

INSTITUTO TECNOLÓGICO DE COSTA RICA  
ESCUELA DE ADMINISTRACIÓN DE EMPRESAS

**“¿CÓMO ELEGIR INVERSIONES QUE SE AJUSTAN A SUS NECESIDADES? UNA  
PROPUESTA DE CATEGORIZACIÓN DE LOS FONDOS DE INVERSIÓN PARA  
MERCADOS EMERGENTES LATINOAMERICANOS, CASO COSTA RICA”**

Proyecto Final de Graduación para optar por el título de Administración de Empresas con el  
grado académico de Bachillerato

MAIKO ZHENG GUO

Cartago, noviembre, 2021

**ESCUELA DE ADMINISTRACIÓN DE EMPRESAS**  
**ACTA DE DEFENSA DE TRABAJO FINAL DE GRADUACIÓN (TFG)**

Al ser las diecisiete horas del veinticinco de noviembre del dos mil veintiuno de manera remota, se procedió a la defensa pública y oral del Trabajo Final de Graduación (TFG) del estudiante: Zheng Guo Maiko, carnet 2018085200.

Título del TFG: “¿CÓMO ELEGIR INVERSIONES QUE SE AJUSTAN A SUS NECESIDADES? UNA PROPUESTA DE CATEGORIZACIÓN DE LOS FONDOS DE INVERSIÓN PARA MERCADOS EMERGENTES LATINOAMERICANOS, CASO COSTA RICA”

El Tribunal Evaluador está integrado por las siguientes personas: Ronald Alvarado Cordero, quien preside y Gustavo Cubillo Salas.

El estudiante realizó la Defensa Oral de su trabajo final de graduación, después de la cual el Tribunal Evaluador hizo las preguntas pertinentes sobre aspectos relacionados con el tema.

Terminada la defensa se determina que el estudiante:

- (  ) Aprobó satisfactoriamente su Trabajo Final de Graduación  
(  ) No aprobó su Trabajo Final de Graduación

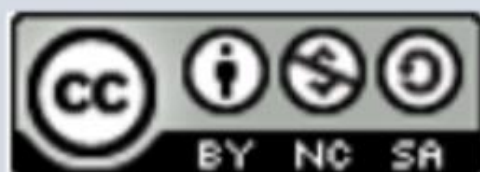
RONALD GERARDO ALVARADO CORDERO (FIRMA)  
Firmado digitalmente por RONALD GERARDO ALVARADO CORDERO (FIRMA)  
Fecha: 2021.11.26 08:16:43 -0500

Firma Evaluador 1

MIGUEL GUSTAVO CUBILLO SALAS (FIRMA)  
Firmado digitalmente por MIGUEL GUSTAVO CUBILLO SALAS (FIRMA)  
Fecha: 2021.11.26 00:17:49 -0500

Firma Evaluador 2

PAULA CECILIA ARZADUN  
Firmado digitalmente por PAULA CECILIA ARZADUN  
Fecha: 2021.11.30  
Dra. Paula Arzadun, Coordinadora



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).

**¿Cómo elegir inversiones que se ajustan a sus necesidades? Una propuesta de categorización de los fondos de inversión para mercados emergentes latinoamericanos, caso Costa Rica**

**Resumen**

Los fondos de inversión son clasificados frecuentemente de acuerdo con su objetivo de inversión; sin embargo, esta metodología no garantiza que los productos conformados en un mismo grupo cuenten con un nivel de rendimiento, riesgo y desempeño similares. Por esta razón, en el presente estudio se propone un método de clasificación empleando la técnica del análisis de conglomerado a 92 fondos de inversión del mercado costarricense. Como resultado, se logró separar los 92 fondos de inversión en 8 grupos diferentes mediante el método de agrupamiento llamado partición alrededor de medoids (PAM). Esta propuesta puede facilitarle a inversionistas y otros actores una mejor planificación estratégica y toma de decisión desde la perspectiva financiera.

*Palabras clave:* Fondos de inversión, análisis de conglomerado, partición alrededor de medoids (PAM), Costa Rica.

**Abstract**

Mutual funds are frequently classified according to their investment objective; however, this methodology doesn't guarantee that the products formed in the same group have a similar level of return, risk, and performance. For this reason, this study proposes a classification method using the conglomerate analysis technique for the 92 mutual funds in the Costa Rican market. As a result, it was possible to separate the 92 mutual funds into 8 different groups by the grouping method called partition around medoids (PAM). This proposal can facilitate investors and other actors better strategic planning and decision making from a financial perspective.

