoft, 2019)

1.12.6.3. Modelo dimensional

Kimball et al. (2006) especifican que un modelo dimensional busca integrar la información para ser comprendida por los usuarios del negocio y permitir una consulta eficiente de datos.

Los repositorios dimensionales se les conoce como sistemas *OLAP (On-Line Analytical Processing)* o Procesamiento Analítico en Línea y se especializan en el análisis de grandes cantidades de datos.

1.12.6.3.1. Tablas de hechos

En un modelo dimensional, la tabla de hechos es una tabla central con medidas numéricas de desempeños, estas se caracterizan por contar con una llave foránea hacia las tablas dimensionales. (Kimball et al., 2006). Es decir, un hecho consiste en métricas que son analizadas al utilizar una o más dimensiones definidas. En este caso, la granularidad en una tabla de hechos se refiere al nivel de detalle en el que se encuentra definida la información.

1.12.6.3.2. Tablas de dimensiones

Una tabla de dimensión se compone de una serie de columnas o atributos que permiten analizar las métricas de los hechos a diferente nivel detalle, por lo que estos atributos pueden ser asociados a filtros. Estas tablas permiten describir el qué, quién, cómo y por qué, asociado a un evento. (Kimball et al., 2006)

1.13. Proceso de extracción, transformación y carga de datos (*ETL*)

Este proceso cuenta con tres principales fases, (Talend , 2020):

- **Extracción de datos:** se extraen los datos desde las diferentes fuentes en la organización, estas fuentes de información pueden ser generados por las aplicaciones operativas o bien provenir de una fuente externa de información, que permita enriquecer y obtener la información solicitada.
- **Transformación de los datos:** llevar a cabo cualquier cambio necesario individualmente en los datos, de manera que sea posible crear información limpia, estandarizada y correcta. Este proceso suele verificar y manejar la existencia de datos duplicados y de valores nulos, así como la necesidad de llevar a cabo conversiones, seleccionar columnas específicas, agregar datos, entre otras actividades.
- **Carga de datos:** carga de esos datos extraídos y transformados a su nuevo destino, ya sea en tablas de hechos, dimensionales o de trabajo.

La **Figura 6**, muestra el flujo que toman los datos para llegar finalmente a un almacén. En este proceso, los datos provienen de diferentes fuentes de datos y pasan por un proceso *ETL* inicial, para ser cargados en un área de trabajo o *Staging Area*, a partir de este paso, los datos pasan nuevamente a través de otro proceso *ETL*, que permite la depuración de la información y la carga de manera dimensional en el almacén de datos.

Figura 6. Almacén de datos con proceso ETL

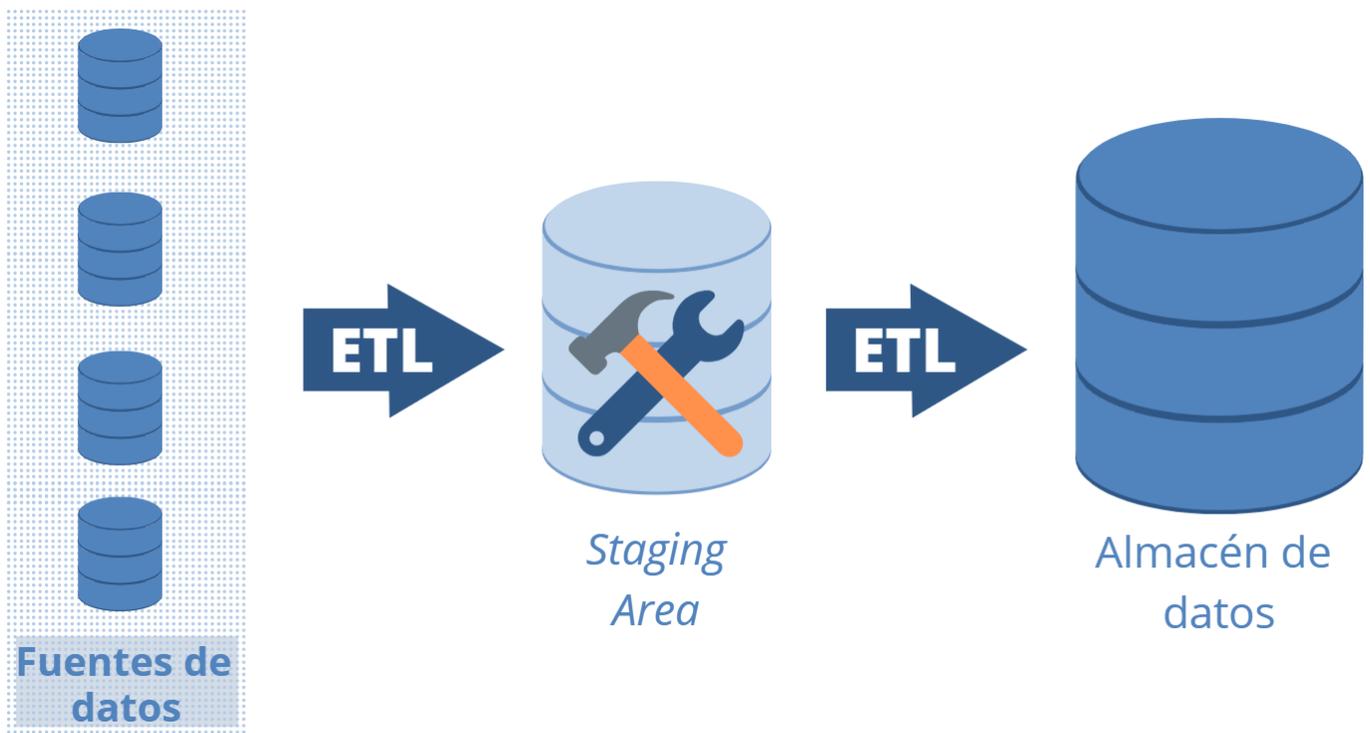


Figura 6. Almacén de datos con proceso ETL. Fuente: elaboración propia. (2020).

CAPÍTULO III:

MARCO METODOLÓGICO

3. CAPÍTULO III:

MARCO METODOLÓGICO

El presente capítulo describe el marco metodológico del Trabajo Final de Graduación, en este se definen aspectos como el tipo de investigación, el diseño, la población de estudio, así como las fuentes, técnicas e instrumentos usados para la recopilación y análisis de información requerida.

Con el marco metodológico se pretende detallar los procedimientos ejecutados para demostrar la validez y autenticidad de la investigación; de igual forma, se determina la manera en la cual se obtiene la información necesaria. (Ulate & Vargas, 2016).

1.14. Tipo de investigación

Las investigaciones son llevadas a cabo, a través de dos enfoques: el enfoque cualitativo y el cuantitativo. Cada uno de estos es utilizado de manera distinta, pero su fin es el mismo: generar conocimiento, a través de procesos cuidadosos, metódicos y empíricos. (Hernández, Fernández, & Baptista, 2014)

Por un lado, el enfoque de investigación de tipo cualitativo se centra en la recopilación y análisis de datos para comprender los fenómenos en un ambiente natural, explorándolos desde la perspectiva de los participantes y en relación con su contexto (Hernández, Fernández, & Baptista, 2014). Por otro lado, el método cuantitativo utiliza la recolección de datos para probar hipótesis y teorías con base en la medición numérica, y con ello, ejecutar análisis estadísticos que permita el establecimiento de patrones de comportamiento. (Cresswell, 2014)

De acuerdo con lo anterior, se determinó que el enfoque empleado para el desarrollo del presente proyecto es de tipo cualitativo. Debido a que la recolección de información no abarca mediciones numéricas ni análisis estadísticos, sino que es obtenida a través de métodos como la observación, entrevistas y revisión documental; de manera que se desarrolle una solución de IN que brinde valor agregado a las personas interesadas en ella.

1.15. Alcance de la investigación

Una vez efectuada la revisión de literatura y un planteamiento del problema de la presente investigación, lo que prosigue es visualizar su alcance, se toma en cuenta las restricciones que posee la información de cada organización y el uso que se necesite dar. (Hernández, Fernández, & Baptista, 2014)

La problemática de la investigación es fundamental para definir el alcance, debido a que fomenta la base de las respuestas buscadas para solventar la situación actual. Debido a esto existen tipos de alcance de investigación, dentro de los cuales es posible clasificar el presente proyecto, estos se describen en la **Tabla 4**.

Tabla 4. Alcances de una investigación

Alcance	Descripción
Explicativo	Está dirigida a responder por las causas de eventos y fenómenos físicos o sociales.
Descriptivo	Busca especificar las propiedades, características y perfiles de personas, grupos, comunidades, procesos, objetos y cualquier otro fenómeno.
Exploratorio	Se realizan cuando el objetivo es examinar un tema o problema de investigación poco estudiado.
Correlacional	Tiene como objetivo comprender la relación o nivel de asociación que existe entre dos o más conceptos, categorías o variables en un determinado contexto.

Nota. Fuente: adaptado de *Metodología de la Investigación*, por R, Hernández, C, Fernández y P, Baptista, 2014.

Para el cumplimiento de lo establecido en los objetivos del presente proyecto, se considera trabajar bajo el enfoque **descriptivo**, debido a que este desarrolla el análisis de requerimientos y necesidades de los interesados con respecto a su perfil y entorno.

Al llevar a cabo la investigación a través del enfoque descriptivo, es posible identificar en el proyecto aspectos relevantes con respecto a los principales involucrados y a las herramientas de recolección de la información, necesarios para el desarrollo de la solución de inteligencia de negocios que brinde visibilidad a los sistemas de la organización.

1.16. Diseño de la investigación

Una vez establecido que el enfoque de la presente investigación es el cualitativo, se debe definir el diseño o abordaje de la investigación. Según Hernández et al. (2014) existen diferentes diseños genéricos para este tipo de enfoque, se detallan en la **Tabla 5**:

Tabla 5. Descripción de los diseños de investigación

Diseño	Descripción
Teoría fundamentada	Es de utilidad cuando existen preguntas sobre procesos y relaciones entre conceptos que conforman un fenómeno y no se dispone de teorías previas o no aplican en el contexto. Se genera información con respecto al fenómeno para concluir en una teoría que explica el problema de la investigación.
Fenomenológico	Pretende analizar preguntas sobre lo que una variedad de personas experimenta en común con respecto a un fenómeno o proceso, de manera que tras acabar la investigación se documenten las experiencias comunes y distintas, así como las categorías que se indican frecuentemente en las experiencias.

Diseño	Descripción
Etnográfico	Pretende investigar las características, estructura y funcionamiento de un sistema social. Genera información que permite describir y explicar los elementos y categorías que integran el sistema social.
Narrativo	Como resultado de este tipo de diseño se obtienen historias sobre procesos, hechos, eventos y experiencias mediante una línea de tiempo. Está relacionado con preguntas de investigación orientadas a comprender una sucesión de eventos, a través de las historias de quienes la vivieron.
Investigación-acción	Se utiliza cuando una problemática de una comunidad necesita resolverse y se pretende lograr el cambio. Como parte de la investigación se realiza un diagnóstico de la problemática para generar información sobre las causas, consecuencias y sus soluciones.

Nota. Fuente: adaptado de *Metodología de la Investigación*, por R, Hernández, C, Fernández y P, Baptista, 2014.

Con base en lo expuesto, se determinó que el tipo de diseño adecuado para el proyecto es el de **investigación-acción**. Algunas de las principales características que relacionan este enfoque al presente trabajo son las expuestas en los siguientes postulados:

- Es participativa, las personas trabajan con la intención de mejorar sus propias prácticas.
- Es un proceso sistemático de aprendizaje, orientado a la práctica.
- Implica registrar, recopilar, analizar juicios propios, reacciones e impresiones en torno a lo que ocurre.
- Procede progresivamente a cambios más amplios.
- Empieza con pequeños ciclos de planificación, acción, observación y reflexión y avanza hacia problemas de más importancia.
- Es colaborativa, se lleva a cabo en grupo por las personas implicadas, estas participan y colaboran en todas las fases del proceso de investigación. (López, 2012)

1.17. Fuentes de información

Para llevar a cabo este trabajo fue necesaria la consulta a diversas fuentes de información, es posible clasificar estas fuentes de información en dos grupos: fuentes de información primarias y fuentes de información secundarias (Ulate & Vargas, 2016). A continuación, se muestra con detalle las principales fuentes de información consultadas de acuerdo con su respectiva clasificación.

1.17.1. Fuentes de información primarias

Las fuentes primarias son aquellas que brindan información directamente de quién la produjo; es decir que proviene directamente de un autor y no es interpretada por otra persona; algunos ejemplos de fuentes primarias son artículos, libros, páginas de internet, entre otros. Para el presente proyecto fueron utilizadas las siguientes fuentes de información:

- The *Data Warehouse Toolkit*. (Kimball, Ross, Thorntwaite, Mundy, & Becker, 2006).
- Libros sobre el desarrollo de metodologías de investigación.
- Publicaciones académicas relacionadas con el tema de inteligencia de negocios.
- Proyectos similares desarrollados dentro del Tecnológico de Costa Rica.

1.17.2. Fuentes de información secundarias

Estas fuentes de información colaboran con el investigador, al brindarle la capacidad de obtener información a través de un tercero; es posible encontrar este tipo de fuentes en comentarios sobre obras originales de un autor, libros que cuentan con su propia recopilación de datos, entre otros. Algunas de las fuentes de información secundarias consultadas para el proyecto, son las siguientes:

- Material del curso TI-6900 Inteligencia de Negocios.
- Libros, foros y revistas sobre inteligencia de negocios.
- Páginas *web*.
- Demás recursos digitales de la Biblioteca del Instituto Tecnológico de Costa Rica, como proyectos de graduación de semestres anteriores.

1.18. Sujetos de información

El proceso de investigación cualitativo del presente proyecto requiere de la definición de los sujetos de investigación, los cuales representan los interesados o personas que aportan la mayor parte de información para la comprensión del problema de estudio y el planteamiento de la solución de inteligencia de negocios adecuada.

En la **Tabla 6** se describe la importancia de cada uno de los sujetos de información del proyecto, de acuerdo con su rol en la organización.

Tabla 6. *Sujetos de información*

Rol del sujeto de investigación	Importancia
Gerente general	<ul style="list-style-type: none">• Es el encargado de tomar las decisiones gerenciales de la organización.• Requiere tener a mano información clara y real sobre el desempeño y los niveles de uso de los sistemas de información de la empresa.• Es el principal contacto con los clientes, por lo cual conoce sus necesidades y quejas.• Cuenta con una visión completa del negocio.
Administrador de proyectos	<ul style="list-style-type: none">• Lidera al equipo para el desarrollo y mantenimiento de soluciones tecnológicas que satisfagan los requerimientos del cliente, por lo que cuenta con conocimiento amplio sobre los sistemas de información.• Conoce la necesidad de mejorar la visibilidad sobre el desempeño y el uso de los sistemas de la organización.• Cuenta con estudios y un amplio conocimiento sobre el tema de inteligencia de negocios.
Desarrolladores de software	<ul style="list-style-type: none">• Encargados de diseñar, desarrollar y mantener los sistemas.

Rol del sujeto de investigación	Importancia
	<ul style="list-style-type: none"> • Cuentan con conocimiento sobre las tecnologías y los procesos de la organización. • Brindan soporte y solucionan los conflictos que se presenten en los ambientes de producción.
Ingenieros <i>Devops</i> (equipo tercerizado)	<ul style="list-style-type: none"> • Encargados de diseñar, configurar y mantener la infraestructura que soporta los sistemas de la organización. • Establecen las medidas de seguridad para los sistemas de información, incluyendo los servidores y las bases de datos.
Analista de negocio (representante del cliente)	<ul style="list-style-type: none"> • Tiene opinión sobre el desempeño, nuevos requerimientos y disconformidades con respecto a los sistemas que consume su organización para el procesamiento de criptomonedas. • Cuenta con interés de visualizar el comportamiento de los sistemas con respecto al manejo del dinero de su organización. • Ha llevado a cabo por su propia cuenta estudios sobre el desempeño de los sistemas de Progressio Digital y de las transacciones en el <i>blockchain</i>. • Tiene la capacidad de controlar el volumen de dinero o transacciones que es procesado a través de los sistemas de información de Progressio Digital.

Nota. Fuente: elaboración propia. (2020).

1.19. Instrumentos de recolección de datos

En esta sección se detallan los instrumentos utilizados para la recolección de la información necesaria para la elaboración del proyecto de investigación. Esta busca obtener datos provenientes de personas, comunidades o situaciones, con la finalidad de analizarlos y comprenderlos para generar conocimiento. (Hernández, Fernández, & Baptista, 2014).

A continuación, se presentan las técnicas e instrumentos utilizadas en el presente proyecto.

1.19.1. Entrevista

Esta técnica le facilita al investigador la obtención de conocimiento, de primera mano, por parte de un interesado que cuente con información relevante. Según Hernández et al. (2014), es posible definir las entrevistas como una reunión para conversar e intercambiar información entre una persona (el entrevistador) y otra (el entrevistado).

Según los autores, la entrevista cuenta con tres clasificaciones:

- Estructurada: se prepara un cuestionario, este guía al entrevistado por medio de un conjunto de preguntas. (Ulate & Vargas, 2016)
- Semiestructurada: este tipo de entrevista cuenta con una base de preguntas o temas, sin embargo, el entrevistador puede realizar preguntas adicionales para detallar ciertos temas o conseguir más información. (Hernández, Fernández, & Baptista, 2014)
- No estructurada o abierta: se basa en una guía general de contenido y el entrevistador posee toda la flexibilidad sobre cómo manejarla. (Ulate & Vargas, 2016)

Para el presente proyecto, fueron utilizadas como técnicas de recolección la entrevista semiestructurada y la entrevista no estructurada o abierta. Para algunos casos, resultó necesaria la recolección de información específica, esta podía derivar en otros temas relevantes para el proyecto, por lo cual se optó por una entrevista semiestructurada; por su parte, el instrumento para aplicar este tipo de entrevista se encuentra en el Apéndice III. Plantilla para entrevistas semiestructuradas.

Por otro lado, en ciertas ocasiones era necesario conocer, de manera abierta, la visión de algunos involucrados, con respecto a la solución de la inteligencia de negocios, por lo que se llevó a cabo una entrevista no estructurada, ver Apéndice IV. Entrevista abierta.

1.19.2. Revisión documental

Existen documentos que representan información relevante y valiosa para llevar a cabo investigaciones con el enfoque cualitativo, estas fuentes de conocimiento permiten al investigador conocer y comprender antecedentes y situaciones que determinan la situación actual. (Hernández, Fernández, & Baptista, 2014)

Para el presente documento, fue necesaria una revisión documental sobre las bases de datos transaccionales de los sistemas de Progressio Digital, con la finalidad de comprender los datos que son almacenados, tanto a nivel de estructura como de composición. Por otro lado, también fue necesario revisar documentación de los *API* de los microservicios ofrecidos por Progressio Digital, con el objetivo de conocer el origen y la finalidad de la información almacenada en cada base de datos. Ver Apéndice V. Revisión documental.

1.19.3. Observación

Ulate y Vargas (2016) mencionan que la observación se utiliza con el objetivo de obtener información de la realidad, mediante la percepción intencionada y selectiva de un objeto o fenómeno determinado. Asimismo, según Hernández et al. (2014) existen cinco principales niveles de participación en un proceso de observación, estos son:

- No participación: el observador no está presente, por ejemplo, cuando se lleva a cabo una observación sobre videos.
- Participación pasiva: el observador está presente, pero no interactúa.
- Participación moderada: el observador participa en algunas actividades, pero no en todas.
- Participación: el observador participa activamente en la mayor parte de las actividades, pero no se mezcla completamente con los participantes.
- Participación completa: el observador es uno de los participantes principales e interactúa de manera completa en la actividad observada.

Con respecto al presente proyecto, el nivel de participación fue moderado, debido a que el observador interactúa, de manera directa, con los encargados de ejecutar las actividades observadas, principalmente referentes a la calidad y gestión de datos. Ver Apéndice VI. Observación.

1.19.4. Tabla comparativa de herramientas

Como parte del desarrollo de la solución de IN, fue necesaria la escogencia de herramientas de *software* para tareas como: el diseño de los procesos ETL y la visualización de los tableros y reportes. Para definir cuál herramienta utilizar se realiza una comparación bajo el método denominado Modelo de Puntuación Ponderada, este se emplea para contrastar herramientas de *software*. (Business Analyst Learnings website, 2014)

En este método se deben elegir los requerimientos que se consideren indispensables para la futura herramienta. Para cada requerimiento se asigna un “peso” que simboliza la importancia que tiene dicho requerimiento en el *software*. Este peso es un porcentaje de 1% a 100% (entre todos los requerimientos debe sumar 100%). Una vez asignados los pesos se define una puntuación para cada requerimiento, basado en su funcionamiento, en cada una de las herramientas (Business Analyst Learnings website, 2014)

El formato con las tablas que se utilizaron para realizar dicha comparación se encuentra en el Apéndice VII. Tabla comparativa de herramientas.

1.19.5. Grupos focales

Un grupo focal se define como una técnica de recolección de datos mediante una entrevista grupal no estructurada, la cual gira alrededor de una temática propuesta por el investigador. Un punto importante de esta técnica es que permite que surjan reacciones, opiniones y pensamientos sobre un tema en específico. En el Apéndice VIII. Grupo Focal se encuentra el material generado para llevar a cabo el grupo focal. (Bonilla-Jimenez & Escobar, 2011).

1.20. Variables de la investigación

A continuación, se muestran las variables de investigación asociadas con los objetivos específicos del proyecto. Las variables surgen de los objetivos específicos, además, una de las principales razones para elaborar el cuadro de variables es que permite descubrir si existen objetivos duplicados, mal planteados, ambiguos o de compleja realización. (Ulate & Vargas, 2016)

En la **Tabla 7** se presenta un conjunto de variables, definidas a partir de los objetivos específicos planteados para este proyecto, una serie de indicadores para determinar su cumplimiento y los instrumentos utilizados para llevar a cabo su ejecución.

Tabla 7. Cuadro de Variables

Objetivo	Variables	Concepto	Indicador	Instrumentos
Analizar los sistemas de información de la empresa, sus procesos y sus correspondientes bases de datos para el planteamiento de la situación actual que oriente el desarrollo de la solución de inteligencia de negocios.	Análisis de los sistemas de información de la empresa.	Especificación de los sistemas de información de la empresa, incluyendo modelos de sus respectivas bases de datos relacionales, y las principales funciones de cada uno.	<ul style="list-style-type: none"> Cantidad de sistemas que se encuentran operando por unidad. Cantidad de espacio de almacenamiento consumido por sistema en las bases de datos. 	<ul style="list-style-type: none"> Entrevistas no estructuradas Revisión Documental Observación moderada

Objetivo	Variables	Concepto	Indicador	Instrumentos
<p>Proponer indicadores clave de desempeño de los sistemas de la organización para que estos sirvan como apoyo en procesos de control, monitoreo y toma de decisiones en la organización.</p>	<p>Indicadores clave de desempeño.</p>	<p>Definición de instrumentos para la medición cuantitativa sobre el desempeño de los sistemas de información de la empresa.</p>	<ul style="list-style-type: none"> • Número de indicadores de desempeño propuestos por sistema. • Cantidad de indicadores clave de desempeño por nivel de prioridad. 	<ul style="list-style-type: none"> • Entrevistas semiestructuradas • Revisión Documental
<p>Diseñar el modelo del <i>datamart</i>, así como los procesos de extracción, transformación y carga de datos (ETL) para soportar la carga,</p>	<p>Modelo del <i>Datamart</i></p>	<p>Modelo que define la organización de las bases de datos correspondientes a un <i>Datamart</i>.</p>	<ul style="list-style-type: none"> • Cantidad de tablas dimensionales del modelo. • Número de tablas de hechos definidas para el modelo. • Cantidad de modelos estrella definidos a 	<ul style="list-style-type: none"> • Revisión documental. • Grupos focales.

Objetivo	Variables	Concepto	Indicador	Instrumentos
estandarización y actualización de la información proveniente de los sistemas de la empresa			partir de la relación entre las tablas de hechos y las tablas dimensionales.	
	Procesos <i>ETL</i>	Proceso mediante el cual se traspasa la información de un modelo relacional a un modelo dimensional, encargándose de las actividades requeridas de extracción, transformación y carga.	<ul style="list-style-type: none"> • Número de procesos de ejecución para las tablas del modelo dimensional. • Cantidad de ejecuciones de los procesos por semana. • Cantidad de días entre cada actualización de la información cambiante. 	<ul style="list-style-type: none"> • Tabla comparativa de herramientas. • Revisión de documental. • Grupos focales.

Objetivo	Variables	Concepto	Indicador	Instrumentos
<p>Elaborar <i>dashboards</i> con vistas gráficas de indicadores de desempeño para el favorecimiento del proceso de análisis de la información y la toma de decisiones gerenciales con respecto al procesamiento de criptomonedas.</p>	<p><i>Dashboards</i> para visualización de <i>KPI</i></p>	<p>Herramienta de reportería que permite la representación gráfica de indicadores clave de desempeño.</p>	<ul style="list-style-type: none"> Intervalos de tiempo para el refrescamiento de datos. Cantidad de indicadores de desempeño graficados en el tablero. Cantidad de filtros interactivos por tablero construido. 	<ul style="list-style-type: none"> Tabla comparativa de herramientas. Grupos focales.

Nota. Fuente: elaboración propia. (2020).

1.21. Procedimiento metodológico

Para realizar el presente proyecto, fue necesario definir un procedimiento metodológico en donde se especificaron las fases realizadas de trabajo. Estas fases del procedimiento metodológico con base en el cual se llevó a cabo el desarrollo de la solución de IN son las mostradas en la **Figura 7**.

Figura 7. Procedimiento metodológico

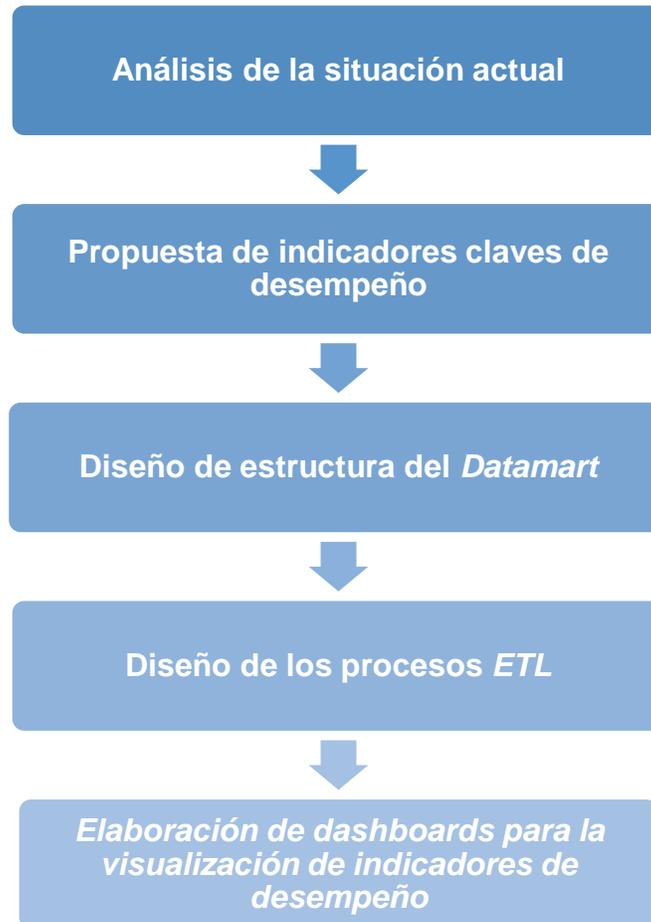


Figura 7. Procedimiento metodológico. Fuente: elaboración propia. (2020).

En las próximas secciones se detalla cada una de las fases establecidos como parte del procedimiento metodológico.

1.21.1. Análisis de la situación actual

En esta etapa se realizó un estudio de la situación actual de la compañía, de forma que, se identificaron las bases para la ejecución y desarrollo de la solución de IN.

Para llevar a cabo este análisis fueron realizadas revisiones documentales de las bases de datos OLTP existentes en la organización, así como de la documentación de los servicios API ofrecidos por la empresa. Por otro lado, también se efectuaron entrevistas no estructuradas a los desarrolladores de los sistemas, y a los interesados del proyecto, con el fin de recolectar información específica sobre los procesos y sus demás características.

1.21.2. Propuesta de indicadores claves de desempeño

Una vez recolectada la información con respecto a los sistemas de la organización, se procedió a realizar una propuesta de indicadores clave de desempeño (*KPI*) para estos sistemas.

Para llevar a cabo la definición de esta lista de indicadores se trabajó por medio de entrevistas semiestructuradas, estas fueron realizadas a la gerencia de la organización y a representantes del cliente; de igual forma, se llevó a cabo una revisión documental de algunas consultas SQL que son utilizadas en la organización, con el fin de llevar un control básico de la información. Finalmente, se realizaron grupos focales con varios interesados, para exponer ideas sobre estos *KPI*.

1.21.3. Diseño de estructura del *datamart*

Luego de haber especificado la situación actual de los sistemas de la empresa e identificado los indicadores clave de desempeño para estos sistemas, se llevó a cabo el modelado o diseño de la estructura del almacén de datos. Para esto, se ejecutaron revisiones documentales sobre la situación de las bases de datos actuales de la empresa, así como grupos focales para conversar sobre las soluciones planteadas.

Para ello, se concretó la arquitectura necesaria para soportar estas bases de datos, tanto para el área de trabajo, como para el almacén de datos como tal. Luego, se identificó y diseñó el modelo dimensional, se definieron las tablas de hechos y las tablas dimensionales, para soportar los modelos estrellas necesarios para la solución.

1.21.4. Diseño de procesos *ETL*

Con la estructura del *datamart* diseñada, se procedió a trabajar en el desarrollo de los procesos *ETL* para el manejo del flujo de los datos hasta llegar a este almacén.

Primero, fue necesaria la identificación de la herramienta con la que se desarrollaron estos procesos, para ello se tomó en cuenta los requerimientos de la organización, asimismo, se llevó a cabo una comparativa de herramientas existentes en el mercado, esto por medio de una tabla donde se pesó cada una de las herramientas con respecto a cada requerimiento de la organización. Para la escogencia de cuáles herramientas pesar, fueron presentadas algunas propuestas, por medio de un grupo focal, al gerente de la organización y al administrador de proyectos.

Luego de haber seleccionado la herramienta, se procedió a llevar a cabo el desarrollo del flujo de extracción, transformación y carga para las tablas de hechos y las tablas dimensionales en el *datamart*. Una vez concluido este desarrollo, se implementó la arquitectura en el ambiente de producción, donde se ejecutarían los procesos, sus versiones, así como la periodicidad y el manejo de la carga incremental de datos.

Por último, se validó el funcionamiento de los procesos por medio de pruebas, para definir la calidad de la información extraída en estos.

1.21.5. Elaboración de *dashboards* para la visualización de indicadores de desempeño

Una vez cargada la información dentro del *datamart* por medio de los procesos *ETL*, se procedió a llevar a cabo el desarrollo de los *dashboards*. Para ello, primero fue definida una herramienta de visualización de datos, se empleó una tabla comparativa de herramientas y un grupo focal para la escogencia de estas, tal como en la fase anterior.

Ya definida la herramienta, se procedió a realizar la conexión con el repositorio del *datamart*, a partir de ahí se comenzó a crear los *dashboards* asociados con los indicadores claves de desempeño, definidos anteriormente, asimismo, para cada uno de los *dashboards* se concretaron filtros de información, se configuró la periodicidad de actualización de la información y del envío de estos *dashboards* a los respectivos interesados por medio de correo electrónico.

Una vez debidamente identificadas y definidas las fases que conforman el procedimiento metodológico del presente proyecto, es importante aclarar la relación entre estas y el ciclo de vida para soluciones de IN propuesto en el Enfoque *Bottom-Up* (Kimball).

1.21.6. Relación entre el procedimiento metodológico y ciclo de vida de Kimball

Si bien el presente proyecto está basado en el Enfoque *Bottom-Up* (Kimball), las fases del procedimiento metodológico implementado no son idénticas a las expuestas por este dentro, del ciclo de vida para el desarrollo de soluciones de inteligencia de negocios que se muestra en la **Figura 5**.

Según Kimball (2006), el ciclo de vida propuesto debe comprenderse como un marco de referencia y debe utilizarse como una base para el desarrollo de proyectos de IN; lo anterior quiere decir que no debe ser visto como una guía absoluta a seguir, sino que se debe adaptar a las necesidades de cada proyecto y de su respectiva organización.

Algunas de las actividades llevadas a cabo para cada una de las fases del proceso metodológico se ejecutan de manera paralela entre sí, tal y como sucede en la metodología propuesta por Kimball. En la **Tabla 8** se muestra, marcando con una “X”, la relación entre las fases del proceso metodológico propuesto para el presente proyecto y las fases correspondientes al ciclo de vida de soluciones de IN expuesto por Kimball (2006).

Tabla 8. Relación entre el proceso metodológico y el ciclo de vida de Kimball.

Fases del ciclo de vida propuesto por Kimball	Fases del proceso metodológico				
	Análisis de la situación actual	Propuesta de indicadores claves de desempeño	Diseño de estructura del <i>datamart</i>	Diseño de procesos ETL	Elaboración de <i>dashboards</i> para la visualización de indicadores de desempeño
Definición de requerimientos de negocio	X	X	X	X	X
Diseño de la arquitectura técnica			X	X	X
Selección de productos e implementación				X	X
Modelado dimensional			X		
Diseño físico			X		
Diseño e implementación del subsistema ETL				X	
Implementación			X	X	X
Especificación de aplicaciones de BI					X
Desarrollo de aplicaciones de BI					X

Nota. Se detalla la relación entre el proceso metodológico y el ciclo de vida de Kimball. Fuente: elaboración propia. (2020).

1.22. Matriz metodológica

Este apartado presenta la matriz metodológica del proyecto, la cual busca enlazar los objetivos del proyecto con las fases definidas, sus respectivas actividades, así como las técnicas/herramientas de recolección y procesamiento de información, para el desarrollo de los entregables. Dicha matriz se muestra en la **Tabla 9**:

Tabla 9. Matriz metodológica del proyecto

Objetivo	Fase	Actividad	Instrumento / Herramienta
<p>Analizar los sistemas de información de la empresa y sus correspondientes bases de datos para el planteamiento de la situación actual que oriente el desarrollo de la solución de inteligencia de negocios.</p>	<p>Análisis de la situación actual</p>	<ul style="list-style-type: none"> • Análisis de la documentación de los sistemas existentes. • Identificación de los procesos relacionados a cada sistema. • Análisis de las bases de datos de la empresa. 	<ul style="list-style-type: none"> • Entrevista no estructurada • Revisión Documental • Observación moderada

Objetivo	Fase	Actividad	Instrumento / Herramienta
<p>Proponer indicadores clave de desempeño sobre los sistemas de la organización para que estos sirvan como apoyo en procesos de control, monitoreo y toma de decisiones en la organización.</p>	<p>Propuesta de indicadores claves de desempeño</p>	<ul style="list-style-type: none"> • Análisis de las actuales consultas a base de datos utilizadas para monitoreo. • Listado de los indicadores clave de desempeño. 	<ul style="list-style-type: none"> • Entrevista semiestructurada • Revisión Documental • Grupos focales
<p>Diseñar el <i>datamart</i>, los procesos de extracción, transformación y carga de datos (ETL); así como la arquitectura necesaria para soportar la carga, estandarización y actualización de la</p>	<p>Diseño de estructura del <i>Datamart</i></p>	<ul style="list-style-type: none"> • Definición de fuentes de información necesarias. • Diseño de tablas de hecho. • Diseño de tablas dimensionales. • Modelado de la estructura del área de trabajo y del <i>Datamart</i>. 	<ul style="list-style-type: none"> • Revisión documental. • Grupos focales.
		<ul style="list-style-type: none"> • Definición de la herramienta utilizada para manejo de los procesos ETL. 	<ul style="list-style-type: none"> • Tabla comparativa de herramientas.

Objetivo	Fase	Actividad	Instrumento / Herramienta
<p>información proveniente de los sistemas de la empresa sobre el procesamiento de criptomonedas.</p>	<p>Diseño de procesos <i>ETL</i></p>	<ul style="list-style-type: none"> • Diseño del flujo de los procesos ETL. • Definición de la arquitectura del ambiente de producción para la ejecución de los procesos ETL. • Manejo de versiones de los procesos ETL en el ambiente de producción. • Definición de la periodicidad y del manejo de la carga incremental de información en el ambiente de producción. • Validación y pruebas de las cualidades de la información extraída con los procesos ETL. 	<ul style="list-style-type: none"> • Revisión de documental. • Grupos focales.
<p>Elaborar reportes y tableros de indicadores para el favorecimiento del proceso de análisis de la información y la toma de decisiones gerenciales con</p>	<p>Elaboración de <i>dashboards</i> para la visualización de indicadores de desempeño.</p>	<ul style="list-style-type: none"> • Determinar la herramienta para el manejo de las visualizaciones. • Llevar a cabo la conexión entre la herramienta escogida y el repositorio del <i>Datamart</i>. • Selección y creación de los <i>dashboards</i> asociados con los indicadores clave de desempeño. • Definición de filtros para los gráficos. 	<ul style="list-style-type: none"> • Tabla comparativa de herramientas. • Revisión documental. • Grupos focales.

Objetivo	Fase	Actividad	Instrumento / Herramienta
respecto al procesamiento de criptomonedas.		<ul style="list-style-type: none">• Configuración de la periodicidad de la actualización de la información para cada <i>dashboard</i>.• Configuración de la periodicidad de envío de los <i>dashboards</i> a los usuarios interesados.	<ul style="list-style-type: none">• Entrevista no estructurada.

Nota. Se detalla la matriz de la metodología a emplear en el presente proyecto de investigación. Fuente: elaboración propia. (2020).

CAPÍTULO IV: ANÁLISIS DE RESULTADOS

4. CAPÍTULO IV: ANÁLISIS DE RESULTADOS

En este capítulo se lleva a cabo el análisis de los resultados obtenidos mediante la aplicación de los instrumentos y técnicas de recolección de información especificados en la sección anterior; de igual forma, se encuentra distribuido de acuerdo con las fases expuestas en el Procedimiento metodológico:

- Fase I: Análisis de la situación actual
- Fase II: Propuesta de indicadores claves de desempeño
- Fase III: Diseño de estructura del *datamart*
- Fase IV: Diseño de procesos *ETL*
- Fase V: Elaboración de *dashboards* para la visualización de indicadores de desempeño

A continuación, se detalla el análisis de resultados para cada una de las anteriores fases donde se obtuvo información que contribuyó a la formulación de la propuesta de solución.

1.23. Fase I: Análisis de la situación actual

Por medio del análisis de resultados de esta fase, se obtuvo conocimiento con respecto a los sistemas de información de Progressio Digital, así como información sobre sus procesos y sus correspondientes bases de datos.

Para llevar a cabo este análisis, se recurrió a una entrevista semiestructurada con el administrador de proyectos de la compañía, para la identificación de los sistemas de información de la organización, así como la observación moderada del funcionamiento de estos sistemas para el análisis de sus procesos. Por otro lado, se realizó una revisión documental de las especificaciones técnicas de implementación de los *API* de estos sistemas, y los diseños lógicos de sus correspondientes bases de datos.

Es posible visualizar en detalle la aplicación de los instrumentos en los siguientes apéndices:

- Apéndice IX. Análisis de la situación actual de los sistemas de la organización
- Apéndice X. Revisión documental de las especificaciones técnicas e implementación de los APIs del sistema de procesamiento de criptomonedas
- Apéndice XI. Revisión documental de las especificaciones técnicas e implementación de los APIS del monedero electrónico.

- Apéndice XII. Revisión documental de los diseños lógicos de las bases de datos del sistema de procesamiento de criptomonedas.
- Apéndice XIII. Revisión documental de los diseños lógicos de las bases de datos del monedero electrónico.

Cómo resultado de la información recolectada a través de los instrumentos fue posible identificar los siguientes hallazgos.

1.23.1. Hallazgos sobre la situación actual de los sistemas de información de la organización y sus procesos

- Para la organización, los sistemas con principal necesidad de análisis son el encargado del procesamiento de criptomonedas y el monedero electrónico de criptomonedas.
- Una de las bases de datos cuenta con un procedimiento almacenado para la carga de tablas de tiempo, este será utilizado para la carga de los datos de la dimensión de tiempo de la solución.
- Los procesos dentro del sistema de procesamiento de criptomonedas incluyen:
 - Depósitos de criptomonedas
 - Retiros de criptomonedas
- Los procesos dentro del sistema de monedero de criptomonedas incluyen:
 - Planificación de retiros: el sistema automáticamente detecta fondos provenientes de un depósito a una cuenta configurada para la automatización del envío de retiros a otras cuentas. Este proceso conlleva la planificación de retiros en diferentes días y horas.
 - Compra de dinero cripto: el sistema permite el procesamiento de tarjetas de crédito para la compra de dinero cripto, generando a su vez transacciones de retiro y de depósito.
- Algunas características técnicas de los sistemas de la organización son las siguientes:
 - Las bases de datos de los sistemas de la organización se encuentran alojadas en los servidores *Amazon Web Services (AWS)* bajo el servicio de *Relational Database Service (RDS)*.
 - Los sistemas de la organización están alojados en servidores *Linux Ubuntu*, configurados dentro de una arquitectura en *Amazon Web Services*.
 - El motor de bases de datos de los sistemas es MySQL, y están actualizados a la última versión disponible: 8.0.
 - La convención para la escritura de los nombres de las tablas es *Snake Case*.
 - La convención para la escritura de los nombres de las columnas en las tablas es *Camel Case*.

1.24. Fase II: Propuesta de indicadores clave de desempeño

En esta fase se determinaron y priorizaron identificadores clave de desempeño, se tomó en cuenta las necesidades con respecto al control y monitoreo de los sistemas de información, por parte de la gerencia y el equipo técnico de desarrollo, así como las del principal cliente para la organización.

Para que fuera posible abarcar los requerimientos de estas partes interesadas del negocio, con respecto a los indicadores clave de desempeño, se llevaron a cabo entrevistas semiestructuras a una serie de sujetos: al gerente de la empresa, al administrador de proyectos, a desarrolladores de la organización y a un representante de la empresa cliente; asimismo, se efectuaron grupos focales con algunos de estos.

Por otro lado, se llevó a cabo la revisión documental de las consultas directas a bases de datos que solían utilizarse normalmente para monitorear aspectos sobre los sistemas en producción, con el fin de identificar los datos de mayor interés dentro de estas consultas.

Los anexos que evidencian la aplicación de los instrumentos de recolección de información son los siguientes:

- Apéndice XIV. Revisión documental consulta para revisión de transacciones del sistema de procesamiento de criptomonedas.
- Apéndice XV. Revisión documental ejecución de retiros planificados del monedero electrónico
- Apéndice XVI. Revisión documental sobre consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.
- Apéndice XVII. Observación del flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico con desarrollador de *software*.

Tras el análisis de la información recolectada, se identificó una serie de indicadores clave de desempeño listados en la **Tabla 10**.

Tabla 10. Listado de indicadores clave de desempeño para los sistemas de procesamiento de criptomonedas.

ID	Nombre	Descripción	Sistema	Fórmula
IND-1	Cantidad de depósitos por instancia de <i>core</i> .	Cantidad de depósitos determinada para cada una de las instancias de <i>core</i> .	Procesador de criptomonedas	
IND-2	Cantidad de retiros por instancia de <i>core</i> .	Cantidad de retiros determinada para cada una de las instancias de <i>core</i> .	Procesador de criptomonedas	
IND-3	Cantidad de transacciones instancia de <i>core</i> .	Cantidad total de transacciones determinada para cada una de las instancias de <i>core</i> .	Procesador de criptomonedas	
IND-4	Cantidad total de direcciones.	Cantidad de direcciones solicitadas y registradas en el sistema.	Procesador de criptomonedas	
IND-5	Porcentaje de uso de las direcciones.	Porcentaje de las direcciones registradas que realmente está	Procesador de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
		relacionada a una o más transacciones en <i>Blockchain</i> .		
IND-6	Promedio de tiempo en procesar retiros.	Tiempo promedio que toma publicarse un retiro en el <i>blockchain</i> .	Procesador de criptomonedas	
IND-7	Cantidad de depósitos pendientes.	Cantidad de depósitos que se encuentran pendientes de confirmaciones.	Procesador de criptomonedas	
IND-8	Cantidad de retiros pendientes.	Cantidad de retiros que ya fueron publicados en <i>blockchain</i> , pero que no han sido minados aún y se encuentran pendientes de confirmar.	Procesador de criptomonedas	
IND-9	Promedio de tiempo entre una solicitud de una dirección y una transacción en <i>blockchain</i> .	Tiempo promedio entre el momento en que una dirección es solicitada y registrada en el sistema, y el momento en que	Procesador de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
		una transacción relacionada con esa dirección es registrada en el <i>blockchain</i> .		
IND-10	Promedio de tiempo entre una solicitud de una dirección y la primera confirmación de una transacción en el <i>blockchain</i> .	Tiempo promedio en minutos entre el momento en que una dirección es solicitada y registrada en el sistema, y el momento en que una transacción relacionada con esa dirección obtiene confirmaciones en el <i>blockchain</i> .	Procesador de criptomonedas	
IND-11	Cantidad de depósitos por categoría de cliente.	Cantidad de depósitos determinada para cada uno de negocios comerciantes.	Procesador de criptomonedas	
IND-12	Cantidad de retiros por categoría de cliente.	Cantidad de retiros determinada para cada uno de negocios comerciantes.	Procesador de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
IND-13	Cantidad de transacciones por categoría de cliente.	Cantidad total de transacciones determinada para cada uno de negocios comerciantes.	Procesador de criptomonedas	
IND-14	Monto de dinero de depósitos por categoría de cliente.	Suma del monto de las transacciones de depósitos para cada uno de negocios comerciantes.	Procesador de criptomonedas	
IND-15	Monto de dinero de retiros por categoría de cliente.	Suma del monto de las transacciones de retiro para cada uno de negocios comerciantes.	Procesador de criptomonedas	
IND-16	Cantidad de depósitos internos por categoría de cliente.	Número de transacciones de tipo depósito para cada categoría de cliente.	Procesador de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
IND-17	Monto de dinero de depósitos internos por categoría de cliente.	Suma de los montos de las transacciones de depósito internas para cada categoría de cliente.	Procesador de criptomonedas	
IND-18	Cantidad de transacciones de retiros internos por categoría de cliente.	Número de las transacciones correspondientes a retiros internos para cada categoría de cliente.	Procesador de criptomonedas	
IND-19	Cantidad mínima pagada por comisión de mineros.	Monto mínimo pagado en una transacción por concepto de comisión para los mineros del <i>blockchain</i> .	Procesador de criptomonedas	
IND-20	Cantidad máxima pagada por comisión de mineros.	Monto máximo pagado en una transacción por concepto de comisión a los mineros del <i>blockchain</i> .	Procesador de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
IND-21	Cantidad promedio pagada por comisión de mineros.	Monto pagado en promedio a los mineros del <i>blockchain</i> .	Procesador de criptomonedas	
IND-22	Monto no procesado para los retiros planificados.	Monto de los retiros planificados que no han sido procesados.	Monedero electrónico de criptomonedas	
IND-23	Cantidad de retiros planificados sin procesar.	Número de transacciones de retiro que fueron planificados, pero no han sido procesadas.	Monedero electrónico de criptomonedas	
IND-24	Monto procesado de los retiros planificados.	Monto referente al dinero que ha sido procesado por medio de los retiros planificados.	Monedero electrónico de criptomonedas	
IND-25	Cantidad de retiros planificados procesados.	Número de transacciones de retiros planificados que ya fueron procesadas exitosamente.	Monedero electrónico de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
IND-26	Monto promedio de los retiros planificados.	Monto de dinero asignado en promedio a las transacciones de retiro planificadas.	Monedero electrónico de criptomonedas	
IND-27	Monto de las transacciones de compra de dinero cripto procesadas.	Sumatoria del monto para las transacciones de tipo compra de dinero cripto.	Monedero electrónico de criptomonedas	
IND-28	Cantidad de transacciones de compra dinero cripto que han sido procesadas.	Número de transacciones de compra de dinero cripto se han procesado a través del sistema.	Monedero electrónico de criptomonedas	
IND-29	Monto promedio de las transacciones de compra de dinero cripto	Promedio del monto de las transacciones de tipo compra de dinero cripto.	Monedero electrónico de criptomonedas	
IND-30	Porcentaje de aprobación de transacciones por parte de los procesadores de tarjeta de crédito.	Porcentaje de transacciones que son aprobadas por parte de los procesadores de tarjetas de crédito con respecto al total de	Monedero electrónico de criptomonedas	

ID	Nombre	Descripción	Sistema	Fórmula
		transacciones de compra de dinero cripto existentes		
IND-31	Porcentaje de aprobación de transacciones por parte del procesador de criptomonedas.	Porcentaje de aprobación de transacciones por parte del procesador de criptomonedas con respecto al total de transacciones que ya se habían aprobado por parte del procesador de tarjetas de crédito.	Monedero electrónico de criptomonedas	

Nota. Se detalla el listado de indicadores identificados para los sistemas de procesamiento de criptomonedas. Fuente: elaboración propia. (2020).

Luego de identificar y detallar los indicadores clave de desempeño para los principales sistemas relacionados con el procesamiento de criptomonedas, se llevó a cabo un grupo focal para la priorización de estos. El detalle de lo conversado en el grupo focal se encuentra en el Apéndice XVIII. Grupo focal para la revisión y priorización de los *KPI* propuestos con gerente general, administrador de proyectos y desarrollador de *software*

La priorización de los indicadores clave de desempeño determina el nivel de urgencia en la implementación de *dashboards* para la visualización de la información. Los niveles de priorización utilizados se describen en la **Figura 8**.

Figura 8. Niveles de prioridad para los indicadores clave de desempeño



Figura 8. Niveles de prioridad para los indicadores clave de desempeño. Fuente: elaboración propia. (2020).

En la **Tabla 11** se muestra el identificador único de cada indicador, así como su nombre y el nivel de prioridad asignado sobre este.

Tabla 11. Priorización de los indicadores clave de desempeño

Identificador del indicador	Nombre del indicador	Nivel de prioridad
IND-1	Cantidad de depósitos por instancia de core.	Medio
IND-2	Cantidad de retiros por instancia de core.	Medio
IND-3	Cantidad de transacciones instancia de core.	Medio
IND-4	Cantidad total de direcciones.	Bajo
IND-5	Porcentaje de uso de las direcciones.	Bajo
IND-6	Promedio de tiempo en procesar retiros.	Medio
IND-7	Cantidad de depósitos pendientes.	Medio
IND-8	Cantidad de retiros pendientes.	Medio
IND-9	Promedio de tiempo entre una solicitud de una dirección y una transacción en <i>blockchain</i> .	Bajo
IND-10	Promedio de tiempo entre una solicitud de una dirección y la primera confirmación de una transacción en el <i>blockchain</i> .	Bajo
IND-11	Cantidad de depósitos por categoría de cliente.	Alto
IND-12	Cantidad de retiros por categoría de cliente.	Alto
IND-13	Cantidad de transacciones por categoría de cliente.	Alto
IND-14	Monto de dinero de depósitos por categoría de cliente.	Alto

Identificador del indicador	Nombre del indicador	Nivel de prioridad
IND-15	Monto de dinero de retiros por categoría de cliente.	Alto
IND-16	Cantidad de depósitos internos por categoría de cliente.	Bajo
IND-17	Monto de dinero de depósitos internos por categoría de cliente.	Bajo
IND-18	Cantidad de transacciones de retiros internos por categoría de cliente.	Bajo
IND-19	Cantidad mínima pagada por comisión de mineros.	Bajo
IND-20	Cantidad máxima pagada por comisión de mineros.	Medio
IND-21	Cantidad promedio pagada por comisión de mineros.	Alto
IND-22	Monto no procesado para los retiros planificados.	Medio
IND-23	Cantidad de retiros planificados sin procesar.	Medio
IND-24	Monto procesado de los retiros planificados.	Medio
IND-25	Cantidad de retiros planificados procesados.	Medio
IND-26	Monto promedio de los retiros planificados.	Medio
IND-27	Monto de las transacciones de compra de dinero cripto procesadas	Alto
IND-28	Cantidad de transacciones de compra dinero cripto que han sido procesadas	Alto

Identificador del indicador	Nombre del indicador	Nivel de prioridad
IND-29	Monto promedio de las transacciones de compra de dinero cripto	Alto
IND-30	Porcentaje de aprobación de transacciones por parte de los procesadores de tarjeta de crédito	Bajo
IND-31	Porcentaje de aprobación de transacciones por parte del procesador de criptomonedas	Bajo

Nota. Priorización de los indicadores clave de desempeño. Fuente: elaboración propia. (2020).

Es importante aclarar que el desarrollo del presente proyecto abarcó el diseño del modelado dimensional, el desarrollo de los procesos *ETL* y la generación de las visualizaciones para los indicadores clave de desempeño identificados con un nivel de prioridad alto, debido a que estos representan una urgencia mayor con respecto a los demás.

1.25. Fase III: Diseño del *datamart*

En esta fase se detallaron las necesidades con respecto al diseño lógico y a la arquitectura que soporta el almacén de datos de la solución de IN. Debido a que el repositorio dimensional representa la base para el desarrollo de esta solución, fue fundamental identificar correctamente las necesidades existentes con respecto a este.

Para la definición de los requerimientos del almacén de datos, se llevó a cabo grupos focales principalmente con el administrador de proyectos de la organización y con los ingenieros *devops*, encargados de mantener y configurar la actual arquitectura de los sistemas de información de la organización.

La información se obtuvo por la aplicación de un grupo focal detallado en el Apéndice XIX. Grupo focal para la definición de requerimientos del *datamart* con administrador de proyectos, gerente general e ingenieros *DevOps*.

1.25.1. Identificación de requerimientos relacionados con el diseño del *datamart*

El análisis de resultados de la aplicación de los instrumentos para esta fase permitió la identificación de los requerimientos que se encuentran especificados detalladamente en la **Tabla 12**.

Tabla 12. *Requerimientos relacionados con el diseño del datamart*

Identificador del requerimiento	Descripción
RDM-1	Se debe de construir el repositorio del <i>datamart</i> en un <i>Relational Database Service (RDS)</i> de <i>Amazon Web Services (AWS)</i> , siguiendo la misma infraestructura de las demás bases de datos de la organización.
RDM-2	Las bases de datos deben ser manejadas por el motor <i>MySQL</i> y contar con la última versión disponible, es decir la 8.0.

Identificador del requerimiento	Descripción
RDM-3	Se debe restringir la conexión a las bases de datos solamente para las aplicaciones que realmente requieran de ellas.
RDM-4	Se debe implementar un área de trabajo o <i>staging area</i> para no aumentar la carga sobre las bases de datos operacionales de los sistemas de información de la organización por consultas directas hacia el <i>datamart</i> .
RDM-5	Para el desarrollo del diseño de la base de datos se debe utilizar la herramienta <i>Navicat</i> , cuya licencia ya posee la organización.

Nota. Detalle de requerimientos relacionados con el *datamart*. Fuente: elaboración propia. (2020).

El análisis de resultados de la Fase IV: Diseño de los procesos *ETL* y de la Fase V: Elaboración de *dashboards* para la visualización de los indicadores clave de desempeño, se basó en los requerimientos presentados anteriormente, debido a que estos son la base del almacén de datos que sostendrá la solución de inteligencia de negocios; así como en los indicadores clave de desempeño clasificados de nivel alto por su urgencia.

1.26. Fase IV: Diseño de los procesos *ETL*

Una vez determinadas las necesidades con respecto al diseño y a la arquitectura del *datamart*, fue posible llevar a cabo la identificación de los requerimientos con respecto al diseño y desarrollo de los procesos de extracción, transformación y carga de datos.

Para la identificación de los requerimientos, se recurrió a grupos focales con el administrador de proyectos y con los ingenieros *devops* de la organización. Con base en estos requerimientos, fueron identificados los criterios evaluados en la Comparativa de herramientas para desarrollo de procesos ETL; de esta manera se aseguró la escogencia de la solución que mejor satisfizo las necesidades de la organización y de la solución de IN. Los requerimientos descritos y el peso asignado a los criterios de escogencia de la solución, se basan en el análisis de los resultados obtenidos en el Apéndice XX. Grupo focal con administrador de proyectos e ingenieros *DevOps* para la definición de requerimientos relacionados al desarrollo de los procesos *ETL*.

1.26.1. Identificación de requerimientos relacionados con el diseño de los procesos *ETL*

En la **Tabla 13** se describe cada uno de los requerimientos identificados referentes al diseño de los procesos *ETL* para la solución de IN.

Tabla 13. *Requerimientos relacionados con el diseño de procesos ETL*

Identificador del requerimiento	Descripción
ETL-1	La ejecución de los procesos ETL debe llevarse a cabo en un servidor Linux Ubuntu de la organización y no solamente en la máquina de trabajo del desarrollador del proyecto.
ETL-2	La ejecución de los procesos ETL debe ser automatizada y calendarizada cada 24 horas.

Identificador del requerimiento	Descripción
ETL-3	La herramienta para los procesos ETL debe ser lo más amigable posible con el usuario para agilizar el proceso de aprendizaje y de desarrollo como tal.
ETL-4	La herramienta de desarrollo de los procesos ETL debe contar con soporte de la industria, es decir, una comunidad de usuarios y soporte por parte de los proveedores.
ETL-5	La herramienta debe tener la capacidad de conectarse a las actuales bases de datos de los sistemas de la organización, así como al almacén de datos. Ambos manejados por MySQL versión 8.0.
ETL-6	Debe considerarse el costo como el criterio de mayor peso en la escogencia de la herramienta.

Nota. Detalle de requerimientos relacionados con el diseño de los procesos *ETL*. Fuente: elaboración propia. (2020).

1.26.2. Comparativa de herramientas para desarrollo de procesos ETL

Con base en los requerimientos descritos, fueron identificados los criterios para la selección de la herramienta con la que se desarrollaron los procesos *ETL*. Para cada criterio se asignó un peso basado en el cual se calificó cada herramienta, es posible observar esta información en la **Tabla 14**.

Tabla 14. Criterios de selección de herramientas para desarrollo de ETL

Criterio	Peso asignado
Costo	40 %
Conexión a las bases de datos de la empresa	30 %
Usabilidad	15 %
Soporte de la industria	15 %

Nota. Detalle de los criterios de escogencia de herramientas de ETL y su peso. Fuente: elaboración propia. (2020).

Asimismo, las herramientas fueron evaluadas para cada uno de los criterios, para esto se basó en los siguientes puntos:

- 0: no cumple con lo definido en el requerimiento.
- 0.25: tiene deficiencias para cumplir con el requerimiento
- 0.5: cumple con algunos aspectos del requerimiento.
- 0.75: cumple el requerimiento.
- 1: cumple de sobremanera el requerimiento.

El criterio relacionado con el costo de las herramientas fue evaluado otorgando los valores anteriormente definidos (0, 0.25, 0.5, 0.75 y 1), pero, de acuerdo con el orden sucesivo de precios de las herramientas; donde se califica 1 a la herramienta con menor costo.

En la **Tabla 15**, se describe la justificación sobre la escogencia de cada una de las herramientas para diseño de procesos ETL que se analizaron para la solución.

Tabla 15. *Justificación de la escogencia de las herramientas de ETL analizadas.*

Herramienta	Justificante
Informatica Data Engineering Integration	Herramienta mejor posicionada en el Cuadrante Mágico de Gartner sobre soluciones para la integración de datos, mostrado en el Anexo I. Cuadrante mágico de Gartner para herramientas de integración de datos
Apache Kafka	Recomendación de ingeniero <i>devops</i> de la organización. Ver Apéndice XX. Grupo focal con administrador de proyectos e ingenieros <i>DevOps</i> para la definición de requerimientos relacionados al desarrollo de los procesos <i>ETL</i> .
Amazon Glue	Herramienta ofrecida por el principal proveedor de servicios de la compañía actualmente (<i>Amazon Web Services</i>).
Talend Open Data Studio for Data Integration	Herramienta <i>Open Source</i> , con versión gratuita, líder en el mercado según el Cuadrante Mágico de Gartner de soluciones de integración de datos, mostrado en el Anexo I. Cuadrante mágico de Gartner para herramientas de integración de datos
IBM InfoSphere	Herramienta líder en el mercado según el Cuadrante Mágico de Gartner de soluciones de integración de datos, mostrado en el Anexo I. Cuadrante mágico de Gartner para herramientas de integración de datos

Nota. Detalle de la justificación de escogencia de las herramientas de ETL analizadas. Fuente: elaboración propia. (2020).

Seguidamente, en la **Tabla 16** se encuentra cada herramienta evaluada con respecto al cumplimiento de los criterios, esta se basa en el Apéndice VII. Tabla comparativa de herramientas

Tabla 16. Tabla comparativa de herramientas para el desarrollo de procesos ETL.

Nombre de herramienta	Criterios			
	Costo	Conexión a las bases de datos de la empresa	Usabilidad	Soporte de la industria
Informatica Data Engineering Integration	El plan más básico es de \$2000 USD mensuales	Permite la conexión con MySQL 8.0	Cuenta con una interfaz <i>drag & drop</i> ; además, se permite realizar la importación y especificación de los datos de manera manual.	Se encuentra posicionada de líder en el cuadrante mágico de <i>Gartner</i> sobre herramientas de integración de datos.
IBM InfoSphere	\$120 USD al mes por dos usuarios, suficientes para el alcance de la solución.	Permite la conexión a MySQL 8 al instalar el conector <i>ODBC</i>	Interfaz gráfica moderna que permite el <i>drag & drop</i> de componentes.	Se encuentra posicionada de líder en el cuadrante mágico de <i>Gartner</i> sobre herramientas de integración de datos.

Nombre de herramienta	Criterios			
	Costo	Conexión a las bases de datos de la empresa	Usabilidad	Soporte de la industria
Talend Open Data Studio for Data Integration	La versión gratuita de la herramienta soporta las necesidades de la organización.	Permite la conexión para diversas fuentes de información y salidas de escritura, entre ellas MySQL 8.0	Cuenta con una interfaz de tipo <i>drag & drop</i> . Algunas personalizaciones de los componentes requieren código Java.	Se encuentra posicionada de líder en el cuadrante mágico sobre herramientas de integración de datos, cuenta con foros de soporte y comunidades para los usuarios.
Amazon Glue	La configuración que satisface las necesidades de la organización está estimada en los \$200 USD.	No permite la conexión con bases de datos MySQL con una versión mayor a la 5.7, al menos de incurrir en configuraciones y cambios en la arquitectura de los servidores de la organización.	Cuenta con interfaz paso a paso, pero la mayor parte del desarrollo debe hacerse manualmente por medio del lenguaje de programación Java.	Cuenta con revisión de referentes importantes como <i>Gartner</i> , sin embargo, no se encuentra en el cuadrante mágico. Por otro lado, AWS cuenta con varios foros sobre su servicio y un equipo de soporte 24/7.

Nombre de herramienta	Criterios			
	Costo	Conexión a las bases de datos de la empresa	Usabilidad	Soporte de la industria
Apache Kafka	Es una solución gratuita y <i>open source</i> .	Permite la conexión con variedad de bases de datos, incluyendo <i>MySQL 8</i> por medio de la instalación de un conector adicional.	Su usabilidad se limita a la ejecución de comandos en una consola <i>Linux</i> .	Hay soluciones basadas en Kafka con tutoriales, foros y recursos relacionados con la herramienta; de igual forma, existen otras organizaciones que le dedican artículos y análisis, tales como Gartner.

Nota. Detalle de la comparativa de herramientas para desarrollo de procesos *ETL*. Fuente: elaboración propia. (2020).

Por otro lado, en la **Tabla 17** se muestra la comparativa realizada, se le otorgan los correspondientes pesos, según lo definido con anterioridad. Esta tabla también se basa en el Apéndice VII. Tabla comparativa de herramientas.

Tabla 17. Comparativa de herramientas de desarrollo de ETL con su respectivo peso

Nombre de herramienta	Criterios				TOTAL
	Costo 40%	Conexión a las bases de datos de la empresa 30%	Usabilidad 15%	Soporte de la industria 15%	
<i>Informatica Data Engineering Integration</i>	$40 \times 0.25 = 10$	$30 \times 1 = 30$	$15 \times 1 = 15$	$15 \times 1 = 15$	70 %
<i>Apache Kafka</i>	$40 \times 1 = 40$	$30 \times 1 = 30$	$15 \times 0.25 = 3.75$	$15 \times 0.75 = 11.2$	84.95%
<i>Amazon Glue</i>	$40 \times 0.5 = 20$	$30 \times 0.75 = 22.5$	$15 \times 0.5 = 7.5$	$15 \times 0.75 = 11.2$	61.2%
<i>Talend Open Studio for Data Integration</i>	$40 \times 1 = 40$	$30 \times 1 = 30$	$15 \times 0.75 = 11.2$	$15 \times 1 = 15$	96.2%
<i>IBM InfoSphere</i>	$40 \times .075 = 30$	$30 \times 1 = 30$	$15 \times 1 = 11.2$	$15 \times 1 = 15$	86.2%

Nota. Detalle de requerimientos relacionados con el *datamart*. Fuente: elaboración propia. (2020).

Como es posible observar en el anterior análisis, la herramienta con mayor puntuación es *Talend Open Studio for Data Integration*, es por esta razón que fue seleccionada para el desarrollo de los procesos *ETL* del proyecto.

1.27. Fase V: Elaboración de *dashboards* para la visualización de los indicadores clave de desempeño

Luego de especificar los requerimientos para el diseño del almacén de datos y los correspondientes al diseño de los procesos *ETL*, se continuó con el análisis de resultados, de la identificación de requerimientos para la elaboración de los *dashboards*, para la visualización de los indicadores clave de desempeño. De igual forma, se procedió a la escogencia de la herramienta para realizar estas visualizaciones de información.

Con el fin de identificar adecuadamente las necesidades de la organización, se llevó a cabo una serie de acciones, entre ellas: la revisión documental de las especificaciones de las herramientas, entrevistas no estructuradas al administrador de proyectos y al gerente general, así como grupos focales para revisión de pesos de los criterios, así como la presentación y revisión de los avances.

El detalle de la aplicación de los instrumentos se presenta en los siguientes apéndices:

- Apéndice XXI. Grupo focal con administrador de proyectos y gerente general para para la definición de requerimientos relacionados las visualizaciones.

1.27.1. Identificación de requerimientos relacionados a la elaboración de los *dashboards*

En la **Tabla 18**, se detallan los requerimientos referentes al diseño de los procesos *ETL*.

Tabla 18. *Requerimientos relacionados con la elaboración de dashboards.*

Identificador del requerimiento	Descripción
DSB-1	La aplicación de visualización debe permitir al menos cinco usuarios de lectura y un usuario administrador o de escritura.
DSB-2	Las visualizaciones deben ser interactivas y permitir el filtrado de la información a los usuarios.
DSB-3	La aplicación de visualización debe tener acceso a los datos provenientes del almacén de datos.

Identificador del requerimiento	Descripción
DSB-4	La herramienta seleccionada debe permitir la calendarización del envío por correo electrónico de <i>dashboards</i> .
DSB-5	La herramienta de visualización debe permitir el refrescamiento de la información de manera automatizada al menos cada 24 horas.
DSB-6	El acceso de la herramienta de visualización debe poder limitarse por una dirección IP privada, esto permite una conexión directa y segura con la instancia RDS del almacén de datos.
DBS-7	Debe considerarse el costo como el criterio de mayor peso en la escogencia de la herramienta.

Nota. Detalle de requerimientos relacionados a las visualizaciones. Fuente: elaboración propia. (2020).

1.27.2. Comparativa de herramientas para elaboración de los *dashboards*

Los siguientes criterios de selección para la herramienta de visualización fueron identificados con base en los requerimientos definidos anteriormente. Para cada criterio se asignó un peso basado en el cual se calificó cada herramienta, es posible observar esta información en la **Tabla 19**.

Tabla 19. Criterios de escogencia de herramientas de *dashboards* y su peso

Criterio	Peso asignado
Costo	40 %
Conexión segura a las fuentes de información	30 %
Envío por correo electrónicos	20 %
Refrescamiento de datos cada 24 horas	10 %

Nota. Detalle de los criterios de escogencia de herramientas de *dashboards* y su peso. Fuente: elaboración propia. (2020).

Las herramientas fueron evaluadas para cada uno de los criterios basándose en lo especificado a continuación:

- 0: no cumple lo definido en el requerimiento.
- 0.25: tiene deficiencias para cumplir con el requerimiento
- 0.5: cumple con algunos aspectos del requerimiento.
- 0.75: cumple el requerimiento.
- 1: cumple de sobremanera el requerimiento.

El criterio relacionado con el costo de las herramientas fue evaluado otorgando los valores anteriormente definidos (0, 0.25, 0.5, 0.75 y 1), pero de acuerdo con el orden sucesivo de precios de las herramientas; donde se califica 1 a la herramienta con menor costo.

En la **Tabla 20** se describe la justificación sobre la escogencia de cada una de las herramientas para el desarrollo de los *dashboards*, estas fueron analizadas para la solución de IN.

Tabla 20. Justificación de escogencia de herramientas de visualización de datos a evaluar

Herramienta	Justificante
Tableau	Segunda herramienta mejor posicionada en el Cuadrante Mágico de Gartner sobre soluciones para la visualización de información, mostrado en el Anexo II.
Microsoft Power BI	Herramienta mejor posicionada en el Cuadrante Mágico de Gartner sobre soluciones para la visualización de información, mostrado en el Anexo II.
Google Data Studio	Petición de parte del administrador de proyectos de la organización. Ver Apéndice XXI. Grupo focal con administrador de proyectos y gerente general para para la definición de requerimientos relacionados las visualizaciones.
Amazon Quicksight	Herramienta actualmente ofrecida por el principal proveedor de servicios de la organización (<i>Amazon Web Services</i>)

Nota. Detalle de la Justificación de escogencia de herramientas de visualización de datos a evaluar. Fuente: elaboración propia. (2020).

En la **Tabla 21** se encuentra cada herramienta evaluada con respecto al cumplimiento de los criterios, esta se basa en el Apéndice VII. Tabla comparativa de herramientas.

Tabla 21. Tabla comparativa de herramientas para el desarrollo de *dashboards*.

Nombre de herramienta	Criterios			
	Costo	Conexión segura a las fuentes de información	Envío por correo electrónico	Refrescamiento de datos cada 24 horas
Tableau	La configuración de un usuario creador y el mínimo de 100 usuarios de lectura dado por <i>Tableau</i> para la versión <i>Tableau Server</i> de sus servicios, sumaría un costo total de \$1270 USD al mes.	Es posible utilizar la IP privada de la licencia de <i>Tableau Server</i> , que es específica de donde se aloja la solución.	Permite el envío por correo de los <i>dashboards</i>	Permite refrescar la información de manera diaria
Google Data Studio	No tiene ningún costo.	Se debe configurar 13 direcciones IP desde el almacén de datos para permitir acceso a esta solución. Todas estas direcciones IP son compartidas por la	Permite el envío por correo de los <i>dashboards</i>	Las fuentes de información se actualizan automáticamente cada vez que se abre el reporte o se cambia el filtro del tiempo, lo que afecta en

Nombre de herramienta	Criterios			
	Costo	Conexión segura a las fuentes de información	Envío por correo electrónico	Refrescamiento de datos cada 24 horas
		comunidad de manera pública, no son definidas para el uso de una solución en específico.		cierta medida al desempeño del reporte.
Microsoft Power BI	El costo es de \$10 USD por mes por usuario para la licencia <i>Power BI Pro</i> que cubre las necesidades de la organización. Este precio es establecido sin importar los permisos con que cuente cada usuario; por lo tanto, el costo mensual sería de \$60 USD para soportar los requerimientos actuales	Se debe realizar varias configuraciones de lado del servicio del almacén de datos para permitir la restricción del acceso para esta herramienta, aunque estas configuraciones no son privadas, por lo que no permiten restringir el acceso únicamente desde una solución, sino que las	Permite el envío por correo de los <i>dashboards</i>	Permite refrescar la información hasta ocho veces por día.

Nombre de herramienta	Criterios			
	Costo	Conexión segura a las fuentes de información	Envío por correo electrónico	Refrescamiento de datos cada 24 horas
	con respecto a los usuarios de la organización.	configuraciones mencionadas son compartidas por la comunidad en general.		
Amazon Quicksight	El precio depende del uso por parte de los usuarios, para cumplir con los requerimientos actuales de la organización, el costo se mantendría entre los \$18 USD y los \$43 USD mensuales.	La dirección IP es privada y se maneja por la organización, debido a que se aloja en los servidores privados de AWS.	Permite el envío por correo de los <i>dashboards</i>	Se permite crear varias calendarizaciones de refrescamiento de la información por día, por semana o por mes.

Nota. Detalle de la comparativa de herramientas para desarrollo de *dashboards*. Fuente: Adaptado de Power BI vs Tableau vs Amazon QuickSight – Which is Best for Your Organization, por K, Adams, 2020; A Comparative Analysis of Top 6 BI and Data Visualization Tools in 2018, por K, Bobriakov, 2018; Conectarse a MySQL, por Data Studio, 2020; Precios de Amazon QuickSight, por Amazon, 2020; Whitelist PowerBI service FQDNs or IP ranges for Connectors to work, por Microsoft, 2018; Add Power BI URLs to your allow list., por Microsoft, 2020; The READ-ONLY License for Power BI., por Radacad, 2020 y Tableau data Management, por Tableau, 2020.

Por otro lado, en la **Tabla 22** se muestra la comparativa realizada otorgando los correspondientes pesos, según lo definido con anterioridad. Esta tabla también se basa en el Apéndice VII. Tabla comparativa de herramientas.

Tabla 22. Comparativa de herramientas de desarrollo de *dashboards* con su respectivo peso

Nombre de herramienta	Criterios				TOTAL
	Costo 40 %	Conexión segura a las fuentes de información 30 %	Envío por correo electrónico 20%	Refrescamiento de datos cada 24 horas 10 %	
<i>Tableau</i>	$40 \times 0.25 = 10$	$30 \times 1 = 30$	$20 \times 1 = 20$	$10 \times 0.75 = 7.5$	67.5 %
<i>Google Data Studio</i>	$40 \times 1 = 40$	$30 \times 0.25 = 7.5$	$20 \times 1 = 20$	$10 \times 0.5 = 5$	72.5 %
<i>Microsoft Power BI</i>	$40 \times 0.5 = 20$	$30 \times 0.25 = 7.5$	$20 \times 1 = 20$	$10 \times 1 = 10$	57.5 %
<i>Amazon Quicksight</i>	$40 \times 0.75 = 30$	$30 \times 1 = 30$	$20 \times 1 = 20$	$10 \times 1 = 10$	90 %

Nota. Detalle de requerimientos relacionados con el *datamart*. Fuente: elaboración propia. (2020).

Como es posible observar en el análisis anterior, la herramienta con mayor puntuación es *Amazon Quicksight*, esta herramienta es brindada por el principal proveedor de servicios de la organización, *AWS*

CAPÍTULO V: PROPUESTA DE SOLUCIÓN

5. CAPÍTULO V: PROPUESTA DE SOLUCIÓN

Este capítulo documenta las fases sobre el desarrollo de soluciones para el cumplimiento de los objetivos propuestos. Para ello, se utiliza como referencia el marco conceptual y los resultados obtenidos después de realizado el proceso para la recolección de datos.

Esta propuesta de solución abarcó la ejecución de las tres últimas fases del Procedimiento metodológico definido en el tercer capítulo:

- Fase III: Diseño de estructura del *datamart*
- Fase IV: Diseño de procesos *ETL*
- Fase V: Elaboración de *dashboards* para la visualización de indicadores de desempeño

Es relevante aclarar que la fase inicial del proceso metodológico:

Análisis de la situación actual, no es abarcada en este capítulo debido a que esta fase se llevó a cabo a través del desarrollo del CAPÍTULO IV: ANÁLISIS DE RESULTADOS como medio para la recolección de información con respecto a los sistemas de información existentes en la organización, y no representa parte del desarrollo de la solución de inteligencia de negocios como un producto. Por otro lado, la Fase II: Propuesta de indicadores claves de desempeño, también fue abarcada en ese capítulo con la identificación y documentación de indicadores clave de desempeño para los sistemas de procesamiento de criptomonedas de la organización, presentada en la sección 1.24.

En las secciones 1.28, 1.29 y 1.30; se detalla la propuesta de solución relacionada a cada una de las fases indicadas con anterioridad.

1.28. Fase III: Diseño del *datamart*

La propuesta de solución para el diseño del *datamart*, se basa en la Identificación de requerimientos relacionados con el diseño del *datamart*, establecido en el capítulo anterior.

1.28.1. Diseño de la arquitectura técnica

El primer aspecto por considerar es la especificación técnica de la arquitectura, se refiere al ambiente de las bases de datos que conforman el *datamart*.

1.28.1.1. Especificaciones técnicas

El alojamiento de las bases de datos de la organización se encuentra en servidores en la nube, brindados por *Amazon Web Services*, específicamente a través del servicio de *Amazon RDS*. Por lo anterior, el almacén de datos encargado de contener la información del procesamiento de criptomonedas fue alojado sobre este mismo servicio. La **Figura 9** muestra los detalles de la instalación realizada por los ingenieros *devops* para el alojamiento del repositorio dimensional.

Figura 9. Detalles de instalación de la instancia del datamart

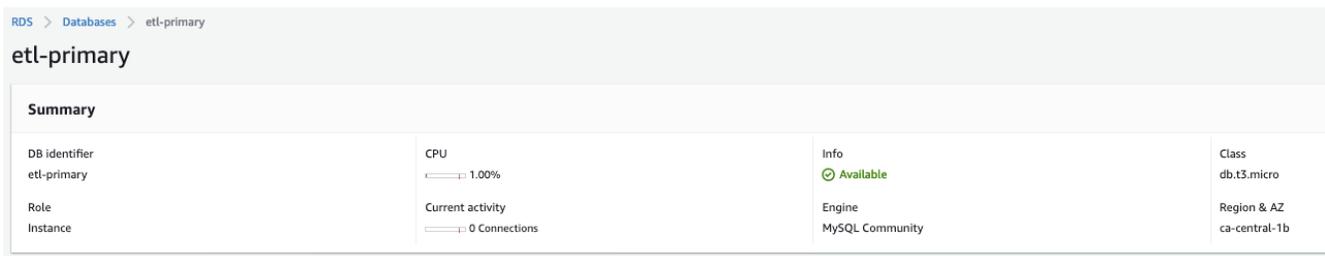


Figura 9. Detalles de instalación de la instancia del datamart. Fuente: Elaboración propia. (2020).

Cabe destacar la configuración de la instancia del servicio *Amazon RDS*, en este caso, el motor de la base de datos es *MySQL*, específicamente por encima de la versión 8.0, al igual que las demás bases de datos. Por otro lado, la clase de la instancia corresponde a la más básica ofrecida dentro de los paquetes *RDS*, que incluye 1GB de memoria *RAM* y un almacenamiento inicial de 20GB, esto fue configurado de esta manera con el fin de ahorrar costos a la organización; sin embargo, la solución permite escalar el servicio sin mayores problemas. Las configuraciones específicas con respecto a la instancia del *RDS* se indican en la **Figura 10**.

Figura 10. Detalles de la configuración de la instancia RDS del *datamart*

Configuration	Instance class	Storage
DB instance id etl-primary	Instance class db.t3.micro	Encryption Enabled
Engine version 8.0.17	vCPU 2	KMS key aws/rds
DB name -	RAM 1 GB	Storage type General Purpose (SSD)
License model General Public License	Availability	IOPS -
Option groups default:mysql-8-0	Master username administrator	Storage 20 GiB
ARN arn:aws:rds:ca-central-1:861677023151:db:etl-primary	IAM db authentication Not Enabled	Storage autoscaling Enabled
Resource id db-S5LLG6FSYJMSANB7WPGW4HVSDM	Multi AZ No	Maximum storage threshold 1000 GiB

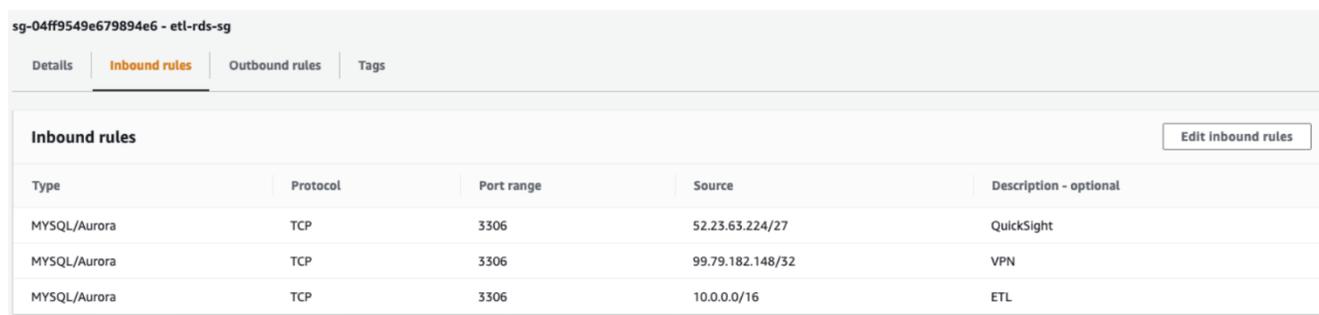
Figura 10. Detalles de la configuración de la instancia RDS del *datamart*. Fuente: elaboración propia. (2020).

1.28.1.2. Gestión de accesos

Uno de los requerimientos relevantes con respecto a la configuración del *datamart* fue el de limitar los accesos sobre la conexión a las bases de datos, con la finalidad de conseguir un nivel de seguridad similar al existente en otras instancias de *RDS* de la organización.

Para configurar esta capa extra de seguridad, se restringió el acceso de acuerdo con las direcciones de IP desde las que se intenta establecer la conexión con las bases de datos, de manera que solamente las que estuvieran en el *white list* contaran con la capacidad de realizar la conexión, de manera exitosa. Las reglas de seguridad donde se restringió el acceso a la conexión se muestran en la **Figura 11**.

Figura 11. Configuración de accesos de la conexión a la instancia *RDS*



sg-04ff9549e679894e6 - etl-rds-sg				
Details Inbound rules Outbound rules Tags				
Inbound rules Edit inbound rules				
Type	Protocol	Port range	Source	Description - optional
MYSQL/Aurora	TCP	3306	52.23.63.224/27	QuickSight
MYSQL/Aurora	TCP	3306	99.79.182.148/32	VPN
MYSQL/Aurora	TCP	3306	10.0.0.0/16	ETL

Figura 11. Configuración de accesos de la conexión a la instancia *RDS*. Fuente: elaboración propia. (2020).

Como es posible observar en la **Figura 11**, se configuró el acceso para tres direcciones IP solamente, estas son las siguientes:

- Dirección IP de la instancia donde está alojada la herramienta de visualización de datos configurada: *Amazon Quicksight*.
Esta herramienta necesita del acceso a las bases de datos del *datamart*, con el fin de consultar la información que posteriormente es transformada en *dashboards* o visualizaciones.
- Dirección IP correspondiente a la *VPN* utilizada por los desarrolladores de la organización.
Este permiso resultaba necesario para que el desarrollador del proyecto tuviese el acceso requerido para el diseño, desarrollo y validación del almacén de datos como tal, así como de los procesos *ETL* desarrollados localmente en la herramienta de escritorio *Talend Open Data Studio for data Integration*.
- Dirección IP de la instancia en *AWS* donde se aloja el servidor de *Linux Ubuntu*.
Este servidor requería del acceso debido a que se encarga de la ejecución y calendarización de los procesos *ETL*, que cargan la información en el *datamart*.

La **Figura 12**, muestra el resultado de un intento de conexión a la base de datos desde una dirección IP de las configuradas en el *white list*. Por otro lado, la **Figura 13** muestra una conexión exitosa a través de la dirección IP de la *VPN* de desarrollo.

Figura 12. Intento fallido de conexión a las bases de datos del *datamart*

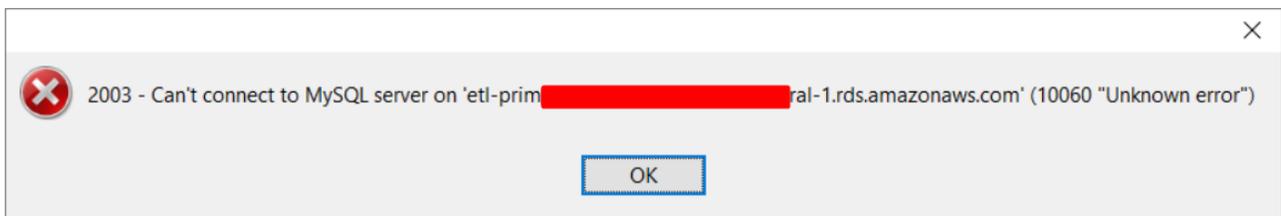


Figura 12. Intento fallido de conexión a las bases de datos del *datamart*. Fuente: elaboración propia. (2020).

Figura 13. Intento exitoso de conexión a las bases de datos del *datamart*.

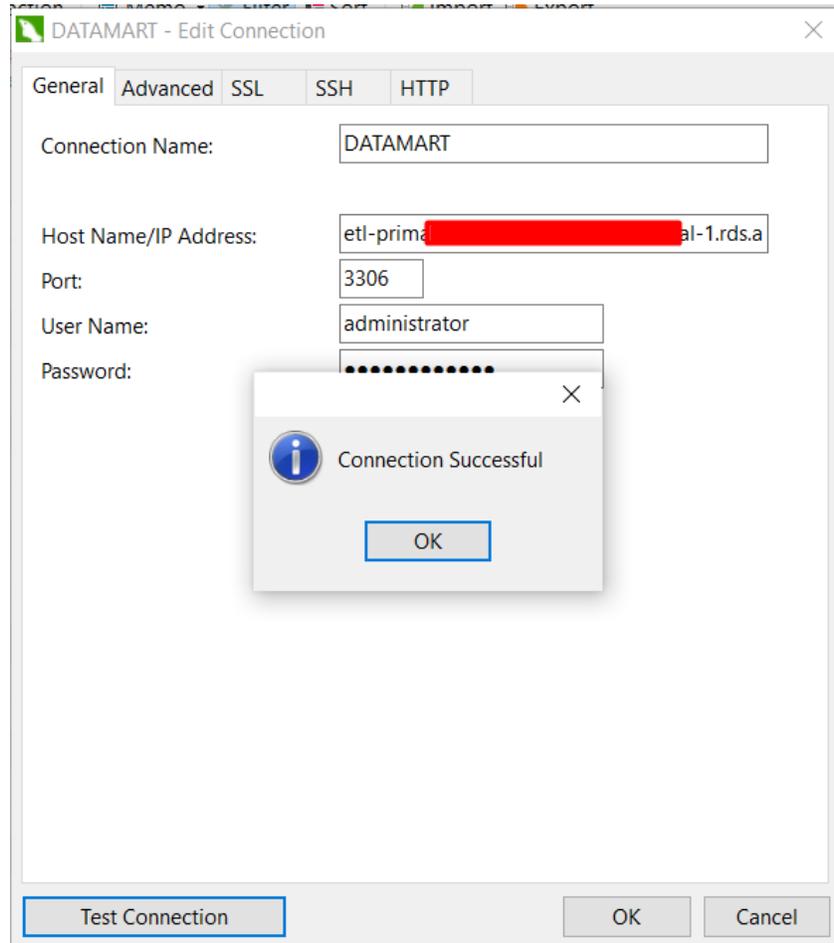


Figura 13. Intento exitoso de conexión a las bases de datos del *datamart*. Fuente: elaboración propia. (2020).

1.28.2. Diseño del modelo dimensional

El diseño del modelo dimensional para la solución de IN propuesta en el presente documento abarca la identificación de las tablas de hechos y de las métricas numéricas que forman parte de esta, así como las dimensiones asociadas con los procesos operacionales de los sistemas de la organización.

Para la identificación de estas tablas de hechos y sus métricas, primero se debe analizar cuáles fueron los indicadores clave de desempeño señalados como de mayor urgencia. En la **Tabla 23**, se enlistan los indicadores que fueron calificados con un nivel de prioridad alto en la Fase II: Propuesta de indicadores clave de desempeño del capítulo anterior.

Tabla 23. Indicadores clave de desempeño con mayor nivel de prioridad

Identificador del indicador	Nombre del indicador	Sistema	Nivel de prioridad
IND-11	Cantidad de depósitos por categoría de cliente.	Procesador de criptomonedas (SB)	Alto
IND-12	Cantidad de retiros por categoría de cliente.	Procesador de criptomonedas (SB)	Alto
IND-13	Cantidad de transacciones por categoría de cliente.	Procesador de criptomonedas (SB)	Alto
IND-14	Monto de dinero de depósitos por categoría de cliente.	Procesador de criptomonedas (SB)	Alto
IND-15	Monto de dinero de retiros por categoría de cliente.	Procesador de criptomonedas (SB)	Alto
IND-21	Cantidad promedio pagada por comisión de mineros.	Procesador de criptomonedas (SB)	Alto

Identificador del indicador	Nombre del indicador	Sistema	Nivel de prioridad
IND-27	Monto de las transacciones de compra de dinero cripto procesadas	Monedero electrónico de criptomonedas (CD)	Alto
IND-28	Cantidad de transacciones de compra dinero cripto que han sido procesadas	Monedero electrónico de criptomonedas (CD)	Alto
IND-29	Monto promedio de las transacciones de compra de dinero cripto	Monedero electrónico de criptomonedas (CD)	Alto

Nota. Fuente: elaboración propia. (2020).

Los anteriores identificadores clave de desempeño permitieron la definición de las tablas de hechos necesarias y sus principales métricas. Seguidamente, en la **Tabla 24** se presentan las tablas de hechos y las métricas identificadas con respecto a los indicadores analizados.

Tabla 24. Tablas de hechos y sus métricas

Nombre del indicador	Tabla de hecho	Métricas
Procesador de criptomonedas	<i>SB Transaction</i>	Monto de dinero
		Comisión de mineros del <i>blockchain</i>
Monedero electrónico de criptomonedas	<i>CD Transaction</i>	Monto de dinero de los procesadores de tarjeta de crédito.
		Monto de dinero de las transacciones

Nota. Fuente: elaboración propia. (2020).

1.28.2.1. Diseño de alto nivel

Una vez identificadas las tablas de hechos y sus respectivas dimensiones, fue posible llevar a cabo un diseño de alto nivel sobre el modelo estrella.

La **Figura 14**, muestra el modelo estrella para la tabla de hechos de las transacciones del sistema de procesamiento de criptomonedas; por otro lado, la **Figura 15** presenta el modelo estrella de alto nivel para la tabla de transacciones del sistema de monedero electrónico de criptomonedas. Para ambas figuras, los nombres de las dimensiones y de la tabla de hechos como tal están escritas en inglés por cómo se definieron dentro del *datamart*.

Figura 14. Modelo estrella de alto nivel sobre transacciones del sistema de procesamiento de criptomonedas



Figura 14. Modelo dimensional de alto nivel sobre el modelo estrella de las transacciones del procesador de criptomonedas. Fuente: elaboración propia. (2020).

Figura 15. Modelo estrella de alto nivel para las transacciones del sistema de monedero electrónico de criptomonedas

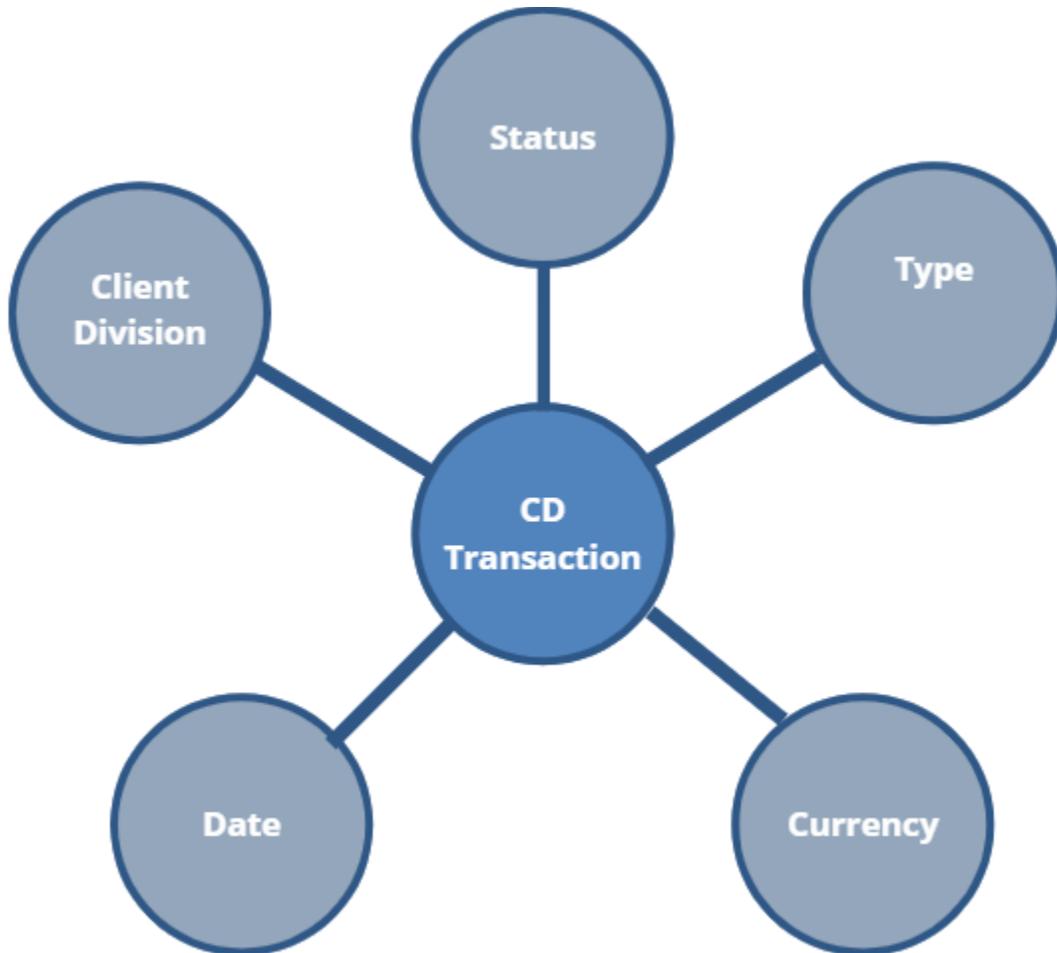


Figura 15. Modelo dimensional de alto nivel sobre el modelo estrella de las transacciones del procesador de criptomonedas. Fuente: elaboración propia. (2020).

1.28.3. Diseño físico

Una vez definidos los modelos estrella para las tablas de hechos, se procedió a llevar a cabo el modelado físico del almacén de datos.

Para estos diseños se siguió la convención existente en las bases de datos de los sistemas de la organización: nombres de tablas escritas en convención *Snake Case*, mientras que los nombres de las columnas de las tablas con la convención *Camel Case*.

1.28.3.1. Modelado dimensional

La **Figura 16**, muestra el modelo físico del modelo estrella propuesto para la tabla de hechos de transacciones del sistema de procesamiento de criptomonedas; por otro lado, la **Figura 17** presenta el diseño correspondiente a la tabla de hechos de transacciones del monedero electrónico de criptomonedas. Es importante resaltar que fueron incluidas algunas otras métricas a las propuestas anteriormente, que van a ser relevantes para la creación futura de visualizaciones de algunos otros indicadores.

Figura 16. Diseño físico del modelo estrella de transacciones del procesador de criptomonedas



Figura 16. Diseño físico del modelo estrella de transacciones del procesador de criptomonedas. Fuente: elaboración propia. (2020).

Figura 17. Diseño físico del modelo estrella de transacciones del monedero de criptomonedas

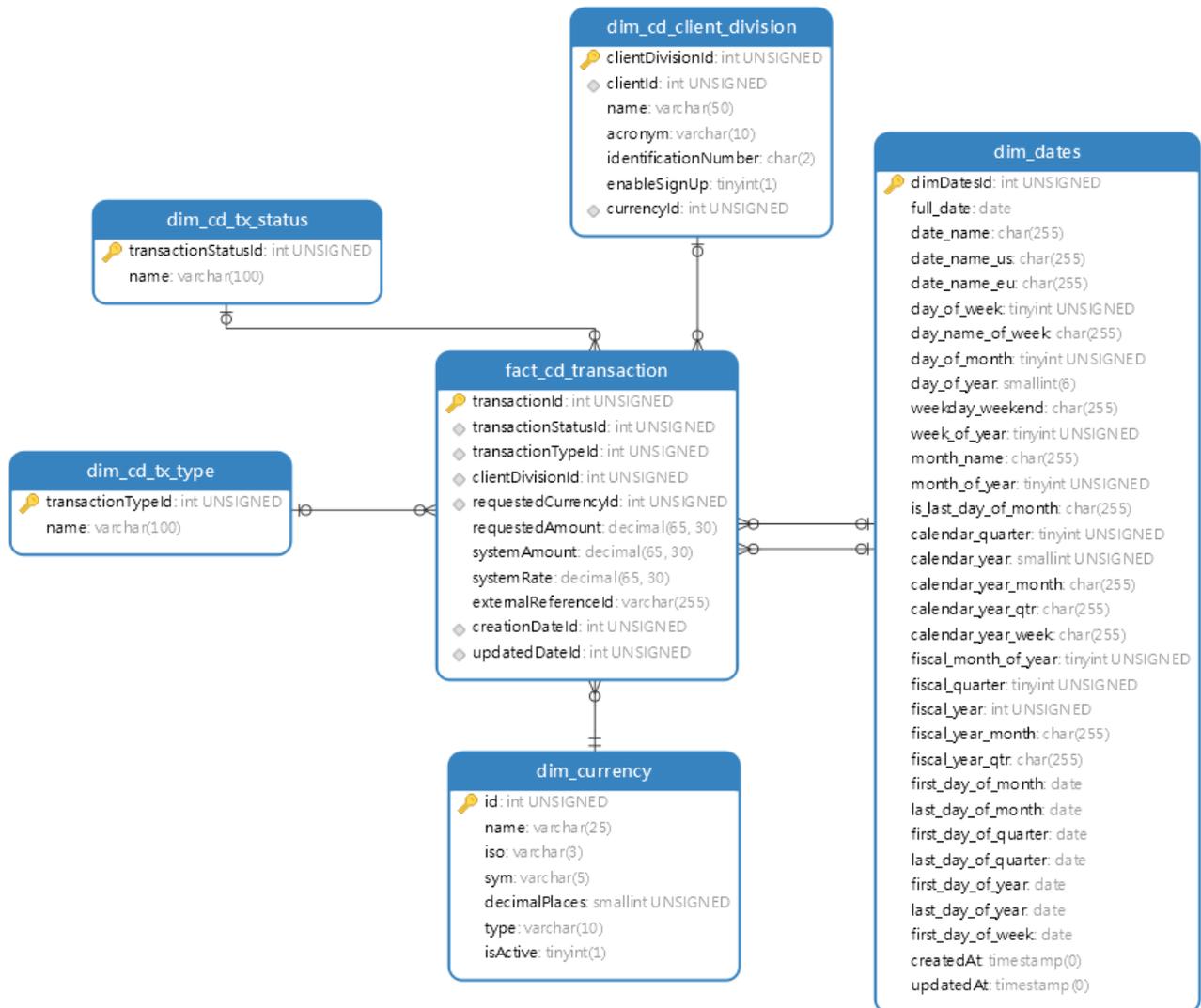


Figura 18. Tablas del área de trabajo. Fuente: elaboración propia. (2020).

1.28.3.2. Staging area o área de trabajo

Una vez debidamente definidos los diseños físicos del modelo dimensional, es posible identificar las tablas a incluir dentro del área de trabajo. La principal función de esta área de trabajo es alojar los datos crudos de las bases de datos operacionales de donde procede la información que necesita ser transferida al *datamart*. La **Figura 6** detalla el flujo de la información a través de esta área de trabajo.

Las tablas creadas para alojar esta información se muestran en la **Figura 18**. Tablas del área de trabajo., estas son una copia de las tablas originales de las bases de datos operacionales donde se almacena la información correspondiente a las dimensiones y las tablas de hechos definidas anteriormente, pero no cuentan con ningún tipo de índices o llaves, sino que solamente contiene los datos “crudos” dentro de las columnas de las diferentes tablas.

Figura 18. Tablas del área de trabajo.

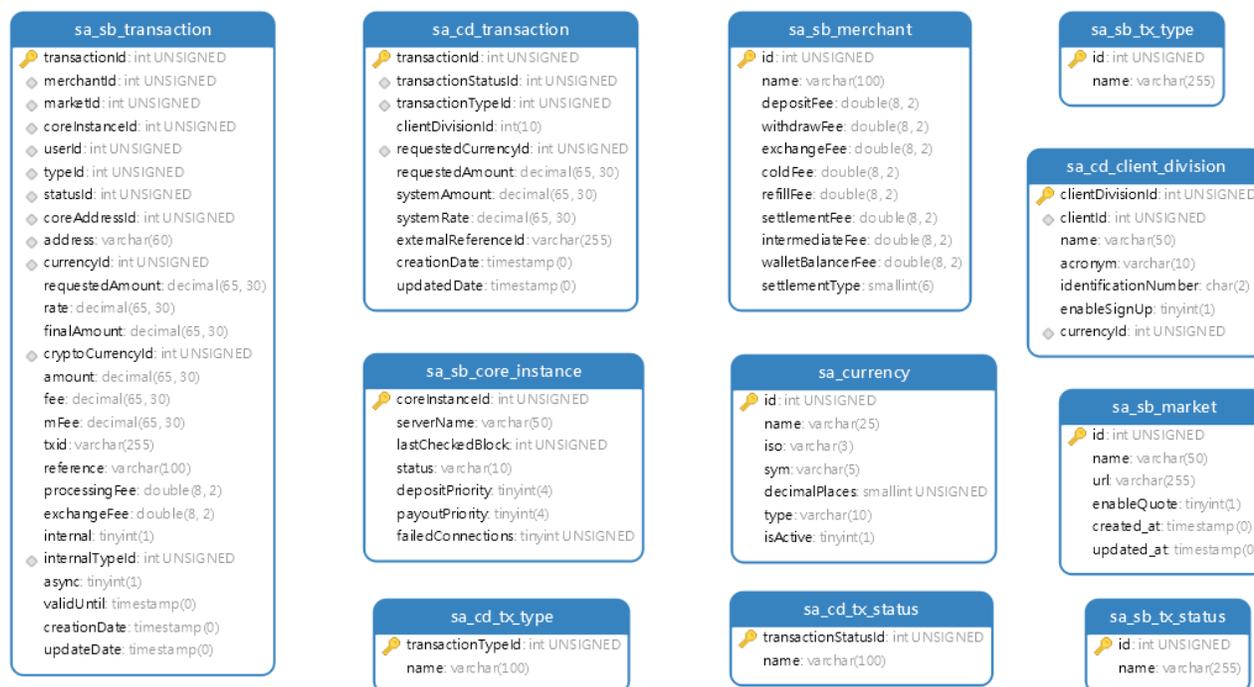


Figura 18. Tablas del área de trabajo. Fuente: elaboración propia. (2020).

1.28.4. Implementación

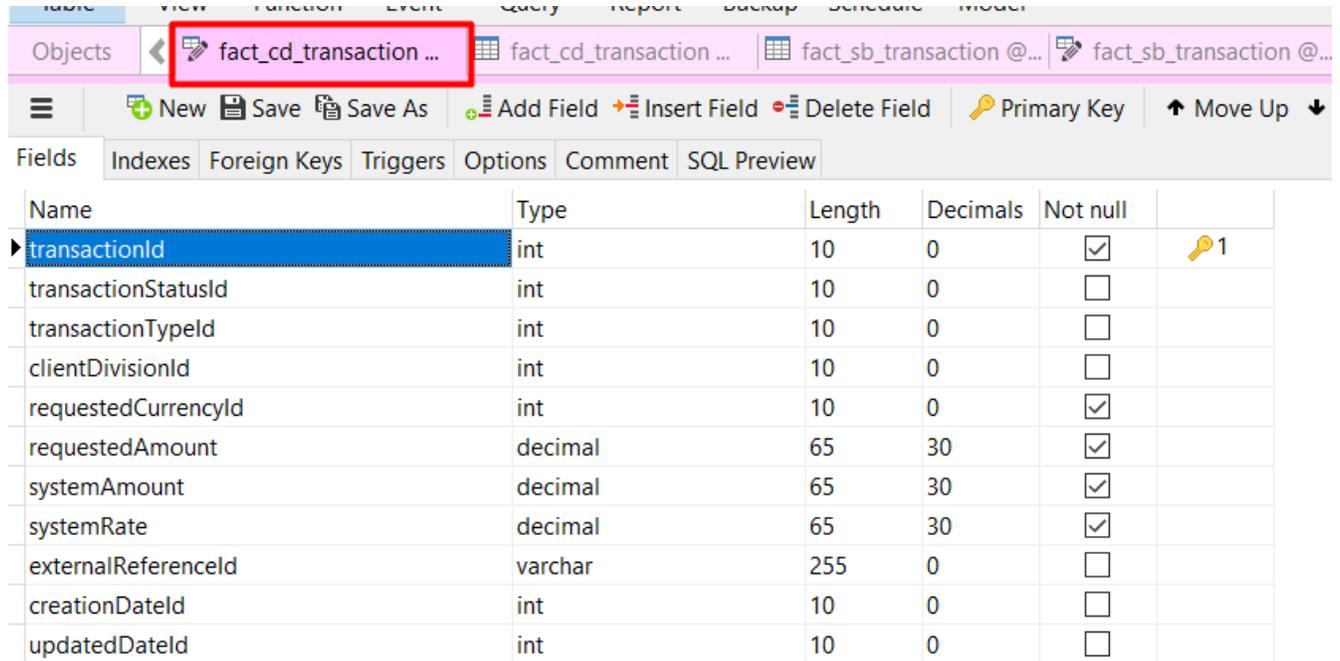
Se llevó a cabo la configuración de las tablas, siguiendo lo establecido en el modelo físico definido en la sección de Diseño físico. La **Figura 19**, muestra la definición de los tipos largos, cantidad de decimales, anulabilidad y la llave primaria de la tabla de hechos de transacción para el sistema de procesamiento de criptomonedas; de manera similar, la **Figura 20** muestra esta información para la tabla de hechos de transacción para el sistema de monedero electrónico de criptomonedas. De esta manera fueron definidas todas las tablas de dimensiones especificadas.

Figura 19. Definición de la tabla de hechos de transacciones del procesador de criptomonedas

Name	Type	Length	Decimals	Not null	
transactionId	int	10	0	<input checked="" type="checkbox"/>	1
merchantId	int	10	0	<input checked="" type="checkbox"/>	
marketId	int	10	0	<input checked="" type="checkbox"/>	
coreInstanceId	int	10	0	<input type="checkbox"/>	
userId	int	10	0	<input checked="" type="checkbox"/>	
typeId	int	10	0	<input checked="" type="checkbox"/>	
statusId	int	10	0	<input checked="" type="checkbox"/>	
coreAddressId	int	10	0	<input type="checkbox"/>	
address	varchar	60	0	<input type="checkbox"/>	
currencyId	int	10	0	<input checked="" type="checkbox"/>	
requestedAmount	decimal	65	30	<input checked="" type="checkbox"/>	
rate	decimal	65	30	<input checked="" type="checkbox"/>	
finalAmount	decimal	65	30	<input checked="" type="checkbox"/>	
cryptoCurrencyId	int	10	0	<input checked="" type="checkbox"/>	
amount	decimal	65	30	<input checked="" type="checkbox"/>	
fee	decimal	65	30	<input checked="" type="checkbox"/>	
mFee	decimal	65	30	<input checked="" type="checkbox"/>	
txid	varchar	255	0	<input type="checkbox"/>	
reference	varchar	100	0	<input type="checkbox"/>	
processingFee	double	8	2	<input checked="" type="checkbox"/>	
exchangeFee	double	8	2	<input checked="" type="checkbox"/>	
internal	tinyint	1	0	<input checked="" type="checkbox"/>	
internalTypeId	int	10	0	<input type="checkbox"/>	
async	tinyint	1	0	<input checked="" type="checkbox"/>	
validUntil	timestamp	0	0	<input checked="" type="checkbox"/>	
creationDateId	int	11	0	<input type="checkbox"/>	
updatedDateId	int	11	0	<input type="checkbox"/>	

Figura 19. Definición de la tabla de hechos de transacciones del procesador de criptomonedas. Fuente: elaboración propia. (2020).

Figura 20. Definición de la tabla de hechos de transacciones del monedero de criptomonedas



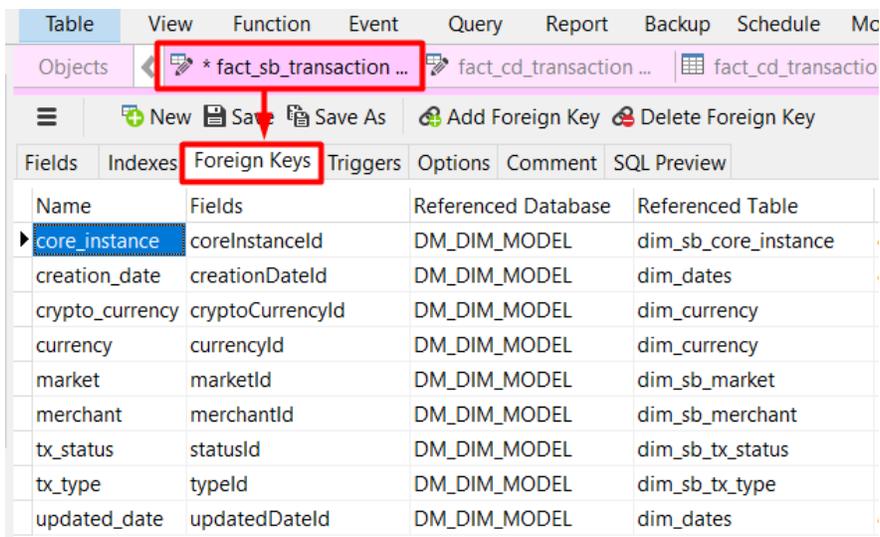
The screenshot shows a database design tool interface. The top toolbar includes options like 'New', 'Save', 'Save As', 'Add Field', 'Insert Field', 'Delete Field', 'Primary Key', and 'Move Up/Down'. Below the toolbar, there are tabs for 'Fields', 'Indexes', 'Foreign Keys', 'Triggers', 'Options', 'Comment', and 'SQL Preview'. The 'Fields' tab is active, displaying a table with the following columns: Name, Type, Length, Decimals, Not null, and a key icon. The 'transactionId' field is selected and highlighted in blue, and a red box highlights the 'fact_cd_transaction ...' tab at the top.

Name	Type	Length	Decimals	Not null	Key
transactionId	int	10	0	<input checked="" type="checkbox"/>	1
transactionStatusId	int	10	0	<input type="checkbox"/>	
transactionTypeId	int	10	0	<input type="checkbox"/>	
clientDivisionId	int	10	0	<input type="checkbox"/>	
requestedCurrencyId	int	10	0	<input checked="" type="checkbox"/>	
requestedAmount	decimal	65	30	<input checked="" type="checkbox"/>	
systemAmount	decimal	65	30	<input checked="" type="checkbox"/>	
systemRate	decimal	65	30	<input checked="" type="checkbox"/>	
externalReferencId	varchar	255	0	<input type="checkbox"/>	
creationDateId	int	10	0	<input type="checkbox"/>	
updatedDateId	int	10	0	<input type="checkbox"/>	

Figura 20. Definición de la tabla de hechos de transacción del monedero de criptomonedas. Fuente: elaboración propia. (2020).

Por otro lado, la relación por medio de llaves foráneas entre las dimensiones y las tablas de hechos es configurada tal como se muestra en la **Figura 21**; y en la **Figura 22**.

Figura 21. Definición de llaves foráneas entre la tabla de hechos de transacción del procesador de criptomonedas y sus dimensiones

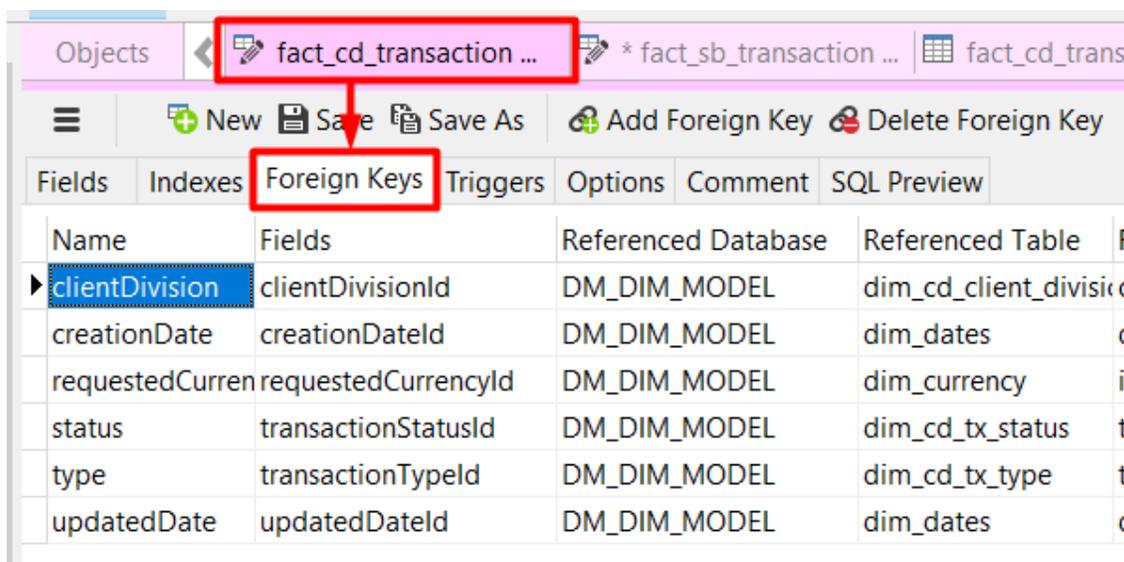


The screenshot shows the 'Foreign Keys' tab in SQL Server Enterprise Manager for the table 'fact_sb_transaction'. The table is highlighted in the Objects pane. The 'Foreign Keys' tab is also highlighted. The table lists the following foreign key relationships:

Name	Fields	Referenced Database	Referenced Table
core_instance	coreInstanceId	DM_DIM_MODEL	dim_sb_core_instance
creation_date	creationDateId	DM_DIM_MODEL	dim_dates
crypto_currency	cryptoCurrencyId	DM_DIM_MODEL	dim_currency
currency	currencyId	DM_DIM_MODEL	dim_currency
market	marketId	DM_DIM_MODEL	dim_sb_market
merchant	merchantId	DM_DIM_MODEL	dim_sb_merchant
tx_status	statusId	DM_DIM_MODEL	dim_sb_tx_status
tx_type	typeId	DM_DIM_MODEL	dim_sb_tx_type
updated_date	updatedDateId	DM_DIM_MODEL	dim_dates

Figura 21. Definición de llaves foráneas entre la tabla de hechos de transacción del procesador de criptomonedas y sus dimensiones. Fuente: elaboración propia. (2020).

Figura 22. Definición de llaves foráneas entre la tabla de hechos de transacción del monedero de criptomonedas y sus dimensiones



The screenshot shows the 'Foreign Keys' tab in SQL Server Enterprise Manager for the table 'fact_cd_transaction'. The table is highlighted in the Objects pane. The 'Foreign Keys' tab is also highlighted. The table lists the following foreign key relationships:

Name	Fields	Referenced Database	Referenced Table
clientDivision	clientDivisionId	DM_DIM_MODEL	dim_cd_client_divisio
creationDate	creationDateId	DM_DIM_MODEL	dim_dates
requestedCurrency	requestedCurrencyId	DM_DIM_MODEL	dim_currency
status	transactionStatusId	DM_DIM_MODEL	dim_cd_tx_status
type	transactionTypeId	DM_DIM_MODEL	dim_cd_tx_type
updatedDate	updatedDateId	DM_DIM_MODEL	dim_dates

Figura 22. Definición de llaves foráneas entre la tabla de hechos de transacción del monedero de criptomonedas y sus dimensiones. Fuente: elaboración propia. (2020).

1.29. Fase IV: Diseño de los procesos *ETL*

En esta sección se abarca el desarrollo de los procesos *ETL*, desde la definición de la arquitectura técnica, hasta la implementación y puesta en producción. Todo lo anterior con base en los requerimientos definidos en la sección de Identificación de requerimientos relacionados con el diseño de los procesos *ETL*.

1.29.1. Diseño de la arquitectura técnica

En esta sección se describe los aspectos técnicos que soportaron el desarrollo la implementación de los procesos *ETL*.

1.29.1.1. Ambiente de ejecución de los *Jobs*

Para configurar la ejecución automática de los procesos *ETL* se utiliza el servidor de *Linux Ubuntu* en los servidores *AWS* de la empresa. Este servidor tiene la labor de calendarizar y ejecutar los flujos de extracción, transformación y carga de datos. La **Figura 23** detalla algunas características de la instancia configurada por los ingenieros *devops*.

Figura 23. Características de la instancia configurada por los ingenieros *devops*

Product Details		
Name	vCPUs	Memory (GiB)
t3.medium	2	4.0

Figura 23. Características de la instancia configurada por los ingenieros *devops*. Fuente: elaboración propia. (2020).

Para ingresar a esta instancia es necesaria la validación de la dirección IP desde donde se realiza la petición de acceso. En este caso, la dirección IP incluida en el *white list* de la instancia es la de la VPN de desarrollo, para el acceso a la calendarización y ejecución de los *Jobs*. En la **Figura 24**, se muestra una conexión fallida al no contar con conexión a la VPN, mientras que la **Figura 25** presenta el resultado de una conexión exitosa al servidor con la VPN activa.

Figura 24. Conexión fallida al no contar con conexión a la VPN

```
josea@MSI C:\Users\josea\Documents\BI PROJECT\Pem
$ "C:/Program Files/Git/bin/sh.exe" --login -i -- "/c/Users/josea/Documents/BI PROJECT\Pem/etl.sh"
ssh: connect to host 10.0.3.213 port 22: Connection timed out
```

Figura 24. Conexión fallida al no contar con conexión a la VPN. Fuente: elaboración propia. (2020).

Figura 25. Conexión exitosa al servidor con la VPN activa

The image shows two overlapping windows. On the left is a terminal window with the following content:

```
josea@MSI C:\Users\josea\Documents\BI PROJECT\Pem
$ etl
josea@MSI C:\Users\josea\Documents\BI PROJECT\Pem
$ cd C:\Users\josea\Documents\BI PROJECT\Pem
josea@MSI C:\Users\josea\Documents\BI PROJECT\Pem
$ "C:/Program Files/Git/bin/sh.exe" --login -i -- "/c/Users/josea/Documents/BI PROJECT\Pem/etl.sh"
Welcome to Ubuntu 18.04.2 LTS (GNU/Linux 5.3.0-1017-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

System information as of Fri Jul 3 03:08:02 UTC 2020

System load:  0.0          Processes:      97
Usage of /:   22.3% of 14.48GB  Users logged in:  0
Memory usage: 6%          IP address for eth0: 10.0.3.213
Swap usage:  0%

 * "If you've been waiting for the perfect Kubernetes dev solution for
macOS, the wait is over. Learn how to install Microk8s on macOS."

https://www.techrepublic.com/article/how-to-install-microk8s-on-macos/

 * Canonical Livepatch is available for installation.
- Reduce system reboots and improve kernel security. Activate at:
https://ubuntu.com/livepatch

72 packages can be updated.
0 updates are security updates.

*** System restart required ***
root@ip-10-0-3-213:~#
```

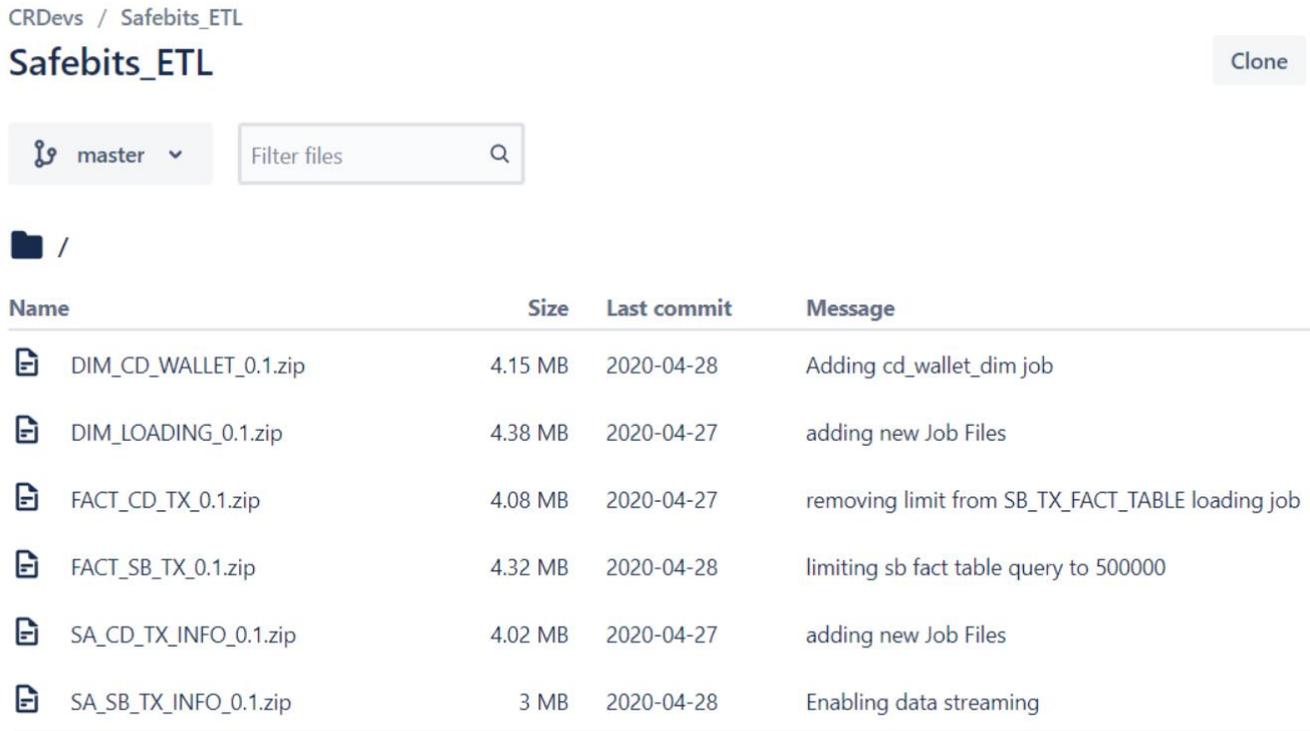
On the right is the OpenVPN Connect application window, titled "OpenVPN Connect Profiles". It shows a "CONNECTED" status for the "OpenVPN Profile" (itops-systems.ionblock.io [jbarboza]). The "CONNECTION STATS" section displays a graph of data transfer over time, with "2.9KB/s" indicated. Below the graph, it shows "BYTES IN 0 KB/S" and "BYTES OUT 0 KB/S". The "DURATION" is "00:00:37" and "PACKET RECEIVED" is "3 sec ago". At the bottom, it says "YOU" and "YOUR PRIVATE IP" with a plus sign icon.

Figura 25. Conexión exitosa al servidor con la VPN activa. Fuente: elaboración propia. (2020).

1.29.1.2. Manejo de versiones

Para el manejo de versiones de los *Jobs* de *ETL*, se utilizó un repositorio basado en *git* llamado *Bitbucket*, el cual es la herramienta actualmente utilizada por la organización para el manejo de los repositorios de otros de los sistemas. En la **Figura 26**, se muestra el repositorio de los *Jobs ETL*, donde se encuentran algunos de los archivos exportados desde la herramienta de desarrollo y demás información con respecto a su carga en dicho repositorio.

Figura 26. Repositorio de los Jobs ETL dentro de *Bitbucket*



Name	Size	Last commit	Message
 DIM_CD_WALLET_0.1.zip	4.15 MB	2020-04-28	Adding cd_wallet_dim job
 DIM_LOADING_0.1.zip	4.38 MB	2020-04-27	adding new Job Files
 FACT_CD_TX_0.1.zip	4.08 MB	2020-04-27	removing limit from SB_TX_FACT_TABLE loading job
 FACT_SB_TX_0.1.zip	4.32 MB	2020-04-28	limiting sb fact table query to 500000
 SA_CD_TX_INFO_0.1.zip	4.02 MB	2020-04-27	adding new Job Files
 SA_SB_TX_INFO_0.1.zip	3 MB	2020-04-28	Enabling data streaming

Figura 26. Repositorio de los *Jobs ETL*. Fuente: elaboración propia. (2020).

1.29.2. Diseño de procesos *ETL*

En esta sección se muestra el diseño y desarrollo de los *Jobs* donde se realiza la extracción, transformación y carga de los datos desde las bases de datos operacionales, hasta el modelo dimensional del almacén de datos. En la **Figura 27**, se muestran las conexiones a las bases de datos configuradas para extraer o cargar los datos.

Figura 27. Conexiones a las bases de datos configuradas para extraer o cargar los datos

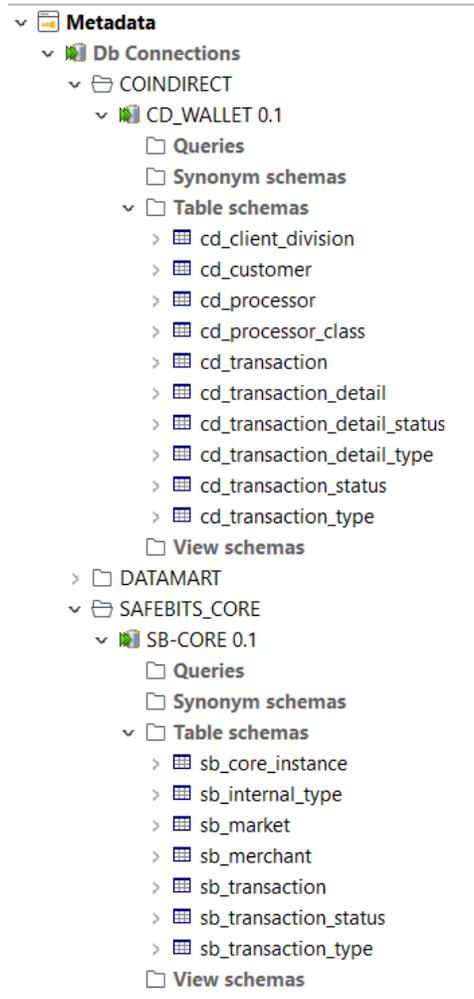


Figura 27. Conexiones a las bases de datos configuradas para extraer o cargar los datos. Fuente: elaboración propia. (2020).

1.29.2.1. Diseño de los Jobs de carga del Staging Area

Estos *Jobs* son los encargados de mover los datos desde las fuentes de datos, hasta el área de trabajo o *staging area*.

En la **Figura 28** es posible observar el *job* encargado de cargar los datos correspondientes a la tabla de transacción para el sistema de procesamiento de criptomonedas. Este proceso cuenta con tres componentes, primero se encuentra un componente correspondiente a la entrada de datos de la tabla originaria de la base de datos operacional del procesador, luego se encuentra un componente

encargado de mapear las columnas entre las tablas de origen y destino, y por último se encuentra la conexión a la tabla destino en el área de trabajo.

Figura 28. Job encargado de cargar los datos correspondientes a la tabla de transacción para el sistema de procesamiento de criptomonedas

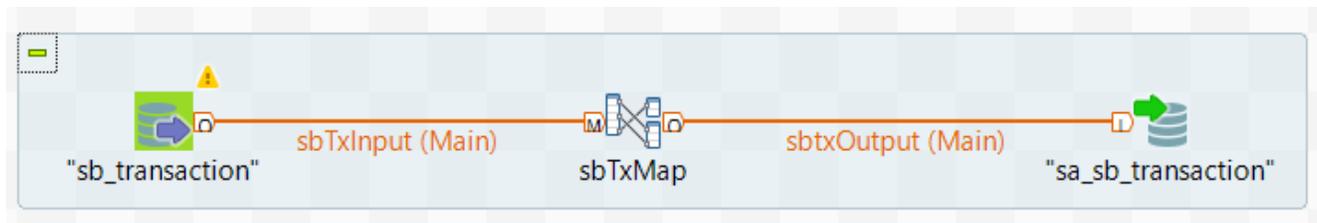


Figura 28. Job encargado de cargar los datos correspondientes a la tabla de transacción para el sistema de procesamiento de criptomonedas. Fuente: elaboración propia. (2020).

Seguidamente, la **Figura 29** muestra el detalle del componente *sbTxMap*, donde las columnas entre la tabla de entrada y la tabla de salida son directamente relacionadas.

Figura 29. Detalle del componente *sbTxMap*

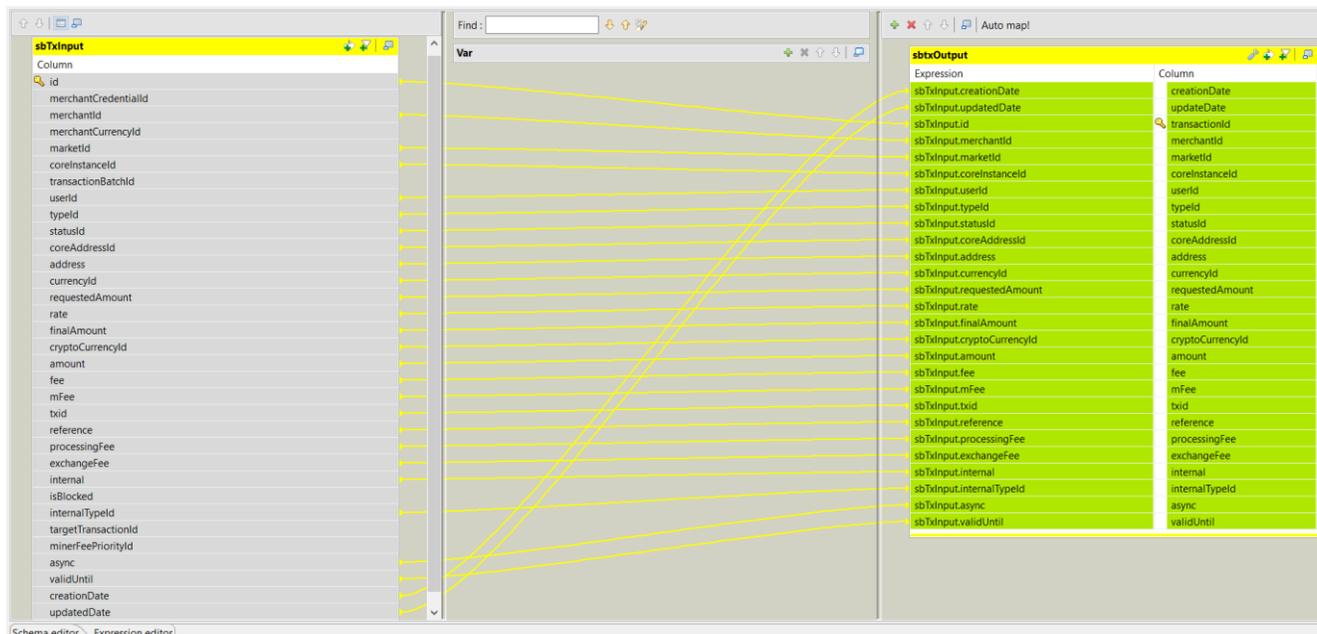


Figura 29. Detalle del componente *sbTxMap*. Fuente: elaboración propia. (2020).

El proceso anteriormente definido es el abarcado para la carga de las demás tablas del área de trabajo: el desarrollo de un *Job*, con un componente relacionado con la fuente de los datos, uno vinculado a la salida de la información, y otro donde se realiza el mapeo entre las columnas de ambas tablas.

1.29.2.2. Diseño de los *Jobs* de carga del almacén de datos

Una vez definidos los procesos para la carga del área de trabajo, se procedió al desarrollo de los *Jobs* para la extracción desde el área de trabajo, la transformación, y la carga de los datos en el modelo dimensional del almacén de datos.

En la **Figura 30**, se muestra la carga de una de las dimensiones del almacén de datos, en este caso, de los tipos de estatus de las transacciones del monedero electrónico de criptomonedas. En este *Job* se toma directamente los datos de la table del área de trabajo, y se mapean a la correspondiente tabla de la dimensión en el almacén de datos.

Figura 30. Carga de una de las dimensiones del almacén de datos

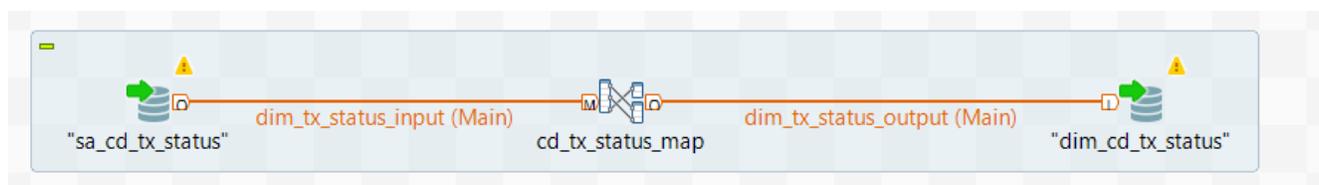


Figura 30. Carga de una de las dimensiones del almacén de datos. Fuente: elaboración propia. (2020).

En la **Figura 31**, se observa cómo la carga de los datos de la tabla de estatus es directa, no tiene ningún proceso de transformación, este es el caso de todos los *Jobs* para la carga de los datos de las dimensiones del modelo dimensional.

Figura 31. Carga de los datos de la tabla de estatus directa



Figura 31. Carga de los datos de la tabla de estatus directa. Fuente: elaboración propia. (2020).

Por otro lado, la carga de las tablas de hecho conlleva mayor complejidad que la carga de las tablas dimensionales, debido principalmente a que estas cuentan con la relación foránea entre ellas y sus respectivas dimensiones. En la

Figura 32, se muestra el *job* para la carga de la tabla de hechos de transacción correspondiente al sistema de monedero de criptomonedas, en esta se detalla las fuentes de la información de las diferentes tablas del área de trabajo, así como el mapeo de los datos y la salida hacia la carga de la tabla encontrada en el modelo dimensional del *datamart*.

Figura 32. Job para la carga de la tabla de hechos de transacción correspondiente al sistema de monedero de criptomonedas

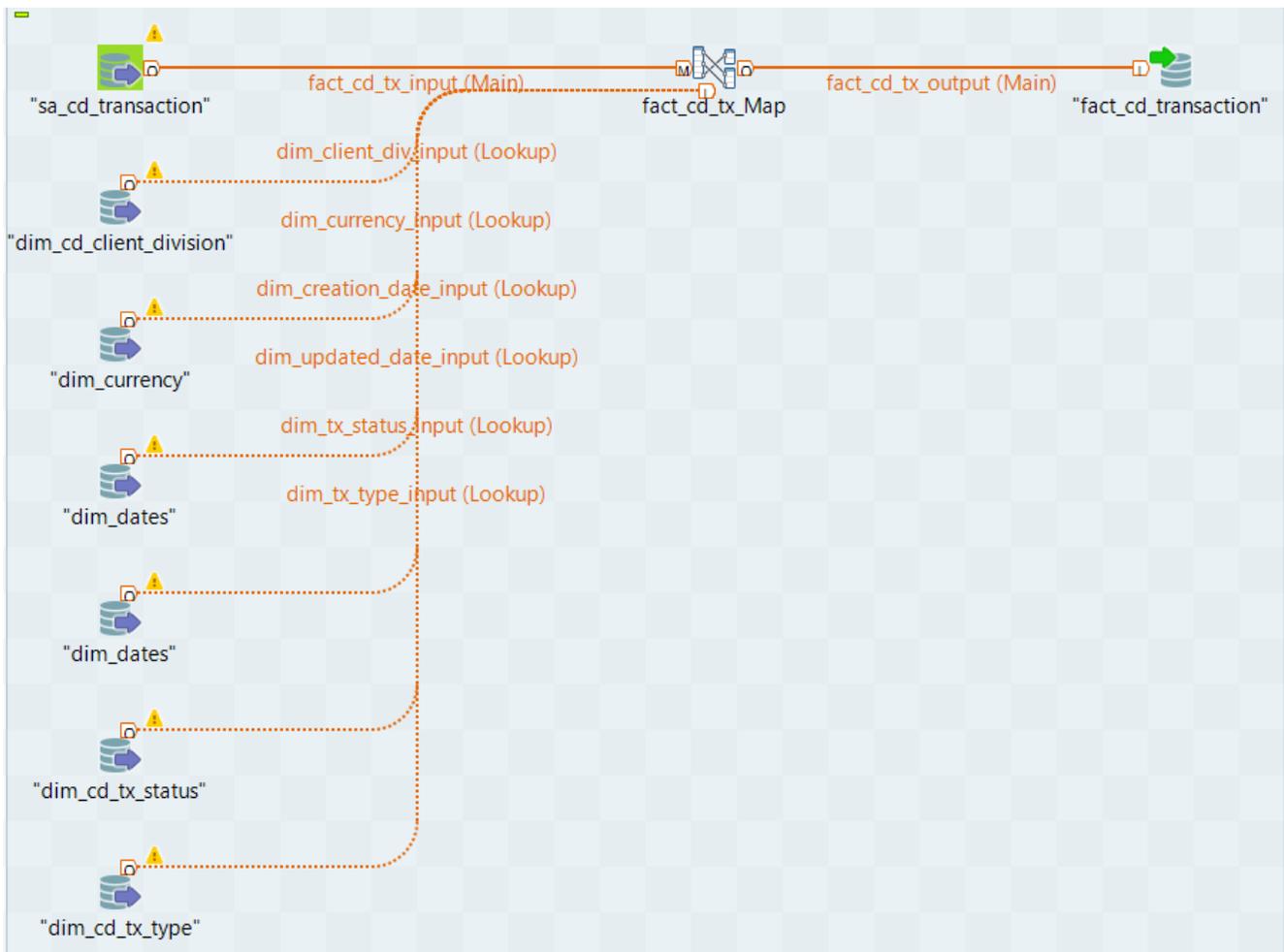


Figura 32. Job para la carga de la tabla de hechos de transacción correspondiente al sistema de monedero de criptomonedas. Fuente: elaboración propia. (2020).

En la **Figura 33**, se muestra la existencia de las relaciones *join* llevados a cabo para la carga de la información, así como el mapeo de los datos hacia la tabla destino.

Figura 33. Mapeo de los datos de la tabla fuente a la tabla de origen

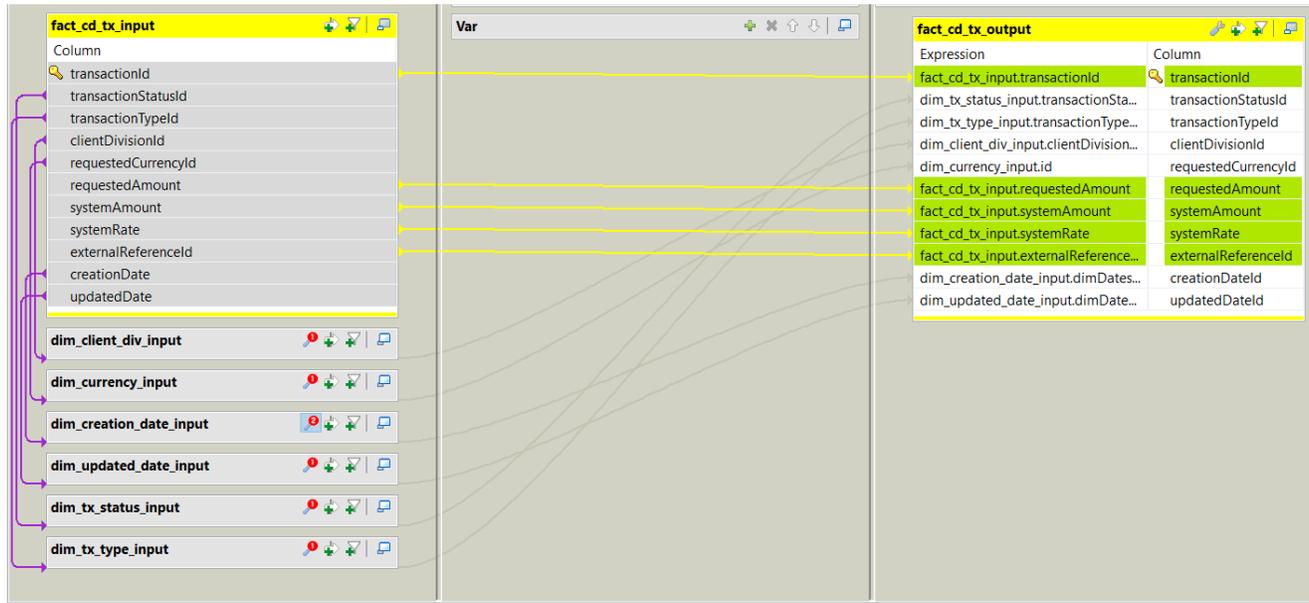


Figura 33. Mapeo de los datos de la tabla fuente a la tabla de origen Fuente: elaboración propia. (2020).

La carga de la tabla de hechos de transacción del procesador de criptomonedas funciona de una manera similar a la del monedero. En la **Figura 34** se muestra el *job* diseñado, mientras que en la **Figura 35** se muestra el mapeo entre el origen de los datos y la tabla destino.

Figura 34. *Job* diseñado para la carga de información de la tabla de hechos de las transacciones del procesador de criptomonedas

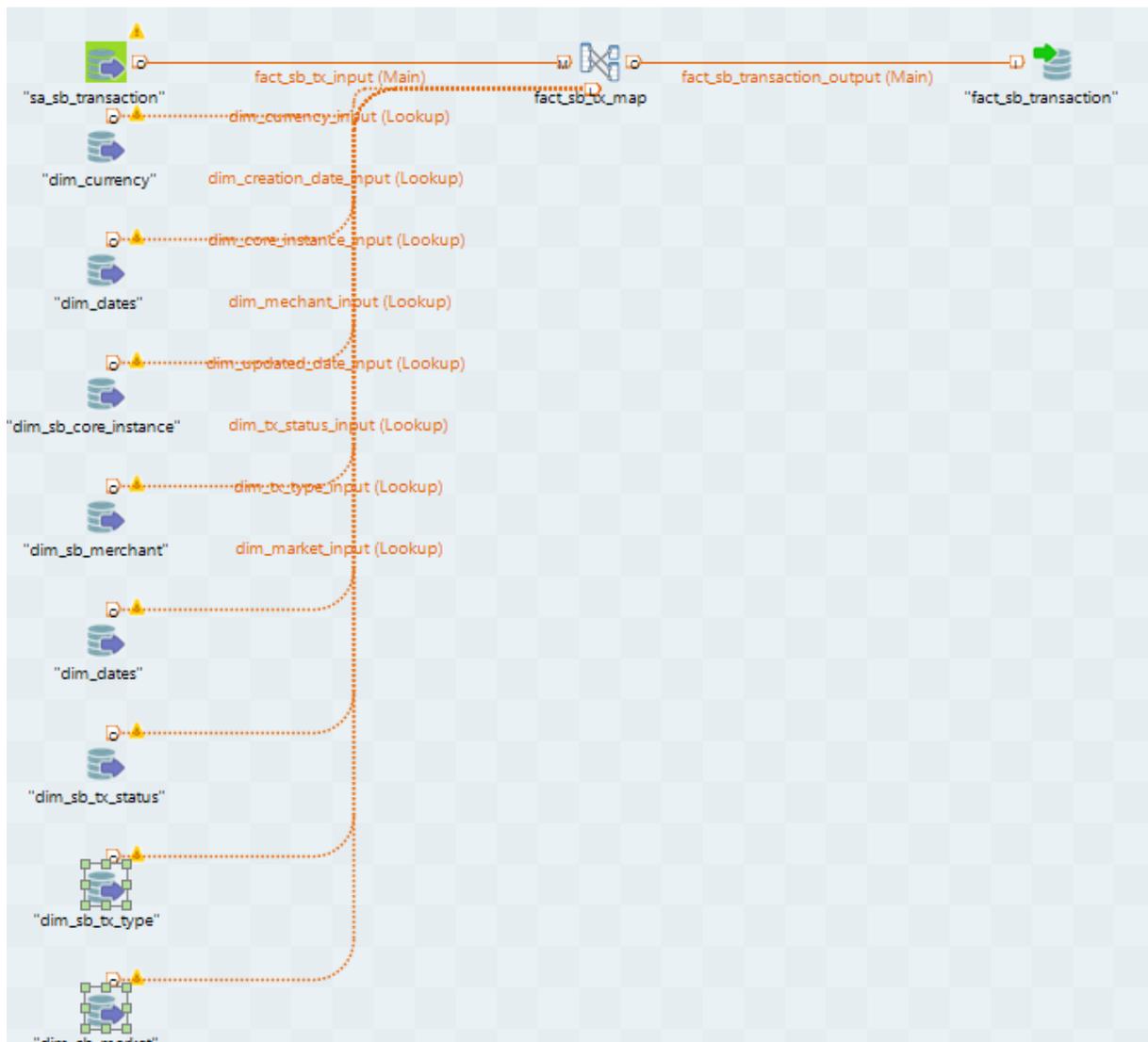


Figura 34. *Job* diseñado para la carga de información de la tabla de hechos de las transacciones del procesador de criptomonedas. Fuente: elaboración propia. (2020).

Figura 35. Mapeo entre el origen de los datos y la tabla destino

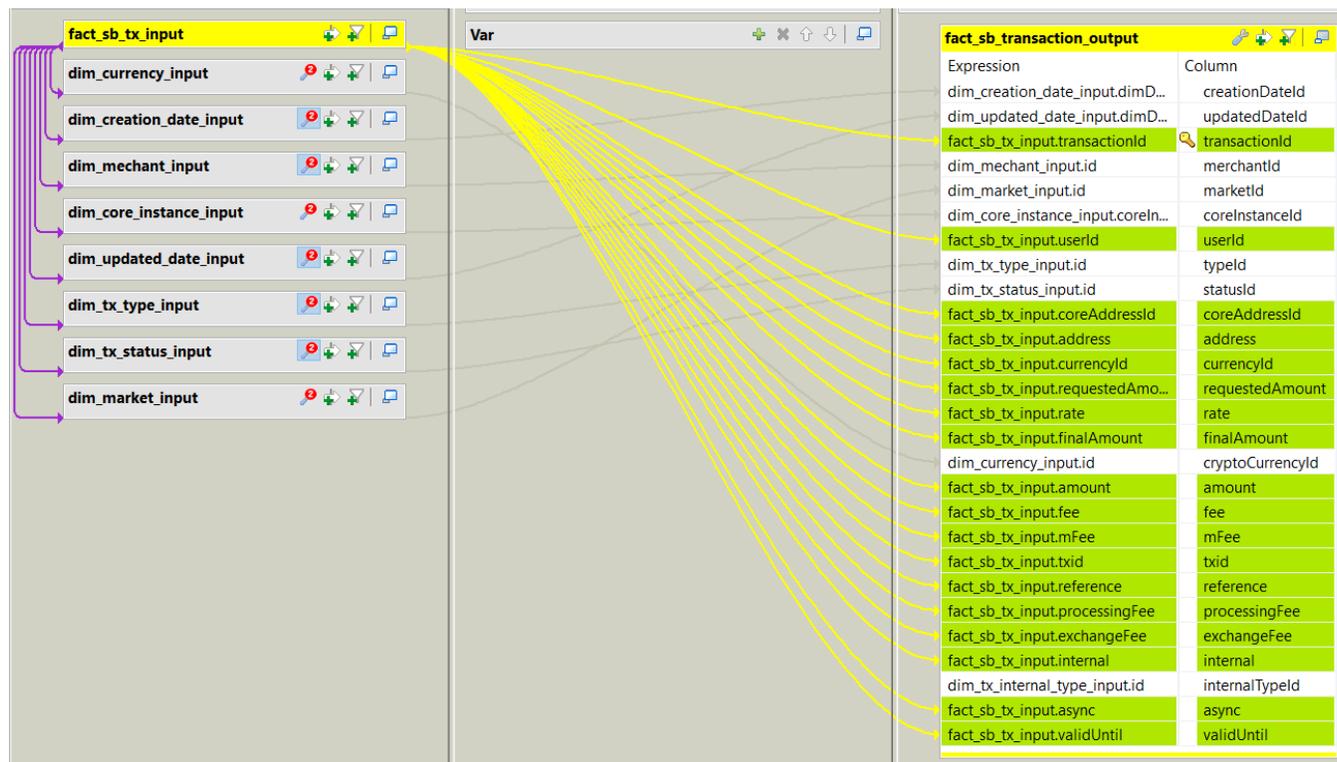


Figura 35. Mapeo entre el origen de los datos y la tabla destino. Fuente: elaboración propia. (2020).

1.29.3. Implementación de los procesos ETL

Una vez diseñados los procesos de carga, se procedió a ejecutar su implementación en el ambiente de producción.

1.29.3.1. Liberación y ejecución de Jobs

En la **Figura 36** se observa que se debe “construir los Jobs” desde dentro de la herramienta y efectuarlo de manera individual, esta acción abre una ventana nueva, mostrada en la **Figura 37**, donde se debe seleccionar la ubicación en donde se exportará el Job, en este caso, la carpeta seleccionada es un clon del repositorio en *Bitbucket*, donde se encuentran las versiones de todos los Jobs.

Figura 36. Opción de construcción de Jobs en Talend para su exportación

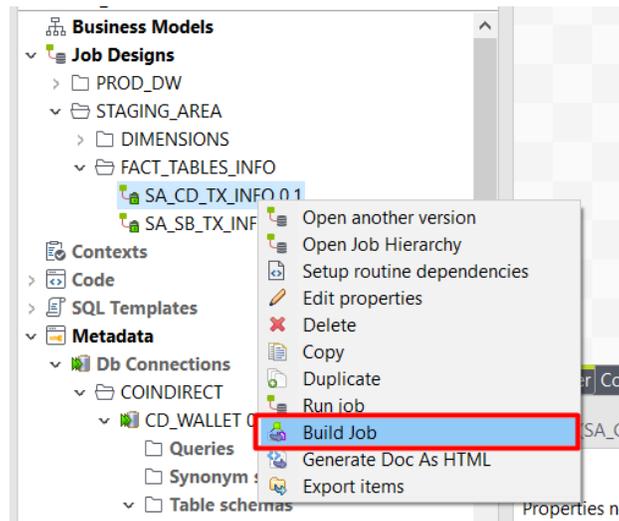


Figura 36. Opción de construcción de Jobs en Talend para su exportación. Fuente: elaboración propia. (2020).

Figura 37. Ventana para la exportación de los Jobs

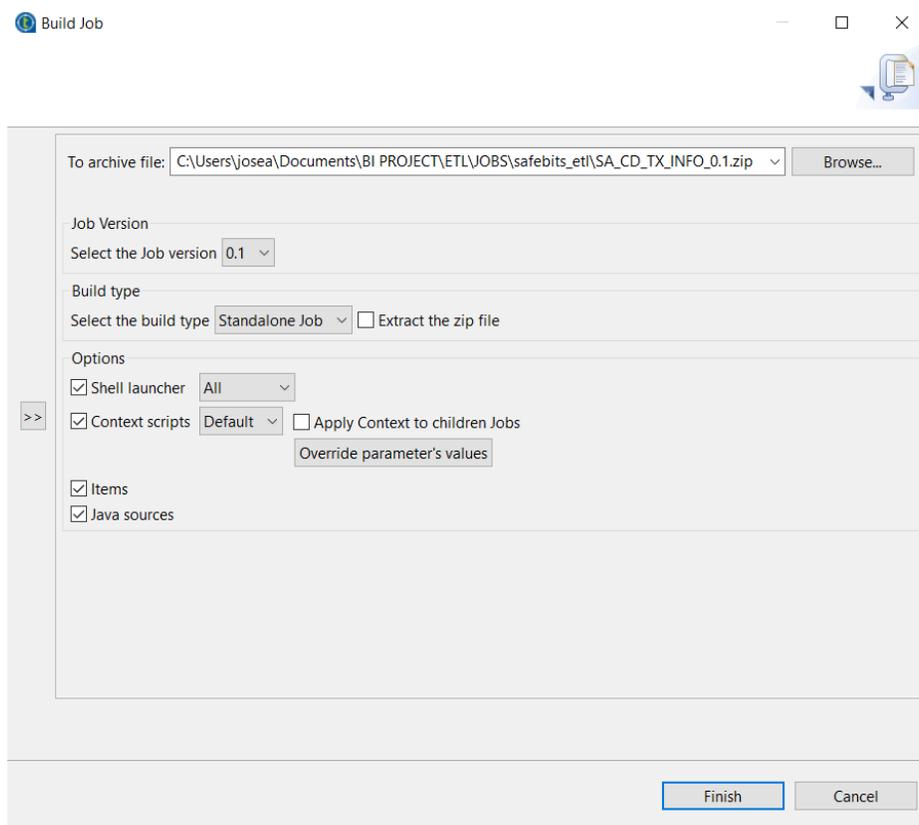


Figura 37. Ventana para la exportación de los Jobs. Fuente: elaboración propia. (2020).

En la **Figura 38**, se muestra el proceso para subir la nueva versión del *job* construido a través de la consola de una terminal de *git*.

Figura 38. Proceso para subir la nueva versión del *job* construido a través de la consola de una terminal de *git*

```
josea@MSI MINGW64 ~/Documents/BI PROJECT/ETL/JOBs/safebits_etl (master)
$ git add SA_CD_TX_INFO_0.1.zip

josea@MSI MINGW64 ~/Documents/BI PROJECT/ETL/JOBs/safebits_etl (master)
$ git commit -m "Uploading Job Changes"
[master f0b890d] Uploading Job Changes
 1 file changed, 0 insertions(+), 0 deletions(-)

josea@MSI MINGW64 ~/Documents/BI PROJECT/ETL/JOBs/safebits_etl (master)
$ git push
Enumerating objects: 5, done.
Counting objects: 100% (5/5), done.
Delta compression using up to 8 threads
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 277.46 KiB | 2.37 MiB/s, done.
Total 3 (delta 2), reused 0 (delta 0), pack-reused 0
To https://bitbucket.org/crdevs/safebits_etl.git
 a13c9f5..f0b890d master -> master

josea@MSI MINGW64 ~/Documents/BI PROJECT/ETL/JOBs/safebits_etl (master)
$ |
```

Figura 38. Proceso para subir la nueva versión del *job* construido a través de la consola de una terminal de *git*. Fuente: elaboración propia. (2020).

Dentro de *Jenkins*, otra de las plataformas tecnológicas de la organización, se simplificó la ejecución de un comando encargado de subir los archivos a la instancia *Linux*, configurada para la ejecución de los *Jobs*, luego de los cambios actualizados en el repositorio de *Bitbucket*. Dentro de las fases del comando se descomprimen los archivos exportados desde *Talend* y luego estos son ejecutados. Asimismo, los *Jobs* fueron actualizados en el proceso.

En la **Figura 39** se muestra la interfaz de *Jenkins*, desde donde se ejecuta el comando encargado de la carga o ejecución de los *Jobs*. Por otra parte, en la **Figura 40**, se muestra el directorio de los archivos dentro del servidor.

Figura 39. Interfaz de Jenkins

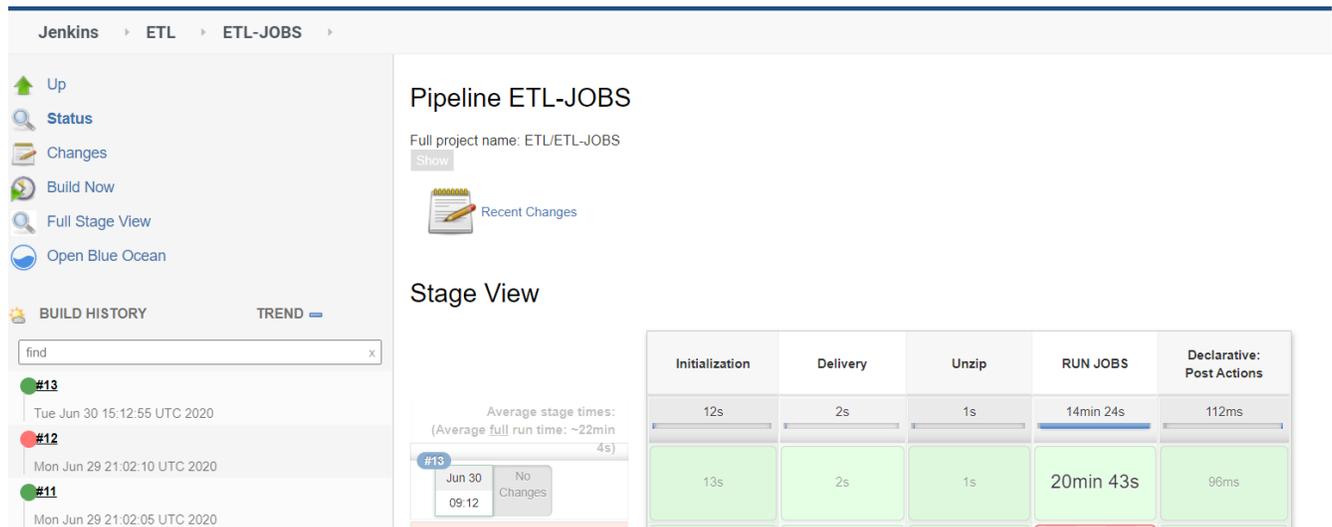


Figura 39. Interfaz de Jenkins. Fuente: elaboración propia. (2020).

Figura 40. Directorio de los archivos dentro del servidor

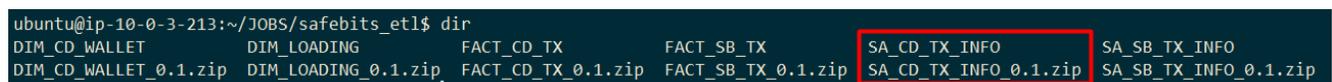


Figura 40. Directorio de los archivos dentro del servidor. Fuente: elaboración propia. (2020).

1.29.3.2. Carga inicial de los datos

Para la carga inicial de la información fue posible la ejecución directa de los *Jobs* del área de trabajo. Sin embargo, la carga de las tablas de hechos del modelo dimensional, como tal, requirió la carga de la información por segmentos de 450 000 filas, debido a que conjuntos de datos mayores requerirían mayor poder computacional en la instancia.

Para definir los segmentos de filas fue necesario realizar cambios en los *Jobs* de carga de las tablas de hechos, específicamente con la modificación de la consulta *MySQL*, que carga la información desde la tabla de origen en el área de trabajo, limitando los resultados a la cantidad especificada y definiendo el inicio del segmento de datos, por medio del identificador de la transacción. En la **Figura 42**, se muestran los cambios realizados en la consulta.

Figura 41. Cambios realizados en la consulta de carga de datos del *job*.

```
SELECT
  `sa_sb_transaction`.`transactionId`,
  `sa_sb_transaction`.`merchantId`,
  `sa_sb_transaction`.`marketId`,
  `sa_sb_transaction`.`coreInstanceId`,
  `sa_sb_transaction`.`userId`,
  `sa_sb_transaction`.`typeId`,
  `sa_sb_transaction`.`statusId`,
  `sa_sb_transaction`.`coreAddressId`,
  `sa_sb_transaction`.`address`,
  `sa_sb_transaction`.`currencyId`,
  `sa_sb_transaction`.`requestedAmount`,
  `sa_sb_transaction`.`rate`,
  `sa_sb_transaction`.`finalAmount`,
  `sa_sb_transaction`.`cryptoCurrencyId`,
  `sa_sb_transaction`.`amount`,
  `sa_sb_transaction`.`fee`,
  `sa_sb_transaction`.`mFee`,
  `sa_sb_transaction`.`txid`,
  `sa_sb_transaction`.`reference`,
  `sa_sb_transaction`.`processingFee`,
  `sa_sb_transaction`.`exchangeFee`,
  `sa_sb_transaction`.`internal`,
  `sa_sb_transaction`.`internalTypeId`,
  `sa_sb_transaction`.`async`,
  `sa_sb_transaction`.`validUntil`,
  `sa_sb_transaction`.`creationDate`,
  `sa_sb_transaction`.`updateDate`
FROM
  `sa_sb_transaction`
WHERE
  `sa_sb_transaction`.`transactionId` > 0
ORDER BY
  `sa_sb_transaction`.`transactionId` ASC
LIMIT 450000
```

Figura 41. Cambios realizados en la consulta de carga de datos del *job*. Fuente: elaboración propia. (2020).

1.29.3.3. Comprobación de resultados de los procesos ETL

En esta sección fueron realizadas pruebas para la validación de los datos cargados a través de los Jobs desarrollados en Talend.

Para corroborar que la información de las dimensiones fuera consistente, bastó con reconocer, directamente, que los datos almacenados en las tablas de origen y destino coincidieran. En la **Figura 42** se muestra una comparativa de los datos para la dimensión de categoría de cliente del sistema procesador de criptomonedas. Es posible observar la misma información en la tabla de origen, como en la tabla del *staging* area y en la tabla destino. Los datos de los nombres fueron modificados para proteger la identidad de los clientes de la organización.

Figura 42. Comparativa de los datos para la dimensión de categoría de cliente del sistema procesador de criptomonedas

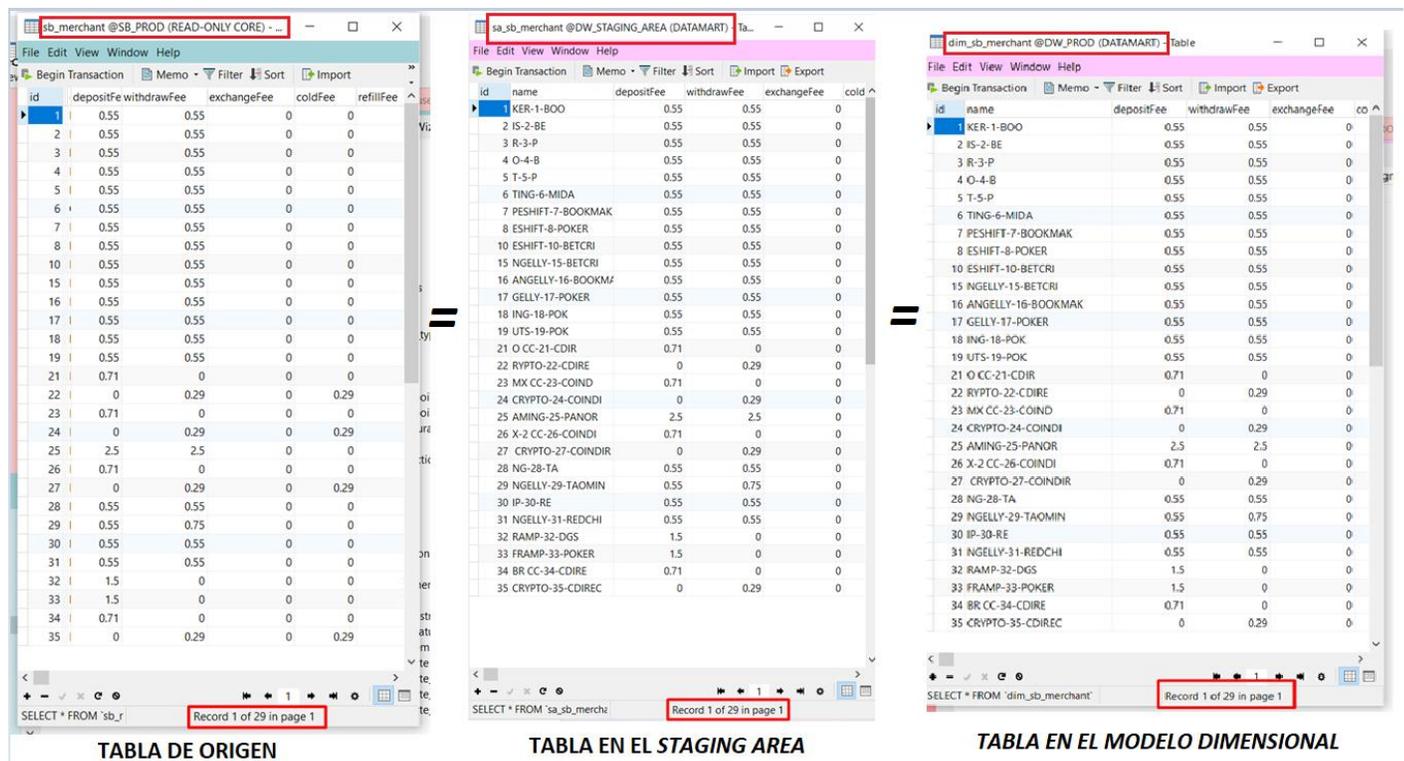


Figura 42. Comparativa de los datos para la dimensión de categoría de cliente del sistema procesador de criptomonedas. Fuente: elaboración propia. (2020).

Por otro lado, para corroborar que la información de las tablas de hechos fue consistente, se realizaron consultas directamente después de terminada la ejecución de los *Jobs*, donde se contabilizaba la cantidad en total de filas; así como la cantidad de estas segmentadas en tres dimensiones distintas.

La **Tabla 25** muestra el resultado de la ejecución de las consultas en las tablas de origen de transacciones del sistema de monedero de criptomonedas, y en su correspondiente tabla de hechos del modelo dimensional del almacén de datos.

Tabla 25. Resultado de la ejecución de las pruebas en las tabla de origen y en la tabla de hechos de las transacciones del procesador de criptomonedas

Monto	Id de tipo			Identificador del estado								Identificador de la moneda				Total
	1	2	3	1	2	3	4	5	6	7	8	1	2	6	7	
Cantidad en tabla de origen	68585	83956	84600	2383	31	13612	220736	25	60	73	221	152541	9	53252	31339	237141
Cantidad en modelo dimensional	68585	83956	84600	2383	31	13612	220736	25	60	73	221	152541	9	53252	31339	237141

Nota. Fuente: elaboración propia. (2020).

La **Tabla 26**, muestra las cantidades de filas para la tabla de origen de las transacciones del procesador de criptomonedas, así como las cantidades de filas para la tabla de hechos en el modelo dimensional, divididas por dimensiones.

Tabla 26. Cantidades de filas para la tabla de origen de las transacciones del procesador de criptomonedas

Monto	Id de tipo		Identificador del estado				Identificador de la moneda					Total
	1	2	1	2	3	4	1	2	3	4	5	
Cantidad en tabla de origen	1954016	446452	395061	838966	580634	585807	160706	2201868	413	337	37144	2400468
Cantidad en modelo dimensional	1954102	446452	395061	838966	580634	585807	160706	2201868	413	337	37144	2400468

Nota. Fuente: elaboración propia. (2020).

1.29.3.4. Automatización de la ejecución de los *Jobs*

De acuerdo con la naturaleza de los sistemas de la organización, las transacciones suelen actualizarse constantemente durante los días próximos a su creación, debido a diferentes factores como las confirmaciones en el *blockchain*. Por consiguiente, se propuso tomar en los últimos 30 días como límite para la consulta de las fuentes de datos de los *Jobs*.

La **Figura 43** muestra el cambio en la consulta para extraer las transacciones que han sido actualizadas en los últimos 30 días.

Figura 43. Cambio en la consulta para actualizar e ingresar las tablas de hechos transacciones

```
SELECT
  `sa_cd_transaction`.`transactionId`,
  `sa_cd_transaction`.`transactionStatusId`,
  `sa_cd_transaction`.`transactionTypeId`,
  `sa_cd_transaction`.`clientDivisionId`,
  `sa_cd_transaction`.`requestedCurrencyId`,
  `sa_cd_transaction`.`requestedAmount`,
  `sa_cd_transaction`.`systemAmount`,
  `sa_cd_transaction`.`systemRate`,
  `sa_cd_transaction`.`externalReferenceId`,
  `sa_cd_transaction`.`creationDate`,
  `sa_cd_transaction`.`updatedAt`
FROM `sa_cd_transaction`
WHERE
  `sa_cd_transaction`.`updatedAt` BETWEEN DATE_SUB(NOW(), INTERVAL 30 DAY) AND NOW();
```

Figura 43. Cambio en la consulta para consultar por las transacciones. Fuente: elaboración propia. (2020).

De igual forma, la **Figura 44** muestra la modalidad de inserción de los *Jobs* configurada en *Talend*, la que identifica por sí misma si los datos deben ser actualizados o si deben de agregar registros nuevos en la tabla destino.

Figura 44. Modalidad de inserción de los *Jobs* configurada en *Talend*



Figura 44. Modalidad de inserción de los *Jobs* configurada en *Talend*. Fuente: elaboración propia. (2020).

1.30. Fase V: Elaboración de *dashboards* para la visualización de los indicadores clave de desempeño

En esta sección se desarrolla la elaboración de los *dashboards* para la visualización de los indicadores clave de desempeño identificados de un nivel alto en la **Tabla 11**. Lo anterior, cumpliendo con lo establecido en la sección de Identificación de requerimientos relacionados a la elaboración de los *dashboards*.

1.30.1. Diseño de la arquitectura técnica

La solución de *AWS* no estaba disponible para la locación de Montreal, donde se ubican los demás servidores de la organización; sin embargo, su instalación se realizó en los servidores de Virginia del norte, sin mayor inconveniente para realizar todas las operaciones realizadas en este proyecto.

A continuación, se indica la **Figura 45**:

Figura 45. Ubicaciones disponibles para el servicio de *Quicksight*

Services Offered:	Northern Virginia	Ohio	Oregon	Northern California	Montreal	São Paulo
Amazon QuickSight	✓	✓	✓			

Figura 45. Ubicaciones disponibles para el servicio de *Quicksight*. Fuente: elaboración propia. (2020).

La **Figura 46**, muestra la notificación de suscripción de la cuenta de *AWS* de la organización, al servicio *Quicksight*.

Figura 46. Notificación de registro del servicio de *Quicksight*

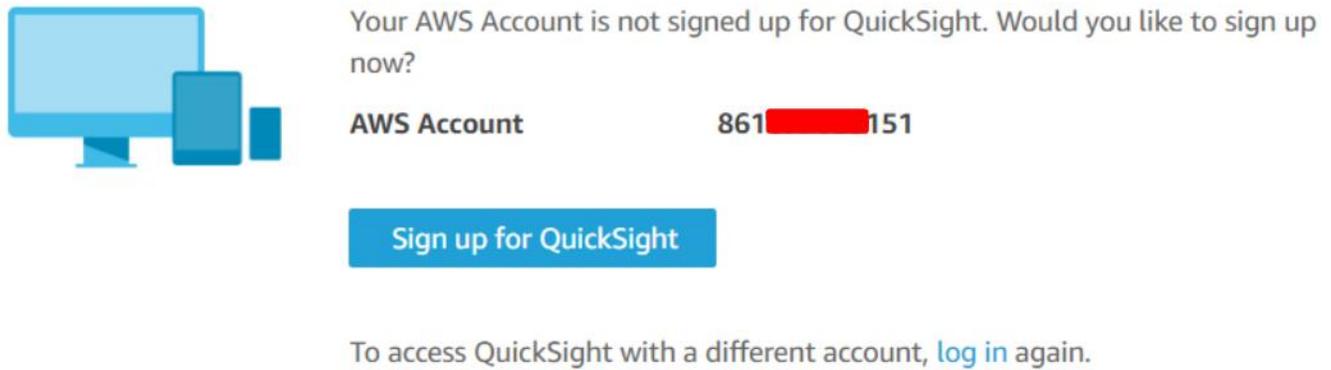


Figura 46. Notificación de registro del servicio de *Quicksight*. Fuente: elaboración propia. (2020).

1.30.2. Desarrollo de la aplicación de visualización de datos

En esta sección se detalla el proceso de carga de información dentro de *Amazon Quicksight*, así como el desarrollo de las visualizaciones para los indicadores clave de desempeño.

1.30.2.1. Carga de información

Para realizar la carga de la información, es necesario realizar una conexión a la base de datos del modelo dimensional, dentro del almacén de datos. La **Figura 47** muestra el detalle de la configuración de una nueva fuente de información MySQL.

Figura 47. La configuración de una nueva fuente de información MySQL

The image shows a web form titled "New MySQL data source" with a close button (X) in the top right corner. The form contains the following fields and controls:

- Data source name:** A text input field with the placeholder text "Enter a name for the data source".
- Connection type:** A dropdown menu currently showing "Public network" with a downward arrow.
- Database server:** A text input field with the placeholder text "Enter hostname or IP address".
- Port:** A text input field with the placeholder text "Port".
- Database name:** A text input field with the placeholder text "Enter a database name".
- Username:** A text input field with the placeholder text "Username".
- Password:** A text input field with the placeholder text "Password".
- Buttons and checkboxes:** At the bottom, there is a "Validate connection" button, a checked checkbox labeled "Enable SSL", and a blue "Create data source" button.

Figura 47. La configuración de una nueva fuente de información MySQL. Fuente: elaboración propia. (2020).

Una vez establecida la conexión, el sistema permite llevar a cabo una carga inicial de información proveniente de una tabla específica. La **Figura 48** muestra el resultado de la carga inicial de información luego de la configuración de la fuente de datos.

Figura 48. Carga inicial de información en el set de datos de Quicksight

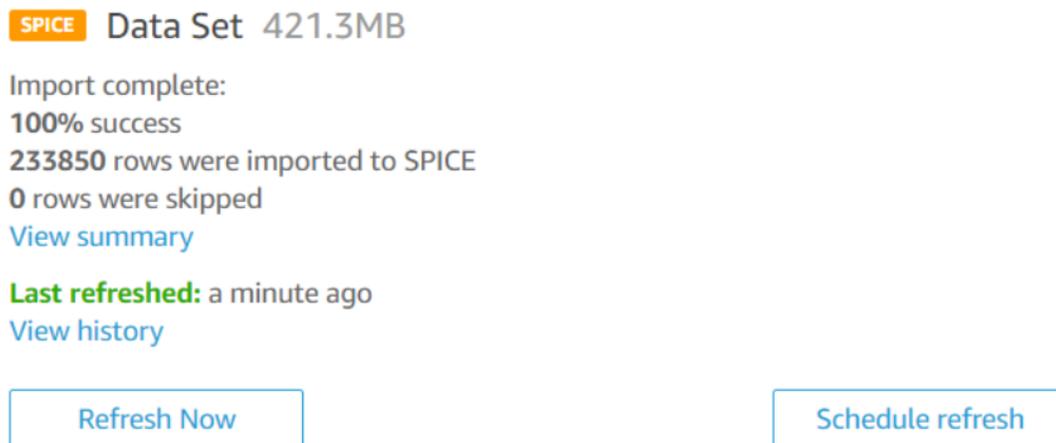


Figura 48. Carga inicial de información. Fuente: elaboración propia. (2020).

La **Figura 49**, muestra el detalle de las fuentes de información importadas para el desarrollo de las visualizaciones:

Figura 49. Fuentes de información importadas para el desarrollo de las visualizaciones

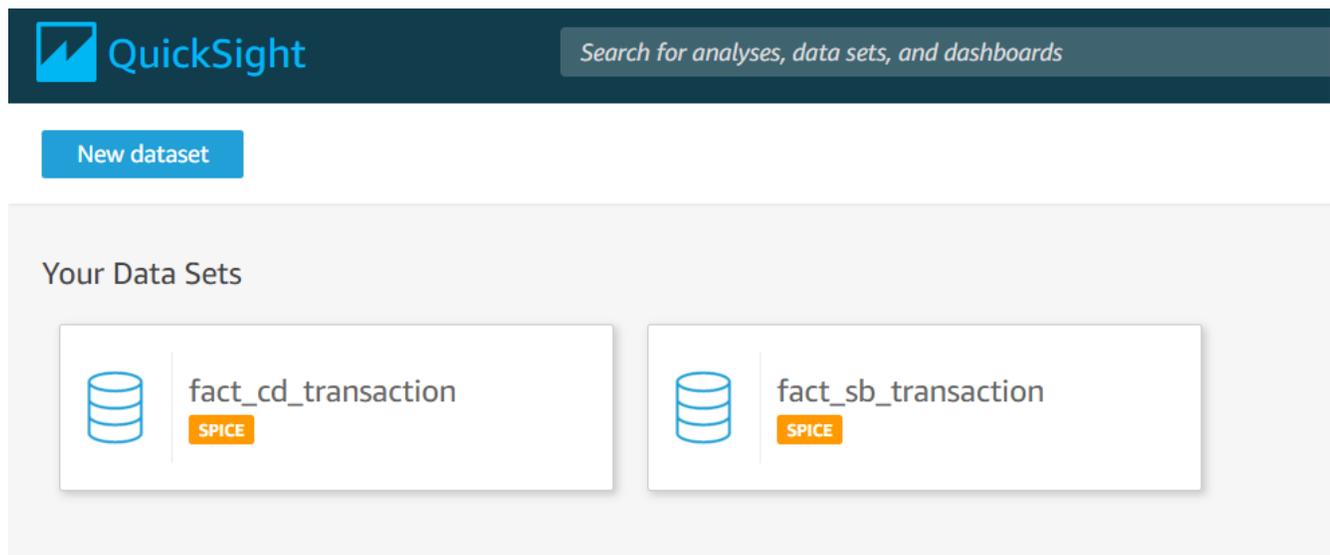


Figura 49. Fuentes de información importadas para el desarrollo de las visualizaciones. Fuente: elaboración propia. (2020).

Una vez realizada una carga inicial de datos, es posible editar la fuente de información para importar otras tablas dentro del sistema. La **Figura 50**, muestra la ventana de importación de nuevas tablas de datos.

Figura 50. Ventana de importación de nuevas tablas de datos

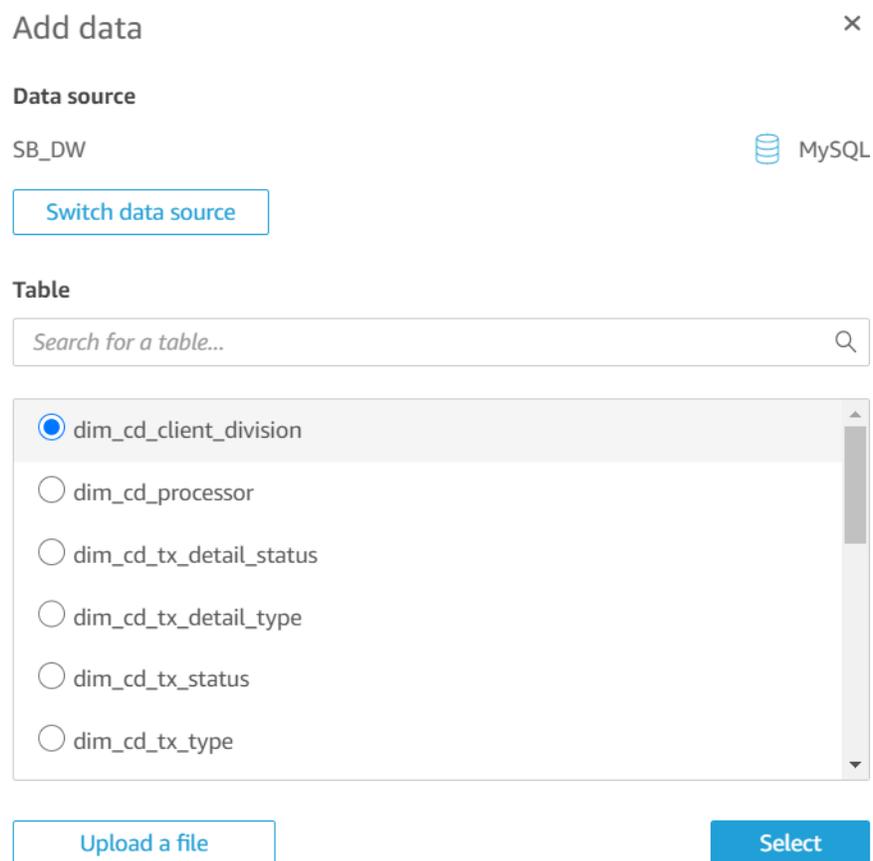


Figura 50. Ventana de importación de nuevas tablas de datos. Fuente: elaboración propia. (2020).

Luego de importar la nueva tabla, es posible especificar directamente la relación con las tablas ya importadas. La **Figura 51**, muestra la configuración de un *join* entre la tabla de hechos y la tabla de dimensión correspondiente a las divisiones de clientes.

Figura 51. Configuración de un *join* entre la dimensión entre dos tablas importas en Quicksight

Join configuration

Join clauses

[+ Add a new join clause](#)

fact_cd_transaction		dim_cd_client_division
clientDivisionId	=	clientDivisionId

Apply

Figura 51. Configuración de un *join* entre la dimensión entre dos tablas importas en Quicksight. Fuente: elaboración propia. (2020).

La **Figura 52**, muestra el diagrama de las fuentes de datos relacionadas con el modelo estrella de las transacciones del monedero de criptomonedas, dentro de *Amazon Quicksight*.

Figura 52. Fuentes de datos relacionadas con el modelo estrella de las transacciones del monedero de criptomonedas

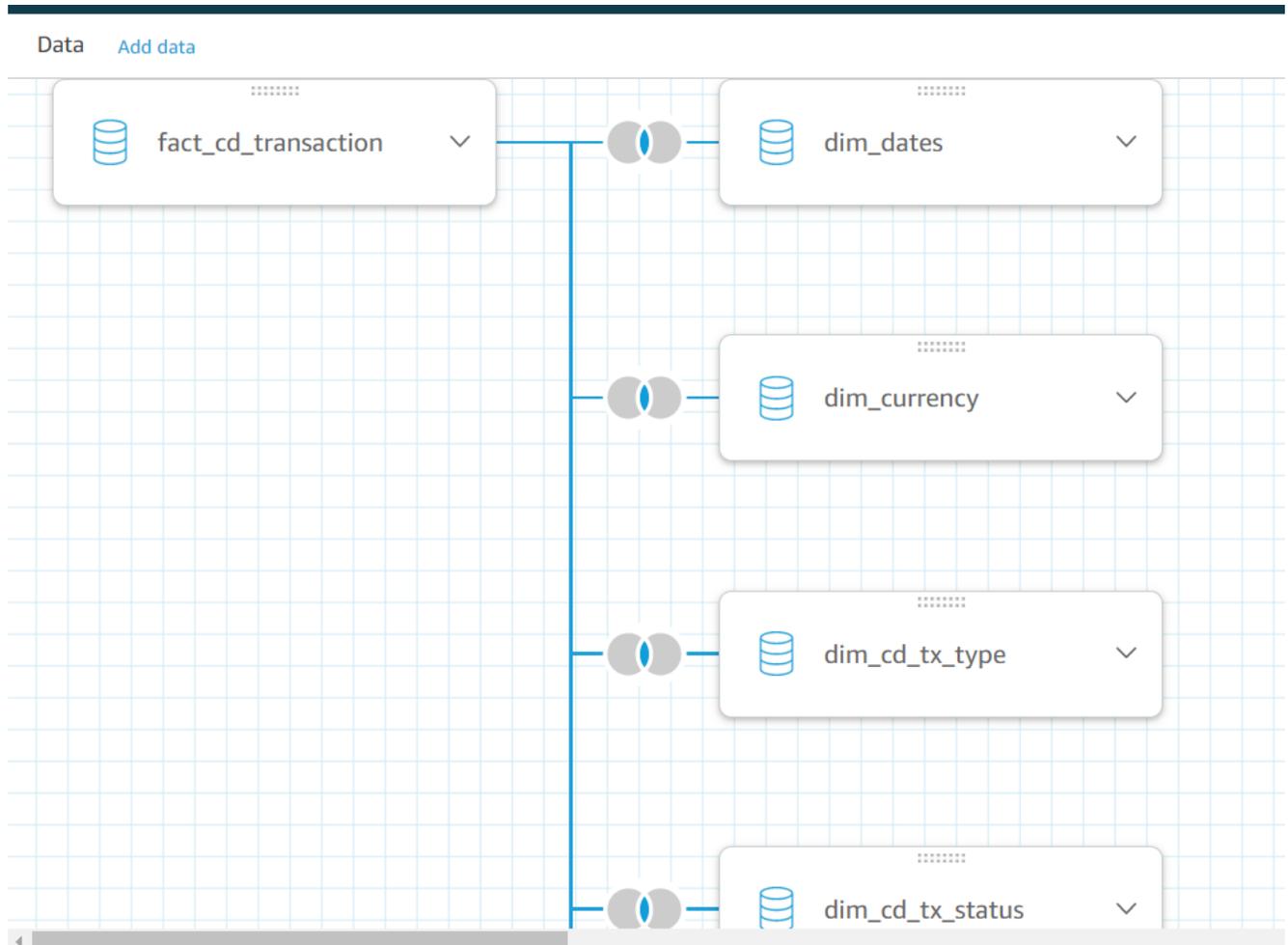


Figura 52. Fuentes de datos relacionadas con el modelo estrella de las transacciones del monedero de criptomonedas. Fuente: elaboración propia. (2020).

La importación de los datos relacionados con el modelo estrella de transacciones del procesador de criptomonedas, sigue un flujo similar al establecido anteriormente para el modelo estrella del monedero.

1.30.2.1. Creación de visualizaciones

Luego de realizar la importación de las fuentes de información, fue posible comenzar con la creación de las visualizaciones dentro del sistema. Es importante recordar que estas se basan en los Indicadores clave de desempeño con mayor nivel de prioridad, definidos en el capítulo anterior. Por otro lado, los valores mostrados en la presente documentación fueron alterados para proteger la confidencialidad de la información de la organización y la de sus clientes.

En primer lugar, la **Figura 53** muestra la visualización para el indicador IND-21: Cantidad promedio pagada por comisión de mineros. En ella se muestra un gráfico de barras separado por meses, donde se presenta el promedio del monto pagado como comisión a los mineros,

Figura 53. Visualización para el indicador IND-21: Cantidad promedio pagada por comisión de mineros

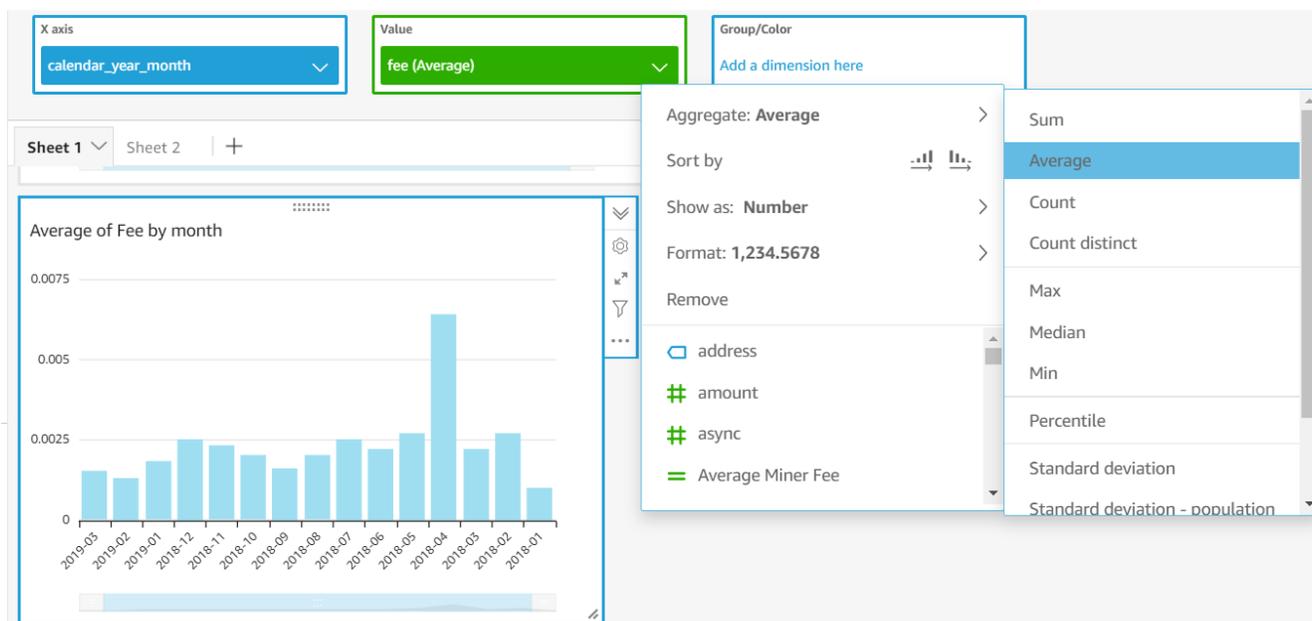


Figura 53. Visualización para el indicador IND-21: Cantidad promedio pagada por comisión de mineros. Fuente: elaboración propia. (2020).

La **Figura 54** muestra un gráfico que abarca los siguientes indicadores:

- IND-12: Cantidad de retiros por categoría de cliente.
- IND-13: Cantidad de transacciones por categoría de cliente.

La visualización se trata de un gráfico de líneas donde cada una de estas representa la cantidad de transacciones realizadas por cada categoría de cliente, a lo largo de una serie de meses. Este gráfico representa ambos indicadores debido a que se agregó un filtro que permite elegir por tipo de transacción, si este filtro no es marcado, se muestra la totalidad de las transacciones.

Figura 54. Cantidad de retiros por categoría de cliente y cantidad de transacciones por categoría de cliente

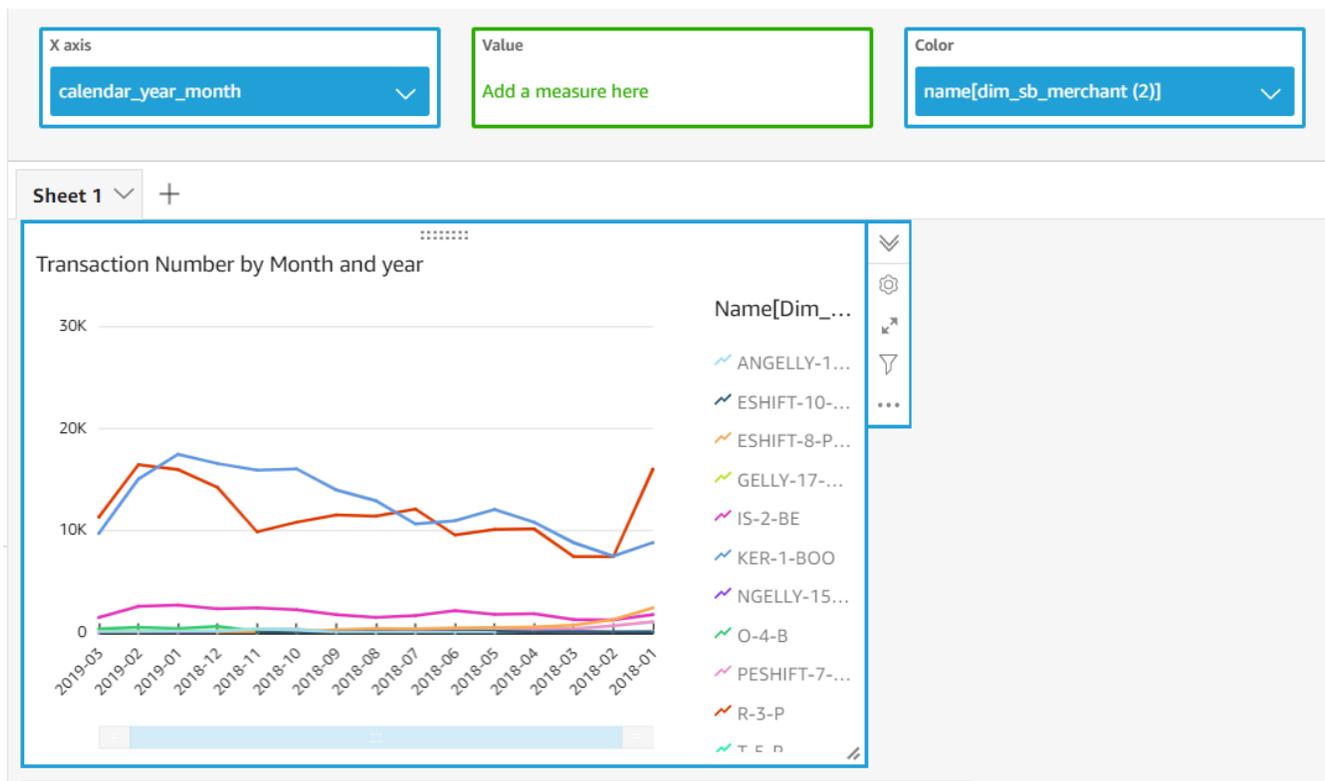


Figura 54. Cantidad de retiros por categoría de cliente y cantidad de transacciones por categoría de cliente
Fuente: elaboración propia. (2020).

Por otro lado, la **Figura 55** representa la visualización de los siguientes indicadores:

- IND-14: monto de dinero de depósitos por categoría de cliente.
- IND-15: monto de dinero de retiros por categoría de cliente.

Esta figura también representa un gráfico líneas, pero en esta ocasión, cada una de estas líneas representa el monto de dinero procesado a través del sistema, y para cada categoría de cliente, a lo largo de una serie de meses. Este gráfico representa ambos indicadores debido a que se agregó un filtro que permite elegir, por tipo de transacción, si este filtro no es seleccionado, se muestra la totalidad del monto de retiros y depósitos sumado.

Figura 55. Monto de dinero de depósitos por categoría de cliente y monto de dinero de retiros por categoría de cliente

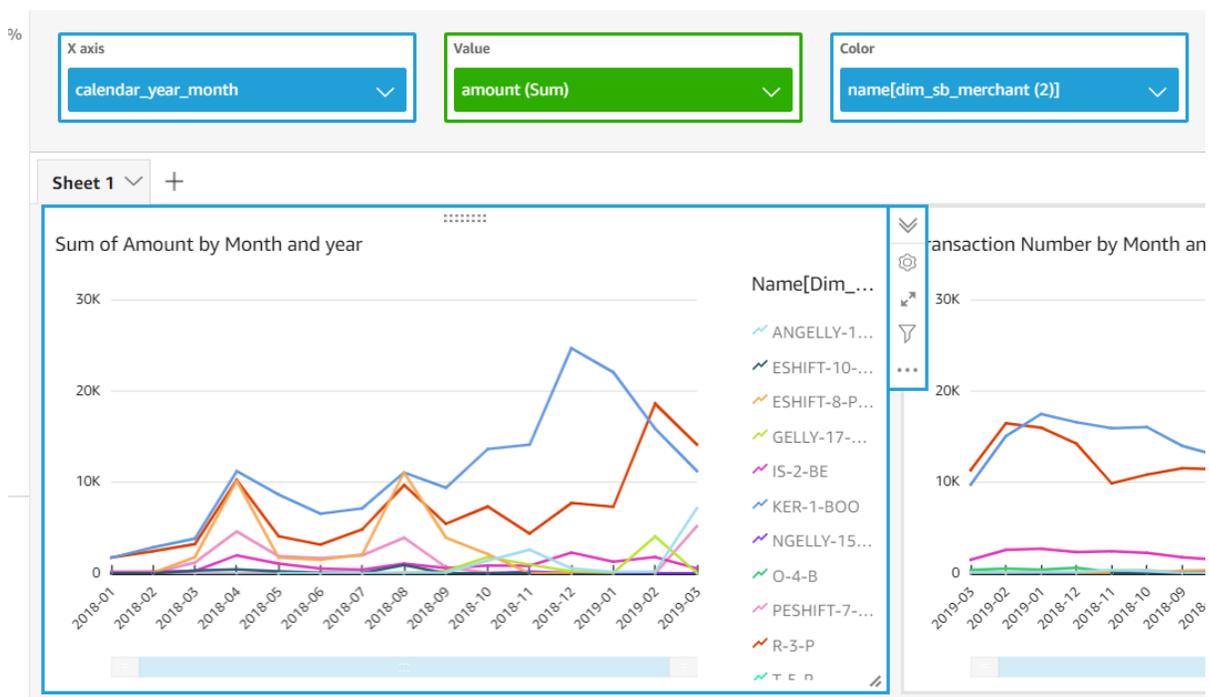


Figura 55. Monto de dinero de depósitos por categoría de cliente y monto de dinero de retiros por categoría de cliente Fuente: elaboración propia. (2020).

Para favorecer el análisis de las anteriores visualizaciones, además del filtro de tipo de transacción, se configuraron los filtros de moneda y de año. La **Figura 56** muestra la configuración de estos filtros.

Figura 56. Filtros de moneda, de tipo de transacción y de año

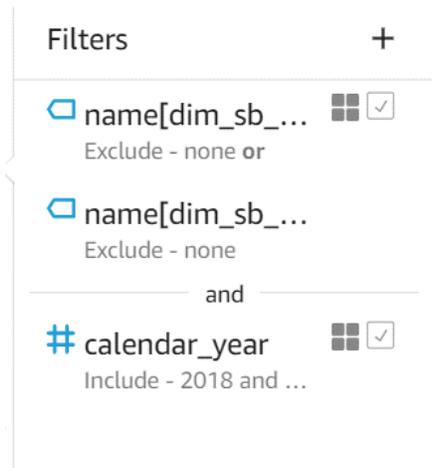


Figura 56. Filtros de moneda, de tipo de transacción y de año. Fuente: elaboración propia. (2020).

La **Figura 57** representa el indicador IND-28: Cantidad de transacciones de compra dinero cripto que han sido procesadas. Se presenta por medio de un gráfico de barras, la cantidad de transacciones para cada división de cliente, a lo largo de los meses.

Figura 57. Indicador IND-28: Cantidad de transacciones de compra dinero cripto



Figura 57. Indicador IND-28: Cantidad de transacciones de compra dinero cripto. Fuente: elaboración propia. (2020).

Por otro lado, la **Figura 58** representa la visualización para el indicador IND-27: Monto de las transacciones de compra de dinero cripto procesadas, por medio de un gráfico de líneas relleno, que permite visualizar la magnitud de diferencias en el procesamiento entre las divisiones de los clientes en el transcurso de los meses.

Figura 58. Indicador IND-27: Monto de las transacciones de compra de dinero cripto procesadas

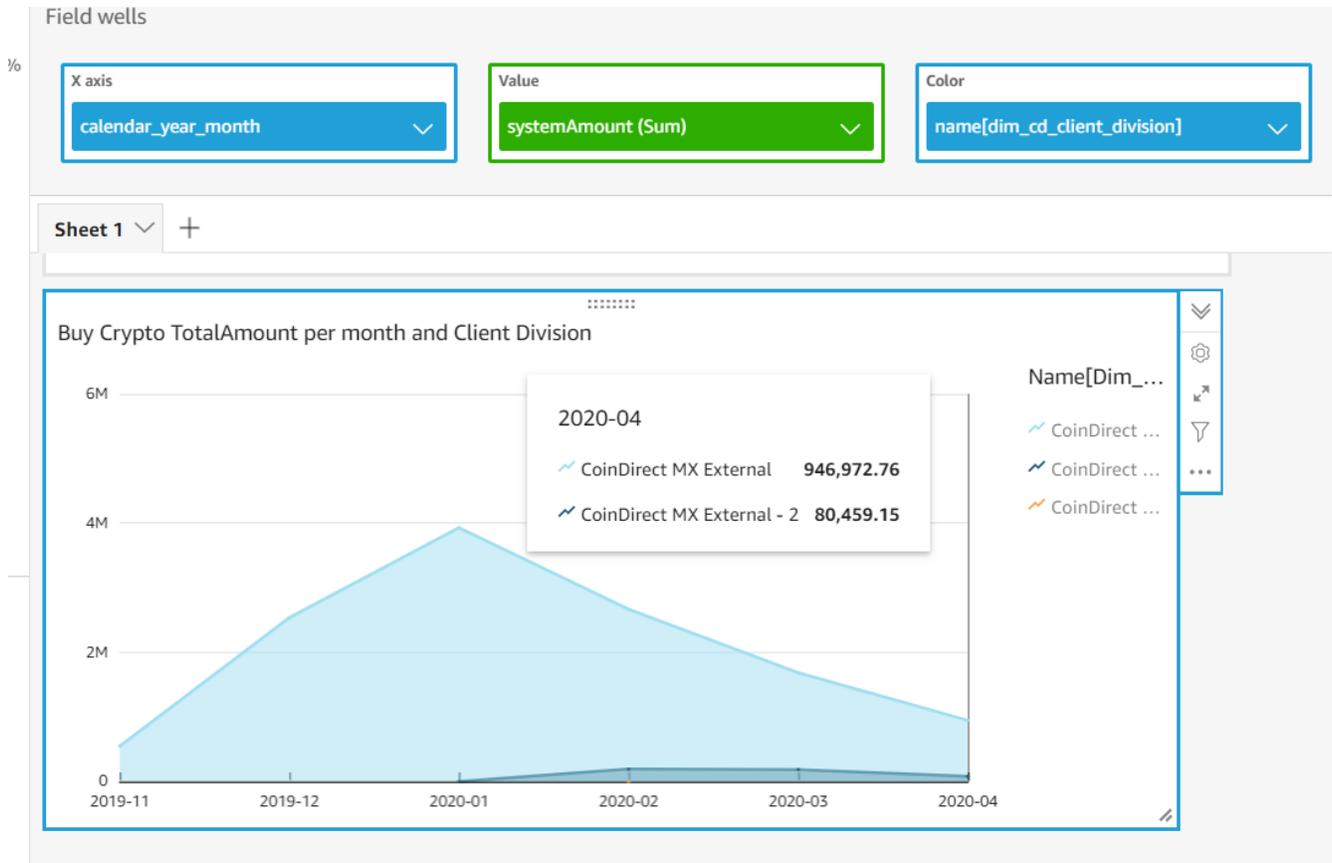


Figura 58. Indicador IND-27: Monto de las transacciones de compra de dinero cripto procesadas. Fuente: elaboración propia. (2020).

Por último, **Figura 59** representa el gráfico para la visualización del indicador IND-29: monto promedio de las transacciones de compra de dinero cripto. Este gráfico permite ver este monto a lo largo de los meses por cada una de las divisiones de clientes del sistema.

Figura 59. Indicador IND-29: monto promedio de las transacciones de compra de dinero cripto

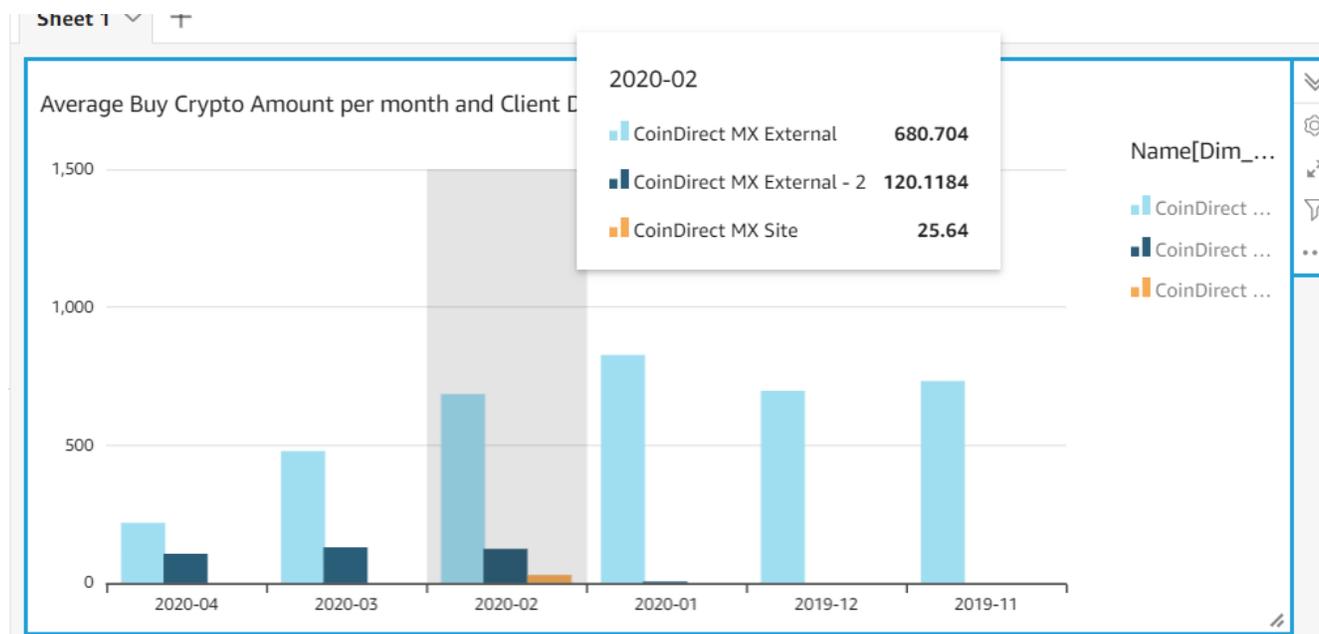


Figura 59. Indicador IND-29: monto promedio de las transacciones de compra de dinero cripto. Fuente: elaboración propia. (2020).

1.30.3. Implementación

Luego de haberse creado las visualizaciones, es momento de configurar aspectos como el refrescamiento automático de los datos, y el envío de correo electrónico a los interesados.

La **Figura 60**, muestra la configuración realizada para el refrescamiento de los datos de manera calendarizada. Si bien es cierto se solicitaba que este se ejecutara cada 24 horas, se propuso realizar dos refrescamientos por día, con 12 horas de diferencia entre ellos, aprovechando que la herramienta lo permite.

Figura 60. Refrescamiento de los datos de manera calendarizada

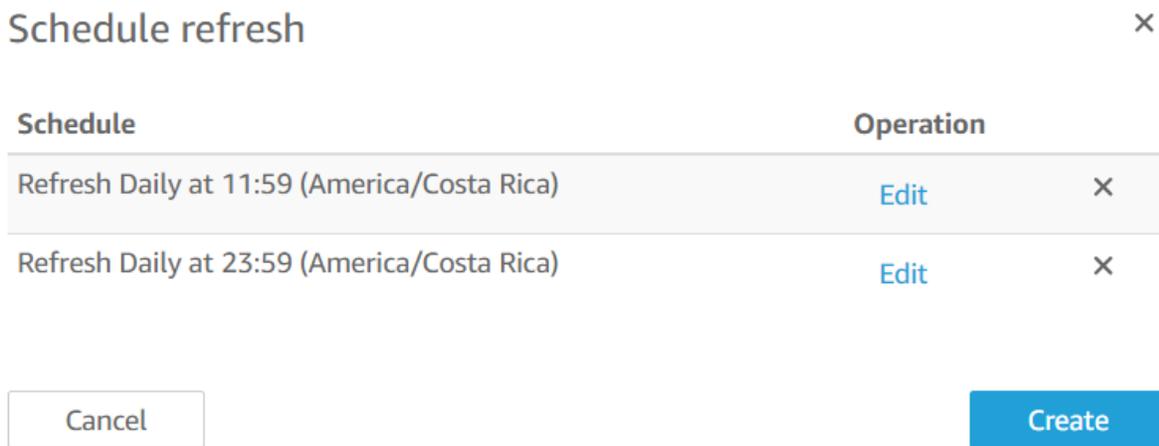


Figura 60. Refrescamiento de los datos de manera calendarizada. Fuente: elaboración propia. (2020).

En la **Figura 61** se muestra un ejemplo del envío por correo electrónico de las visualizaciones, donde se le configura al usuario, únicamente permisos de lectura del *dashboard*.

Figura 61. Ejemplo de configuración de envío por correo electrónico de las visualizaciones.

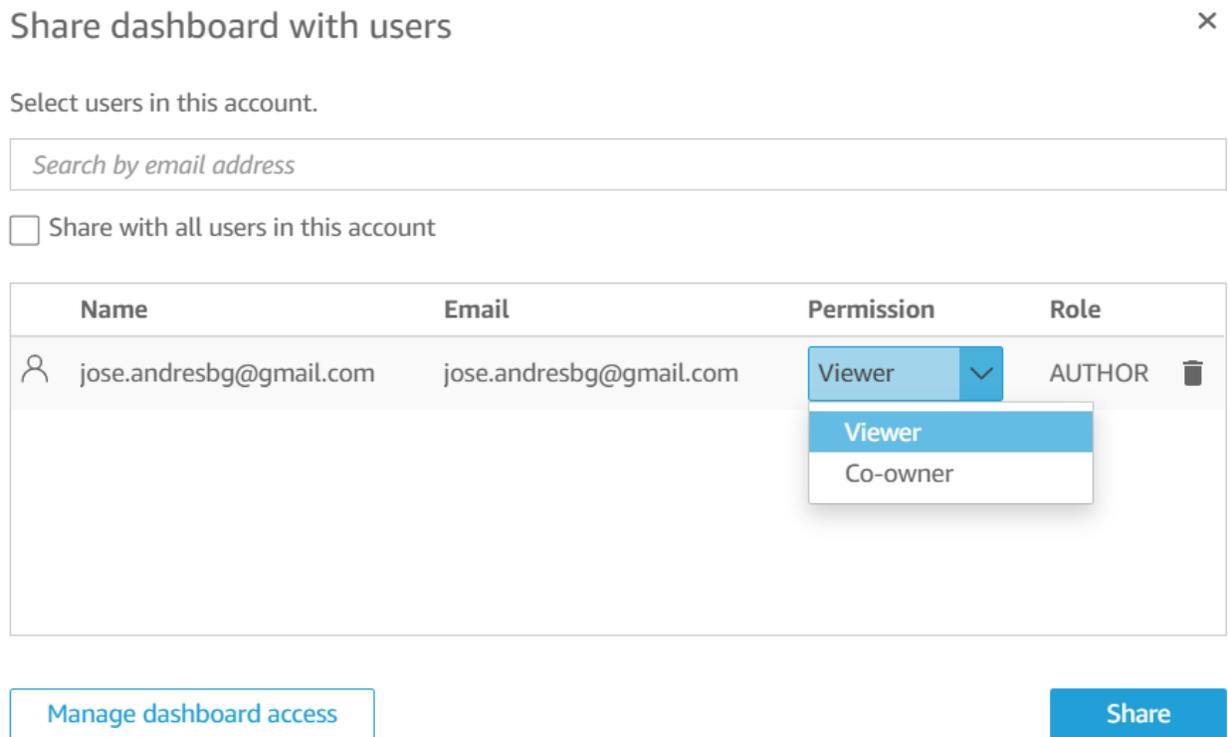


Figura 61. Ejemplo de configuración de envío por correo electrónico de las visualizaciones. Fuente: elaboración propia. (2020).

La **Figura 62**, muestra cómo son recibidas las invitaciones al correo electrónico, mientras que la **Figura 63** muestra la vista de los *dashboards* para los usuarios con permisos de lectura solamente.

Figura 62. Invitaciones al correo electrónico

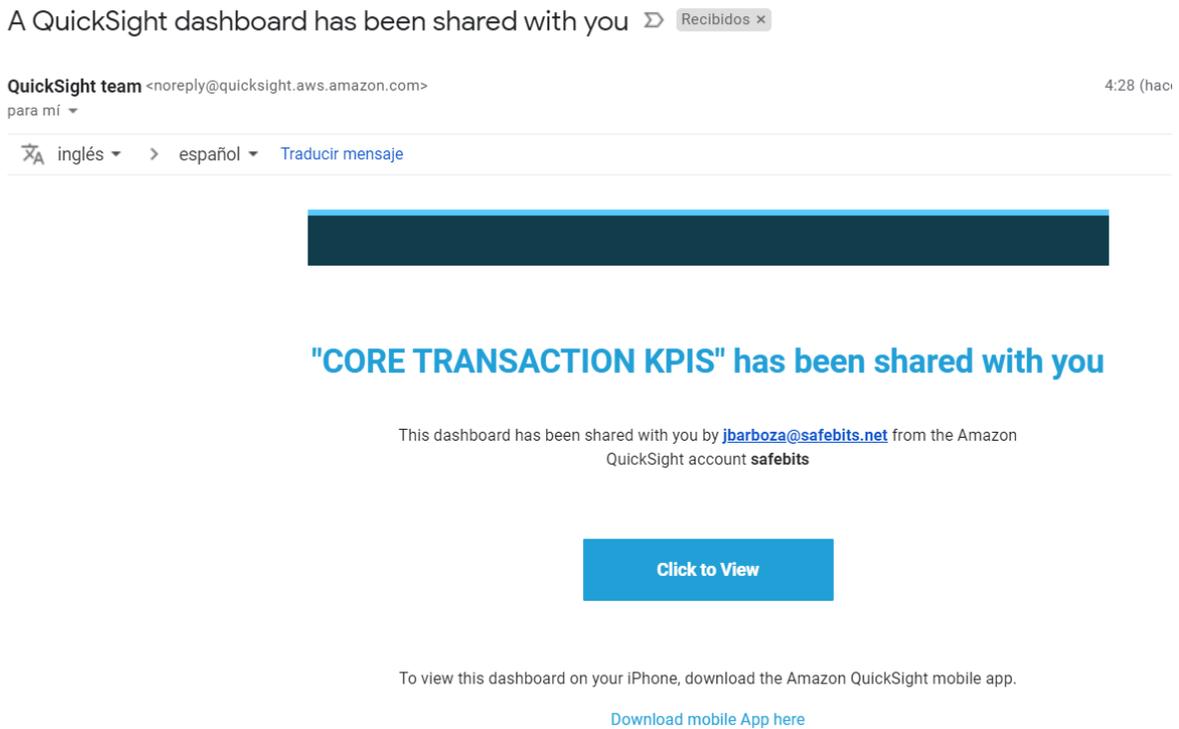


Figura 62. Invitaciones al correo electrónico fuente: elaboración propia. (2020).

Figura 63. Vista de los dashboards para los usuarios con permisos de lectura solamente

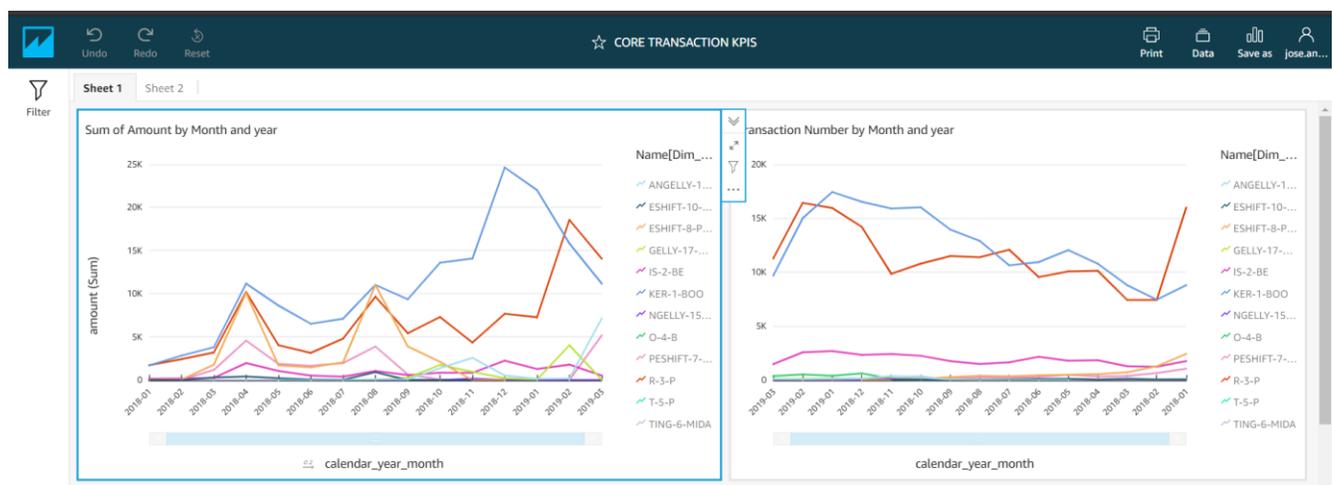


Figura 63. Vista de los dashboards para los usuarios con permisos de lectura solamente. Fuente: elaboración propia. (2020).

CAPÍTULO VI: CONCLUSIONES

6. CAPÍTULO VI: CONCLUSIONES

En este capítulo se presentan las conclusiones de la investigación llevada a cabo, estas tienen la finalidad de detallar, con base en los resultados obtenidos, el cumplimiento de los Objetivos específicos propuestos en el desarrollo del primer capítulo del presente documento.

A continuación, se encuentran estas conclusiones debidamente detalladas para cada uno de los objetivos específicos:

- I. Analizar los sistemas de información de la empresa, sus procesos y sus correspondientes bases de datos para el planteamiento de la situación actual que oriente el desarrollo de la solución de inteligencia de negocios.***
1. Para la organización, los sistemas con principal necesidad de análisis son el encargado del procesamiento de criptomonedas y el monedero electrónico de criptomonedas, fue debido a ello que el desarrollo de la solución se enfocó en abarcar las necesidades identificadas para estos dos sistemas.
2. A partir de la identificación de los sistemas con mayor prioridad de análisis, se determinó que las principales operaciones básicas del sistema de procesamiento de criptomonedas son el depósito y el retiro de dinero criptomonedas, con lo cual fueron determinados *KPIs* relacionados a estos.
3. El sistema de monedero cuenta con dos procesos principales, el de compra de criptomonedas, el cual hace referencia al dinero fiat a partir de tarjetas de crédito; y el de la generación y ejecución automatizada de retiros de dinero, a través del procesador de criptomonedas; por ello, la identificación de *KPIs* para este monedero se basó en estos dos procesos.
4. Los sistemas, arquitectura y herramientas que se identificaron como fundamentales para la solución de inteligencia de negocios, se encuentran integrados entre sí, lo que facilita que la solución se ajuste a los objetivos de la organización.

II. Proponer indicadores clave de desempeño sobre los sistemas de la organización para que estos sirvan como apoyo en procesos de control, monitoreo y toma de decisiones en la organización.

1. Con base en la recolección de información llevada a cabo para los sistemas seleccionados, fue posible identificar 32 indicadores clave de desempeño, estos indicadores son una base para el monitoreo, control, seguimiento y mejora del desempeño de los sistemas de información estudiados en la investigación.
2. A partir de la lista de indicadores identificados, se llevó a cabo la asignación de prioridades, a partir de dicha priorización, se decidió abarcar el desarrollo de nueve indicadores que se clasificaron con prioridad alta, brindando al negocio la información de mayor relevancia.

III. Diseñar el modelo del datamart, así como los procesos de extracción, transformación y carga de datos (ETL) para soportar la carga, estandarización y actualización de la información proveniente de los sistemas de la empresa.

1. El desarrollo de la solución con respecto al *datamart* se basó en la configuración de las bases de datos de los sistemas de la organización, las que se encuentran alojadas en los servidores de *Amazon Web Services (AWS)* bajo el servicio de *Relational Database Service (RDS)*, específicamente manejadas por el motor *MySQL* versión 8.0.
2. En la instancia *RDS* de *AWS*, se ejecutó el diseño de dos bases de datos, la primera correspondiente al área de trabajo o *staging area* donde se almacenó la información sin ningún tipo de transformación, de manera que esta sirvió de fuente para la segunda de las bases de datos, correspondiente al *datamart* como tal, de esta manera se logró evitar la carga directa sobre las bases de datos operacionales de los sistemas de la organización.
3. La identificación de las tablas de las bases de datos se basó en la información requerida para efectuar el cálculo de los identificadores clave de desempeño identificados y priorizados en la fase anterior.

4. Se identificó principalmente dos tablas de hechos, una relacionada a las transacciones del sistema de monedero y otra del sistema de procesamiento de criptomonedas, para las cuales se establecieron las dimensiones correspondientes.
5. Respecto al *ETL*, se creó una instancia en *Ubuntu Linux* donde están los demás servidores de la empresa, donde se ejecutaron y calendarizaron los *Jobs* creados por medio de la herramienta Talend.
6. Para el manejo de las versiones de *job*, se utilizó un repositorio de *git* configurado en la herramienta de la organización "*Bitbucket*". Por otro lado, se utilizó la herramienta *Jenkins* para automatizar el paso del repositorio a la instancia de los *Jobs*.
7. Se definió un periodo de treinta días para la actualización o carga de nuevos datos en las bases de datos, asegurando así la veracidad en la información almacena que, por la naturaleza de las soluciones en el *blockchain*, es cambiante.
8. Para la escogencia de la herramienta de *ETL*, se realizó una evaluación de los líderes de mercado contra algunas herramientas recomendadas por los desarrolladores de *DevOps*, donde la seleccionada fue Talend, la cual cuenta con una versión gratuita que se ajusta a las necesidades de la organización.

IV. Elaborar dashboards con vistas gráficas de indicadores de desempeño para el favorecimiento del proceso de análisis de la información y la toma de decisiones gerenciales, con respecto al procesamiento de criptomonedas.

1. Se realizó una comparativa de herramientas contemplando los líderes del mercado, una herramienta proveída por los servidores de *AWS* y una propuesta del administrador de proyectos, donde la selecciona fue *Quicksight*, brindando al negocio la alternativa más acorde a sus necesidades.

2. Se desarrollaron los dashboard en la herramienta atendiendo los indicadores claves de desempeño categorizados como de urgencia alta; favoreciendo así el proceso de toma de decisiones de la organización, con respecto a sus prioridades.
3. Para la calendarización y refrescamiento de los datos, se establecieron dos corridas diarias y el envío por correo de los *dashboards*, con la finalidad de asegurar la disponibilidad e integridad de la información para lo usuarios de la solución.
4. En la solución, se configuraron accesos de lectura y escritura para usuarios; de esta manera, cada uno obtuvo acceso a información que le es pertinente, asegurando así que la información y el diseño de los *dashboards* no fuera expuesto de manera descuidada.

CAPÍTULO VII: RECOMENDACIONES

7. CAPÍTULO VII: RECOMENDACIONES

El presente capítulo tiene como objetivo documentar las recomendaciones identificadas durante toda la ejecución del Trabajo Final de Graduación, lo cual, se espera que sirva como insumo para que la organización pueda brindarle continuidad al proyecto.

Como se mencionó anteriormente, se refiere a recomendaciones que puede tomar la organización, para mejorar el producto final entregado, pero, no se ve en la obligación de aplicarlas a corto plazo.

1. Se aconseja incluir dentro del proceso las fases restantes que propone la metodología Kimball (2006), con el fin de dar mantenimiento a las soluciones de inteligencia de negocios, y, a su vez, permitir el crecimiento de soluciones, en el futuro, hacia otros sistemas.
2. Si bien es cierto, el alcance de la presente solución no contemplaba la integración de datos entre los diferentes sistemas, se recomienda aprovechar las bases establecidas con respecto al manejo de procesos ETL para el diseño de tablas de base de datos y *jobs* de la herramienta Talend, que genere información nueva para el análisis del negocio por parte de la gerencia de la organización.
3. Se recomienda a los usuarios de los dashboards el uso de la aplicación móvil para contar con acceso a las visualizaciones, desde cualquier lugar donde tengan disponible su teléfono. La aplicación como tal no cuenta con un costo adicional, por ende, facilita la accesibilidad para que los colaboradores hagan uso de esta.
4. A través de la solución propuesta por medio del Trabajo Final de Graduación, la cual permite medir el desempeño de los sistemas de información, se propone que la organización incluya, en un corto plazo, acuerdos de nivel de servicio, para fortalecer la relación con el cliente, mediante pautas que definan específicamente el servicio que se les brinda.
5. Se recomienda analizar en el futuro la capacidad y rendimiento que tenga el Datamart. Lo anterior es debido a que la organización está creciendo constantemente, lo que puede implicar que se deba optar por una solución con mayor capacidad, como lo es un *datawarehouse*.

6. Es importante monitorear el crecimiento que tenga la solución de inteligencia de negocios, debido a que, eventualmente, será necesario incorporar mejoras en la arquitectura técnica tanto del Datamart como de la instancia de ejecución de los *ETLs*, de manera que se pueda mantener una continua disponibilidad de la solución.
7. Se sugiere tomar en cuenta los indicadores claves de desempeño que quedaron fuera del alcance del proyecto, debido a que, aunque su prioridad no era alta, agregan mucho valor al proceso de toma de decisiones.
8. Finalmente, es relevante considerar la inclusión de herramientas o soluciones que permitan llevar a cabo predicciones sobre el comportamiento de las soluciones con base en los datos recolectados con la solución de IN, se podrían considerar para ello tecnologías como inteligencia artificial o *machine learning*.

Apéndices

Apéndice I. Plantilla de minutas para las reuniones del proyecto

Tabla 27. *Plantilla para la gestión de minutas de reunión*

Minuta de reunión			
Número de reunión			
Fecha		dd/mm/aaaa	
Lugar			
Objetivo de la reunión			
Participantes		Presentes:	
		Ausentes:	
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1			
2			
Próxima Reunión			
Temas a tratar		Fecha	Convocados
		dd/mm/aaaa	

Nota. Fuente: elaboración propia. (2020).

Apéndice II. Plantilla para la gestión de cambios

Tabla 28. *Plantilla para la gestión de solicitudes de cambio*

Gestión de solicitudes de cambio	
Numero de solicitud	
Fecha	dd/mm/aaaa
Solicitante	
Descripción del cambio	
Detalle de aspectos a modificar	
Involucrados	
Observaciones sobre la implementación del cambio	
Aprobación por parte del estudiante a cargo del proyecto	
Aprobación por parte del tutor a cargo del proyecto	
Aprobación por parte del representante de la organización	
Fecha de aprobación o rechazo	

Nota. Fuente: elaboración propia. (2020).

Apéndice III. Plantilla para entrevistas semiestructuradas

Tabla 29. Plantilla para entrevistas semiestructuradas

Entrevista semiestructurada	
Participantes	
Objetivo de la entrevista	
Pregunta:	Respuesta:
Observaciones adicionales:	

Nota. Fuente: elaboración propia. (2020).

Apéndice IV. Entrevista abierta

Tabla 30. *Entrevista abierta*

Entrevista abierta	
Participantes	
Objetivo de la entrevista	
Tema abordado:	Observaciones:
Observaciones adicionales:	

Nota. Fuente: elaboración propia. (2020).

Apéndice V. Revisión documental

Tabla 31. *Revisión documental*

Revisión documental	
Objetivo	
Nombre del documento	
Observaciones	

Nota. Fuente: elaboración propia. (2020).

Apéndice VI. Observación

Tabla 32. *Observación*

Observación	
Tipo de participación	
Objetivo	
Comentarios	

Nota. Fuente: elaboración propia. (2020).

Apéndice VII. Tabla comparativa de herramientas

Tabla 33. *Tabla comparativa de herramientas*

Nombre de la herramienta	Criterios				
	Criterio 1 (# %)	Criterio 2 (# %)	Criterio 3 (# %)	Criterio 4 (# %)	Criterio 5 (# %)
Herramienta 1					
Herramienta 2					
Herramienta 3					
Herramienta 4					
Herramienta 5					

Nota. Fuente: elaboración propia. (2020).

Apéndice VIII. Grupo Focal

Tabla 34. *Grupo Focal*

Grupo focal	
Objetivo	
Temas tratados	
Comentarios	
Cantidad de participantes	

Nota. Fuente: elaboración propia. (2020).

Apéndice IX. Análisis de la situación actual de los sistemas de la organización

Tabla 35. Entrevista sobre situación actual de los sistemas de la organización

Entrevista semiestructurada	
Participantes	Edgar Cordero; administrador de proyectos
Objetivo de la entrevista	Recopilar información general sobre los sistemas de la organización.
Pregunta: ¿Cuál es su rol para en la organización y cuáles son sus principales responsabilidades?	Respuesta: Soy el administrador de proyectos, me encargo de dirigir la ejecución de los proyectos y de aplicar la metodología SCRUM, en la cual se definen tareas, se priorizan, se analizan las arquitecturas y se ejecutan, etc.
Pregunta: ¿Con qué frecuencia monitorea los factores clave de éxito?	Respuesta: No existen factores claves de éxito para monitorear, existen algunas consultas de bases de datos que usualmente se revisan, pero sin ninguna frecuencia, solo ocasionalmente.
Pregunta: ¿Qué informes utiliza actualmente?	Respuesta: Se utilizan consultas de bases de datos para la revisión del sistema de procesamiento de criptomonedas y otro para el sistema de monedero de criptomonedas.
Pregunta: ¿Qué datos del informe son importantes?	Respuesta: Actualmente, se obtiene información relacionada a las transacciones de depósito y retiro del sistema de procesamiento de criptomonedas y del sistema de monedero de criptomonedas relacionadas a lo que es el proceso de compras de criptomonedas por medio de tarjetas de crédito y también se monitorea información relacionada a la ejecución de los retiros planificados.
Pregunta:	Respuesta: utilizamos las bases de datos operacionales de los sistemas.

¿En general cuáles son las fuentes de información que se utilizan?	
Pregunta: ¿Qué análisis le gustaría realizar?	Respuesta: nos gustaría realizar un análisis visual de las transacciones de los sistemas por un período de tiempo.
Observaciones adicionales: se coordina una reunión con los ingenieros de DevOps para analizar requerimientos técnicos de la solución.	

Nota. Fuente: elaboración propia. (2020).

Apéndice X. Revisión documental de las especificaciones técnicas e implementación de los APIs del sistema de procesamiento de criptomonedas

Tabla 36. *Detalle de la revisión documental del detalle del APIs del sistema de procesamiento de criptomonedas*

Revisión documental	
Objetivo	Analizar el flujo de las operaciones del sistema del procesamiento de criptomonedas.
Nombre del documento	API de procesamiento de criptomonedas del sistema de Safebits
Observaciones	<ul style="list-style-type: none">- Se observa el flujo principal que sigue el cliente con respecto a los retiros y depósitos de dinero.- Se especifican los significados de algunos datos devueltos en las respuestas de este API. Por ejemplo: los montos, la comisión de dinero, tipos y estatus de la transacción, entre otros.

Nota. Fuente: elaboración propia. (2020).

Apéndice XI. Revisión documental de las especificaciones técnicas e implementación de los APIS del monedero electrónico.

Tabla 37. *Revisión documental de las especificaciones técnicas e implementación de los APIS del monedero electrónico*

Revisión documental	
Objetivo	Analizar el flujo de las operaciones del monedero de criptomonedas.
Nombre del documento	API de compra de criptomonedas del sistema de monedero.
Observaciones	<ul style="list-style-type: none">- Se observa el flujo principal que sigue el cliente con respecto a los retiros y depósitos de dinero.- Se especifican los significados de algunos datos devueltos en las respuestas de este API. Por ejemplo: los montos, la comisión de dinero, tipos y estatus de la transacción, entre otros.

Nota. Fuente: elaboración propia. (2020).

Apéndice XII. Revisión documental de los diseños lógicos de las bases de datos del sistema de procesamiento de criptomonedas.

Tabla 38. Revisión documental de los diseños lógicos de las bases de datos del sistema de procesamiento de criptomonedas.

Revisión documental	
Objetivo	Identificar las fuentes de información y los detalles técnicos de la base de datos del procesador de criptomonedas
Nombre del documento	Modelo físico de la base de datos del procesador de criptomonedas.
Observaciones	<ul style="list-style-type: none">- Se identificaron las tablas y columnas relacionadas a los procesos de las transacciones. Por ejemplos: tipos de datos y tabla original de transacciones. Se encontró que el nombre de las tablas está escrito en <i>Snake Case</i> y el nombre de las columnas se encuentra en <i>Camel Case</i>.- La base de datos se encuentra alojada en un servidor de RDS de AWS. El acceso se encuentra limitado por dirección IP de las conexiones.- La tecnología de manejo de la base de datos es MySQL, versión 8.0.

Nota. Fuente: elaboración propia. (2020).

Apéndice XIII. Revisión documental de los diseños lógicos de las bases de datos del monedero electrónico.

Tabla 39. Revisión documental de los diseños lógicos de las bases de datos del monedero electrónico.

Revisión documental	
Objetivo	Identificar tipos de datos y las fuentes de datos del sistema de procesamiento de monedero electrónico.
Nombre del documento	Modelo físico de la base de datos del monedero de criptomonedas.
Observaciones	<ul style="list-style-type: none"> - Las observaciones son similares a la base de datos anterior. Primeramente, se identificaron las tablas y columnas relacionadas a los procesos de las transacciones. Por ejemplos: tipos de datos y tabla original de transacciones. Se encontró que el nombre de las tablas está escrito en <i>snakecase</i> y el nombre de las columnas se encuentra en <i>Cammel</i>. - La base de datos se encuentra alojada en un servidor de RDS de AWS. El acceso se encuentra limitado por dirección IP de las conexiones. - La tecnología de manejo de la base de datos es MySQL, versión 8.0.

Nota. Fuente: elaboración propia. (2020).

Apéndice XIV. Revisión documental consulta para revisión de transacciones del sistema de procesamiento de criptomonedas.

Tabla 40. Consulta para revisión de transacciones del sistema de procesamiento de criptomonedas.

Revisión documental	
Objetivo	Identificar los datos de relevancia y filtros que se toman en cuenta para la ejecución de la consulta.
Nombre del documento	Consulta para la revisión estadística de transacciones del sistema procesador de criptomonedas.
Observaciones	<ul style="list-style-type: none">- Se identificaron varios indicadores relacionados al tema de las transacciones de criptomonedas.- Se identificaron algunos filtros como el tipo de transacción, categoría de clientes, etc.

Nota. Fuente: elaboración propia. (2020).

Apéndice XV. Revisión documental ejecución de retiros planificados del monedero electrónico

Tabla 41. *Ejecución de retiros planificados del monedero electrónico*

Revisión documental	
Objetivo	Identificar cuáles eran algunos de los indicadores o aspectos que se evalúan en las consultas
Nombre del documento	Consulta de análisis de retiros planificados
Observaciones	<ul style="list-style-type: none">- Se identificó que algunos datos analizados son: montos, cantidad de transacciones y estados de procesamiento.

Nota. Fuente: elaboración propia. (2020).

Apéndice XVI. Revisión documental sobre consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.

Tabla 42. *Revisión documental sobre consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.*

Revisión documental	
Objetivo	Identificar la arquitectura de relevancia para leer indicadores.
Nombre del documento	Consulta para el análisis de transacciones de compras de criptomonedas en el monedero electrónico.
Observaciones	<ul style="list-style-type: none">- Existen tipos de transacciones.- Existen otros datos de relevancia como los clientes.- Lo que se observa principalmente son los montos y la cantidad de transacciones.

Nota. Fuente: elaboración propia. (2020).

Apéndice XVII. Observación del flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico con desarrollador de *software*.

Tabla 43. Observación del flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico.

Observación	
Tipo de participación	Moderada
Objetivo	Observar el flujo de las operaciones de los sistemas de procesamiento de criptomonedas y del monedero electrónico
Comentarios	<ul style="list-style-type: none">- Se observa la información que se devuelve en las respuestas del proceso. Se muestra información relacionada al procesador de tarjeta como al procesador de criptomonedas. Dentro de esta información se encuentran montos fíat, montos de criptomonedas, estatus del flujo, entre otros.

Nota. Fuente: elaboración propia. (2020).

Apéndice XVIII. Grupo focal para la revisión y priorización de los *KPI* propuestos con gerente general, administrador de proyectos y desarrollador de *software*

Tabla 44. Grupo focal para la revisión y priorización de los *KPI* propuestos

Grupo focal	
Objetivo	Revisar los indicadores clave de desempeño propuestos para la definición de los niveles de prioridad.
Temas tratados	Definición de los niveles de prioridades que serán asignados a los indicadores clave de desempeño.
Comentarios	<ul style="list-style-type: none"> - Se llegó al acuerdo de definir tres niveles de prioridad: alto, medio y bajo. - Se decidió que para el proyecto se van a abarcar aquellos indicadores clave de desempeño que obtenga un nivel alto de prioridad.
Cantidad de participantes	3

Nota. Fuente: elaboración propia. (2020).

Apéndice XIX. Grupo focal para la definición de requerimientos del *datamart* con administrador de proyectos, gerente general e ingenieros *DevOps*

Tabla 45. Grupo focal para la definición de requerimientos del *datamart*

Grupo focal	
Objetivo	Definir los requerimientos del diseño del Datamart.
Temas tratados	Requerimientos del diseño del Datamart.
Comentarios	<ul style="list-style-type: none"> - Se llegó al acuerdo que se debe utilizar la misma arquitectura RDS de AWS de los otros sistemas de la organización. - Las bases de datos deben utilizar el mismo motor de MySQL, versión 8.0. - Las bases de datos deben permitir la conexión a ellas por medio de la IP como lo hacen los demás sistemas. - Se debe implementar un área de trabajo para no aumentar la carga de las demás bases de datos. - Para el diseño del Datamart se debe usar la herramienta Navicat.
Cantidad de participantes	5

Nota. Fuente: elaboración propia. (2020).

Apéndice XX. Grupo focal con administrador de proyectos e ingenieros *DevOps* para la definición de requerimientos relacionados al desarrollo de los procesos *ETL*.

Tabla 46. Grupo focal para la definición de requerimientos relacionados al desarrollo de los procesos *ETL*

Grupo focal	
Objetivo	Definir los requerimientos de la infraestructura técnica del ETL.
Temas tratados	Requerimientos de la infraestructura técnica del ETL.
Comentarios	<ul style="list-style-type: none"> - Se definió que la ejecución de los procesos se debe llevar a cabo en un servidor Linux Ubuntu como los otros sistemas. - Los ETL se deben calendarizar para que corran cada 24 horas - La herramienta debe ser fácil de utilizar para que el proyecto no se prolongue. - La herramienta de ETL debe tener un buen soporte de la industria, es decir, una buena comunidad de usuarios y un buen soporte por parte de los proveedores. - La herramienta de ETL debe tener la capacidad de conectarse a las bases de datos actuales, manejando MySQL, versión 8.0.

	<ul style="list-style-type: none">- Unos de los aspectos más importantes para seleccionar la herramienta es el costo.- Uno de los ingenieros de DevOps hizo la recomendación de utilizar Apache Kafka.
Cantidad de participantes	4

Nota. Fuente: elaboración propia. (2020).

Apéndice XXI. Grupo focal con administrador de proyectos y gerente general para para la definición de requerimientos relacionados las visualizaciones.

Tabla 47. Grupo focal para para la definición de requerimientos de relacionados las visualizaciones

Grupo focal	
Objetivo	Definir los requerimientos correspondientes a las visualizaciones de información
Temas tratados	Requerimientos de visualizaciones de información.
Comentarios	<p>Se documentaron una serie de requerimientos relacionados a las visualizaciones:</p> <ul style="list-style-type: none"> - La aplicación de visualización debe permitir al menos cinco usuarios de lectura y un usuario administrador o de escritura. - Las visualizaciones deben ser interactivas y permitir el filtrado de la información a los usuarios. - La aplicación de visualización debe tener acceso a los datos provenientes del almacén de datos. - La herramienta seleccionada debe permitir la calendarización del envío por correo electrónico de <i>dashboards</i>.

	<ul style="list-style-type: none">- La herramienta de visualización debe permitir el refrescamiento de la información de manera automatizada al menos cada 24 horas.- El acceso de la herramienta de visualización debe poder limitarse por una dirección IP privada, esto permite una conexión directa y segura con la instancia RDS del almacén de datos.- Debe considerarse el costo como el criterio de mayor peso en la escogencia de la herramienta.
Cantidad de participantes	3

Nota. Fuente: elaboración propia. (2020).

Apéndice XXII. Bitácora de reuniones y comunicaciones con la organización.

Bitácora de reuniones y comunicaciones con la organización

Proyecto: Desarrollo de solución de inteligencia de negocios para la interpretación, análisis y evaluación de indicadores clave de desempeño relacionados al procesamiento de criptomonedas
Caso: Progressio Digital.

Estudiante: José Andrés Barboza González

Contraparte de la organización: Edgar Cordero

Información general de la reunión			
Reunión No.	01	Fecha	17/02/2020
Objetivo de la reunión	Realizar la reunión inicial con la profesora tutora y la contraparte de la organización.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza• Jacqueline Solís• Federico López-Calleja• Edgar Cordero		

Información general de la reunión			
Reunión No.	02	Fecha	07/04/2020
Objetivo de la reunión	Analizar la situación actual de los sistemas.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza González• Edgar Cordero		

Información general de la reunión			
Reunión No.	03	Fecha	15/04/2020
Objetivo de la reunión	Revisar los indicadores clave de desempeño propuestos para la definición de los niveles de prioridad.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza González• Edgar Cordero		

Información general de la reunión			
Reunión No.	04	Fecha	18/05/2020
Objetivo de la reunión	Definir los requerimientos del diseño del Datamart.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza González• Edgar Cordero• Warren Castro• Ricardo Rodríguez		

Información general de la reunión			
Reunión No.	05	Fecha	18/05/2020
Objetivo de la reunión	Realizar la segunda reunión con la tutora y la contraparte de la organización.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza González• Edgar Cordero• Jacqueline Solís• Federico López-Calleja		

Información general de la reunión			
Reunión No.	06	Fecha	22/05/2020
Objetivo de la reunión	Definir los requerimientos de la infraestructura técnica de ETL.		
Participantes	<ul style="list-style-type: none">• Edgar Cordero• Warren Castro• Ricardo Rodríguez• José Andrés Barboza González		

Información general de la reunión			
Reunión No.	07	Fecha	04/06/2020
Objetivo de la reunión	Definir los requerimientos correspondientes a las visualizaciones de información		
Participantes	<ul style="list-style-type: none">• Edgar Cordero• Warren Castro• Ricardo Rodríguez• José Andrés Barboza González• Federico López-Calleja		

Información general de la reunión			
Reunión No.	08	Fecha	26/06/2020
Objetivo de la reunión	Realizar la última presentación del trabajo final de graduación con profesora tutora y la contraparte de la organización.		
Participantes	<ul style="list-style-type: none">• José Andrés Barboza González• Edgar Cordero• Jacqueline Solís• Federico López-Calleja		

Yo, Edgar Cordero Sancho con cédula de identidad 304040691, en pleno uso de mis facultades legales e intelectuales, como contraparte de la organización Progressio Digital, por este medio indico que las minutas resumidas en el presente documento fueron ejecutadas durante el desarrollo del trabajo final de graduación a cargo del estudiante, José Andrés Barboza González, en la organización.



Firma de la contraparte de Progressio Digital



Firma del estudiante del ITCR

Apéndice XXIII. Minuta de reunión 02 con la organización

Minuta de reunión			
Número de reunión	02		
Fecha	7/04/2020		
Lugar	Llamada virtual (Slack)		
Objetivo de la reunión	Analizar la situación actual de los sistemas.		
Participantes	Presentes: Edgar Cordero y José Andrés Barboza		
	Ausentes: -		
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1	Generalidades de los colaboradores de la organización	Soy el administrador de proyectos, me encargo de dirigir la ejecución de los proyectos y de aplicar la metodología SCRUM, en la cual se definen tareas, se priorizan, se analizan las arquitecturas y se ejecutan, etc.	No aplica.
2	Informes actuales	Actualmente, se obtiene información relacionada a las transacciones de depósito y retiro del sistema de procesamiento de criptomonedas y del sistema de monedero de criptomonedas relacionadas a lo que es el proceso de compras de criptomonedas por medio de tarjetas de crédito y también se monitorea información relacionada a la ejecución de los retiros planificados.	Mediante el desarrollo del TFG se debe permitir realizar análisis visuales de las transacciones de los sistemas por períodos de tiempo específicos.
3	Bases de datos utilizadas	Se utilizan consultas de bases de datos para la revisión del sistema de procesamiento de criptomonedas y otro para el sistema de monedero de criptomonedas.	Tomar como referencia las bases de datos operacionales.
Próxima Reunión			
Temas a tratar	Fecha	Convocados	
Revisar los indicadores clave de desempeño propuestos para la definición de los niveles de prioridad.	15/04/2020	Edgar Cordero, Federico López-Calleja y José Andrés Barboza	

Apéndice XXIII. Minuta de reunión 03 con la organización

Minuta de reunión			
Número de reunión	03		
Fecha	15/04/2020		
Lugar	Llamada virtual (Slack)		
Objetivo de la reunión	Revisar los indicadores clave de desempeño propuestos para la definición de los niveles de prioridad.		
Participantes	Presentes: Edgar Cordero, Federico López-Calleja y José Andrés Barboza		
	Ausentes: -		
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1	Niveles de prioridad	Se deben definir una serie de niveles de prioridad para identificar cuales son los indicadores clave de desempeño que se van a cubrir mediante la ejecución del trabajo final de graduación.	<ul style="list-style-type: none"> - Se llegó al acuerdo de definir tres niveles de prioridad: alto, medio y bajo. - Se decidió que para el proyecto se van a abarcar aquellos indicadores clave de desempeño que obtenga un nivel alto de prioridad.
Próxima Reunión			
Temas a tratar		Fecha	Convocados
Definir los requerimientos del diseño del Datamart.		18/04/2020	Edgar Cordero, Warren Castro, Ricardo Rodríguez y José Andrés Barboza

Apéndice XXIV. Minuta de reunión 04 con la organización

Minuta de reunión			
Número de reunión	04		
Fecha	18/05/2020		
Lugar	Llamada virtual (Slack)		
Objetivo de la reunión	Definir los requerimientos del diseño del Datamart.		
Participantes	Presentes: Edgar Cordero, Warren Castro, Ricardo Rodríguez y José Andrés Barboza		
	Ausentes: -		
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1	Requerimientos del diseño del Datamart.	Existen una serie de aspectos que se deben tomar en cuenta para que el diseño del Datamart se ajuste a las necesidades y arquitectura de la organización.	<ul style="list-style-type: none"> - Se llegó al acuerdo que se debe utilizar la misma arquitectura RDS de AWS de los otros sistemas de la organización. - Las bases de datos deben utilizar el mismo motor de MySQL, versión 8.0. - Las bases de datos deben permitir la conexión a ellas por medio de la IP como lo hacen los demás sistemas. - Se debe implementar un área de trabajo para no aumentar la carga de las demás bases de datos. - Para el diseño del Datamart se debe usar la herramienta Navicat
Próxima Reunión			
Temas a tratar		Fecha	Convocados
Requerimientos de la infraestructura técnica del ETL.		22/05/2020	Edgar Cordero, Warren Castro, Ricardo Rodríguez y José Andrés Barboza

Apéndice XXV. Minuta de reunión 05 con la organización

Minuta de reunión			
Número de reunión	05		
Fecha	22/05/2020		
Lugar	Llamada virtual (Slack)		
Objetivo de la reunión	Definir los requerimientos de la infraestructura técnica de ETL.		
Participantes	Presentes: Edgar Cordero, Warren Castro, Ricardo Rodríguez y José Andrés Barboza		
	Ausentes: -		
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1	Requerimientos de la infraestructura técnica del ETL.	Uno de los ingenieros de DevOps hizo la recomendación de utilizar Apache Kafka para el desarrollo del ETL.	<ul style="list-style-type: none"> - Se definió que la ejecución de los procesos se debe llevar a cabo en un servidor Linux Ubuntu como los otros sistemas. - Los ETL se deben calendarizar para que corran cada 24 horas - La herramienta debe ser fácil de utilizar para que el proyecto no se prolongue. - La herramienta de ETL debe tener un buen soporte de la industria, es decir, una buena comunidad de usuarios y un buen soporte por parte de los proveedores. - La herramienta de ETL debe tener la capacidad de conectarse a las bases de datos actuales, manejando MySQL, versión 8.0. - Unos de los aspectos más importantes para seleccionar la herramienta es el costo.
Próxima Reunión			
Temas a tratar	Fecha	Convocados	
Definir los requerimientos correspondientes a las visualizaciones de información	4/06/2020	Edgar Cordero, Warren Castro, Ricardo Rodríguez, Federico López-Calleja y José Andrés Barboza	

Apéndice XXVI. Minuta de reunión 06 con la organización

Minuta de reunión			
Número de reunión		06	
Fecha		04/06/2020	
Lugar		Llamada virtual (Slack)	
Objetivo de la reunión		Definir los requerimientos correspondientes a las visualizaciones de información	
Participantes		Presentes: Edgar Cordero, Warren Castro, Ricardo Rodríguez, Federico López-Calleja y José Andrés Barboza	
		Ausentes: -	
Temas Tratados			
No.	Asunto	Comentarios	Acuerdos
1	Requerimientos de visualizaciones	Los acuerdos definidos son de gran relevancia para configurar la herramienta de visualización según la arquitectura de la organización.	<ul style="list-style-type: none"> - La aplicación de visualización debe permitir al menos cinco usuarios de lectura y un usuario administrador o de escritura. - Las visualizaciones deben ser interactivas y permitir el filtrado de la información a los usuarios. - La aplicación de visualización debe tener acceso a los datos provenientes del almacén de datos. - La herramienta seleccionada debe permitir la calendarización del envío por correo electrónico de <i>dashboards</i>. - La herramienta de visualización debe permitir el refrescamiento de la información de manera automatizada al menos cada 24 horas. - El acceso de la herramienta de visualización debe poder limitarse por una dirección IP privada, esto permite una conexión directa y segura con la instancia RDS del almacén de datos. - Debe considerarse el costo como el criterio de mayor peso en la escogencia de la herramienta. -
Próxima Reunión			
Temas a tratar		Fecha	Convocados
			Edgar Cordero, Warren Castro, Ricardo Rodríguez, Federico López-Calleja y José Andrés Barboza

ANEXOS

Anexo I.

Cuadrante mágico de Gartner para herramientas de integración de datos

Figure 1. Magic Quadrant for Data Integration Tools



Source: Gartner (August 2019)

Fuente: (Talend, 2019)

Anexo II.

Cuadrante mágico de Gartner para herramientas de visualización de datos



Fuente: (Cuadrante Mágico de Gartner 2020 y 2019 Analítica e Inteligencia de Negocio (BI), 2020)

Anexo III.

Ejemplo de matriz bus de dimensiones de alto nivel para un almacén de datos

Business Process / Event	Common Dimensions								
	Date	Policyholder	Coverage	Covered Item	Agent	Policy	Claim	Claimant	Payee
Underwriting Transactions	X	X	X	X	X	X			
Policy Premium Billing	X	X	X	X	X	X			
Agents' Commissions	X	X	X	X	X	X			
Claims Transactions	X	X	X	X	X	X	X	X	X

Fuente: (Kimball, Ross, Thorntwaite, Mundy, & Becker, 2006)

GLOSARIO

A continuación, se presenta una lista de términos con sus respectivos significados, su objetivo es lograr un adecuado entendimiento de este documento dentro del contexto del desarrollo de la solución de IN.

- **Blockchain:** una red distribuida de bloques cifrados, en él se registran transacciones que son propiedad de todos los nodos de la red, y no solamente de una entidad centralizada.
- **Criptomoneda:** es una moneda digital diseñada para funcionar como medio de intercambio. Utiliza la criptografía para asegurar y verificar transacciones, así como para controlar la creación de nuevas unidades de una criptomoneda particular.
- **Dinero fiat:** es dinero de curso legal cuyo valor no deriva del hecho de ser un bien físico o mercancía, sino por ser emitido y respaldado por un gobierno. (Binance Academy, s.f.)
- **Framework de desarrollo:** estructura de soporte definida, en la cual un proyecto de software puede ser organizado y desarrollado. (Alegsa.com, 2010)
- **Hardware:** Conjunto de elementos físicos que componen una solución de inteligencia de negocios.
- **Ingenieros devops:** ingenieros en sistemas encargados de aprovisionar los recursos o infraestructura requeridos para los proyectos de software, asegurar la integración y despliegue continuo, monitorización continua de los procesos de desarrollo y operaciones, etc. (Rico, 2019)
- **Software:** Conjunto de elementos tecnológicos no físicos que componen una solución de inteligencia de negocios.
- **Transacción en el blockchain:** Una transacción es una transferencia de valor entre las billeteras de Bitcoin que se incluye en el *blockchain*

BIBLIOGRAFÍA

- Acuña, L. D. (2018). *Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes. Caso: Empresa Áltica*.
- Alegsa.com. (02 de Julio de 2010). *Definición de Framework de desarrollo*. Obtenido de Alegsa: <http://www.alegsa.com.ar/Dic/framework.php>
- Binance Academy. (s.f.). *¿Qué es el Dinero Fíat?* Obtenido de Binance Academy: <https://www.binance.vision/es/economics/what-is-fiat-currency>
- Bitcoin. (2020). *Some Bitcoin words you might hear*. Obtenido de Bitcoin: www.bitcoin.org
- Bitcoin Cash. (2020). *Peer-to-Peer Electronic Cash*. Obtenido de Bitcoin Cash: <https://www.bitcoincash.org/index.html>
- Bonilla-Jimenez, F., & Escobar, J. (2011). Grupos focales: una guía conceptual y metodológica. En F. Bonilla, & J. Escobar, *CUADERNOS HISPANOAMERICANOS DE PSICOLOGÍA* (págs. 51-67).
- Business Analyst Learnings website. (2014). *Weighted Scoring Model: A Technique for Comparing Software Tools*. Obtenido de Business Analyst Learnings website: <https://businessanalystlearnings.com/ba-techniques/2014/2/27/weighted-scoring-model-a-technique-for-comparing-software-tools>
- Cresswell, J. (2014). *Research design: qualitative, quantitative, and mixed methods approaches. Cuarta Edición*. California: SAGE Publications, Inc.
- Cuadrante Mágico de Gartner 2020 y 2019 Analítica e Inteligencia de Negocio (BI)*. (2020). Obtenido de Software & Hardware: <https://softwarehardware.com/software/cuadrante-magico-de-gartner-2019-analitica-e-inteligencia-de-negocio-bi/>
- Dolader, C., Bel, J., & Muñoz, J. (2017). *LA BLOCKCHAIN: FUNDAMENTOS*. Catalunya: Universitat Politècnica de Catalunya.
- Ethereum. (10 de Febrero de 2020). *What is Ethereum?* Obtenido de Ethereum: <https://ethereum.org/what-is-ethereum/>
- Gartner. (2020). *Blockchain*. Obtenido de Gartner Glossary: <https://www.gartner.com/en/information-technology/glossary/blockchain>
- Gartner. (2020). *Business Intelligence (bi) Services*. Obtenido de Gartner Glossary: <https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi-services>
- Gupta, M. (2017). *Blockchain for Dummies IBM Limited Edition*. New Jersey: John Wiley & Sons, Inc.
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la Investigación*. Obtenido de Universidad Florencio del Castillo: <https://www.uca.ac.cr/wp-content/uploads/2017/10/Investigacion.pdf>
- Hileman, G., & Michel, R. (2017). *GLOBAL CRYPTOCURRENCY*.

- Inmon, W. (2002). *Building a Data Warehouse, Third Edition*. Obtenido de HMCM University of Technology and Education: <http://fit.hcmute.edu.vn/Resources/Docs/SubDomain/fit/ThayTuan/DataWH/Bulding%20the%20Data%20Warehouse%204%20Edition.pdf>
- Kimball, R., Ross, M., Thorntwaite, W., Mundy, J., & Becker, B. (2006). *The Kimball Group Reader: Relentlessly Practical Tools*. Wiley.
- Litecoin. (2020). *WHAT IS LITECOIN?* Obtenido de Litecoin: <https://litecoin.org/>
- López, A. (2012). *La investigación-acción*. Puerto Rico: Centro de Investigaciones Educativas, Facultad de Educación Universidad de Puerto Rico.
- Microsoft. (27 de Julio de 2019). *Procesamiento de transacciones en línea (OLTP)*. Obtenido de Microsoft: <https://docs.microsoft.com/es-es/azure/architecture/data-guide/relational-data/online-transaction-processing>
- Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*.
- Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *BITCOIN AND CRYPTOCURRENCY TECHNOLOGIES*. New Jersey: Prensa de la Universidad Princeton.
- Pérez, D. (2019). *Desarrollo de herramienta de inteligencia de negocios para el área de ventas. Caso: Equifax USA*.
- Progressio Digital. (2019). Obtenido de Progressio Digital: <http://www.progressiodigital.com/>
- Rauchs, M., Blandin, A., Klein, K., Pieters, G., Recanatini, M., & Zhang, B. (2018). *2ND GLOBAL CRYPTOASSET BENCHMARKING STUDY*.
- Rico, A. (17 de diciembre de 2019). *¿Qué es un ingeniero DevOps?* Obtenido de Ambit: <https://www.ambit-bst.com/blog/qu%C3%A9-es-un-ingeniero-devops>
- Sansu, G. (14 de Abril de 2012). *Inmon or Kimball: Which approach is suitable for your data warehouse?* Obtenido de ComputerWeekly: <https://www.computerweekly.com/tip/Inmon-or-Kimball-Which-approach-is-suitable-for-your-data-warehouse>
- Scheps, S. (2008). *Business Intelligence for Dummies*. Indiana: Wiley Publishing, Inc.
- Talend . (2020). *¿En qué consiste un proceso de ETL (Extraer, Transformar y Cargar)?* Obtenido de Talend Resource Center: <https://es.talend.com/resources/what-is-etl/>
- Talend. (2019). *2019 Gartner Magic Quadrant for Data Integration Tools*. Obtenido de Talend: <https://es.talend.com/resources/2019-gartner-magic-quadrant-data-integration-tools/>
- Top 100 Cryptocurrencies by Market Capitalization*. (11 de 2019). Obtenido de CoinMarketCap: <https://coinmarketcap.com/>
- Ulate, I., & Vargas, E. (2016). *Metodología para elaborar una tesis*. San José: Editorial Universidad Estatal a Distancia.
- Adams, K. (2020). *Power BI vs Tableau vs Amazon QuickSight – Which is Best for Your Organization?*. Recuperado de <https://key2consulting.com/power-bi-vs-tableau-vs-amazon-quicksight-which-is-best-for-your-organization/>

- Bobriakov, K. (2018). A Comparative Analysis of Top 6 BI and Data Visualization Tools in 2018. Recuperado de <https://medium.com/activewizards-machine-learning-company/a-comparative-analysis-of-top-6-bi-and-data-visualization-tools-in-2018-658490665973>
- Data Estudio. (2020). Conectarse a MySQL. Recuperado de <https://support.google.com/datastudio/answer/7088031?hl=es>.
- Amazon. (2020). Precios de Amazon QuickSight. Recuperado de <https://aws.amazon.com/es/quicksight/pricing/?nc=sn&loc=4>
- Microsoft. (2018). Whitelist PowerBI service FQDNs or IP ranges for Connectors to work. Recuperado de <https://community.powerbi.com/t5/Power-Query/Whitelist-PowerBI-service-FQDNs-or-IP-ranges-for-Connectors-to/td-p/336847>.
- Microsoft. (2020). Add Power BI URLs to your allow list. Recuperado de <https://docs.microsoft.com/en-us/power-bi/admin/power-bi-whitelist-urls>
- Radacad. (2020). The READ-ONLY License for Power BI. Recuperado de <https://radacad.com/the-read-only-license-for-power-bi>.
- Tableau. (2020). Tableau data Managment. Recuperado de <https://www.tableau.com/pricing/teams-orgs#server>.