*Keywords:* Mutual funds, cluster analysis, partition around medoids (PAM), Costa Rica.

**Cumplimiento al Objetivo de Desarrollo Sostenible (ODS)**

Objetivo 17: fortalecer los medios de implementación y revitalizar la alianza mundial para el desarrollo sostenible / finanzas.

Meta vinculada: 17.5 Adoptar y aplicar sistemas de promoción de las inversiones en favor de los países menos adelantados.

### **Introducción**

En los últimos 3 años, el número de cuentas abiertas de los fondos de inversión ha incrementado significativamente en Costa Rica; en un 39% según las estadísticas de la industria (Superintendencia General de Valores [SUGEVAL], 2021), lo cual refleja que cada vez más, los inversionistas adoptan estos productos para generar ingresos pasivos a sus ahorros. Asimismo, el mercado costarricense se encuentra en un ritmo de desarrollo y crecimiento constante, que llegó a un aumento de aproximadamente 50% de los activos netos invertidos en el sector de los fondos de inversión en el año 2021 con respecto al año 2018 (SUGEVAL, 2021). Además la cantidad de opciones disponibles llegó a 102 fondos de inversión, por lo tanto, ante estos crecimientos, los consumidores podrían enfrentar dificultades para seleccionar correctamente aquellos productos que se ajustan a sus necesidades personales.

En el mercado costarricense, los fondos de inversión están regulados por la Superintendencia General de Valores (SUGEVAL) y se encuentran clasificados principalmente por el objetivo de inversión. Sin embargo, actualmente se carece de una metodología de clasificación que permita identificar cuáles fondos son homogéneos de acuerdo a sus características de rendimiento, riesgo y desempeño para que los inversionistas puedan determinar cuáles productos son aptos a sus necesidades personales y así diversificar el riesgo.

Tradicionalmente, los fondos de inversión son clasificados de acuerdo con su objetivo de inversión (Moreno et al., 2006), pero los investigadores del mercado internacional afirman que este método no es una manera óptima para que los inversionistas puedan diversificar su portafolio, debido a que no garantiza que una categoría sea suficientemente distinta en comparación con otra categoría en términos financieros, lo cual dificulta la diversificación (Marathe y Shawky, 1999; Moreno et al., 2006). Por esta razón, en el mercado exterior, el método de análisis de conglomerado es empleado para clasificar los productos de inversión de acuerdo a su nivel de homogeneidad, y así facilitar la estrategia de diversificación (Kumar y Rama, 2011; Shanmugham, 2011; Menardi y Lisi, 2015; Sakakibara et al., 2015).

Debido a la falta de un sistema de clasificación adecuado que tome en cuenta el nivel de similitud en términos financieros entre grupos de fondos de inversión en el mercado costarricense, se considera relevante proponer un método de clasificación basado en el análisis de conglomerado. Por este motivo, el objetivo del presente estudio es desarrollar una propuesta de agrupamiento para los fondos de inversión costarricenses, que permita identificar cuales productos en específicos son homogéneos entre sí desde la perspectiva de los indicadores de rendimiento, riesgo y desempeño.

La finalidad es facilitarle a los inversionistas una mejor selección de los productos y al mismo tiempo proporcionarle opciones similares en los indicadores de rendimiento, riesgo y desempeño para la diversificación del portafolio. De igual manera, servirá como una herramienta para simplificar la toma de decisiones financieras, como el establecimiento de estrategias de inversión, determinar opciones que logran satisfacer sus expectativas y la comprensión de las características financieras que conforma la cartera. Finalmente, el resultado podrá servir como una fuente de referencia para las sociedades administradores de fondos de inversión, permitiéndoles revisar cómo se encuentra actualmente el desempeño de su producto en comparación con los similares de la competencia, analizando la diferencia de los ratios financieros del mismo conglomerado.

### **Agrupamiento de Inversiones: Una Revisión Literaria**

El análisis de conglomerado, es una herramienta estadística multivariante que intenta identificar grupos de casos relativamente homogéneos en función de un conjunto de variables seleccionadas mediante la estimación de la distancia o el nivel de similitud entre ellos. En otras palabras, busca clasificar objetos en grupos de tal manera que el grado de asociación entre dos objetos es máximo si pertenece al mismo grupo y mínimo en caso contrario. (González y Álvarez, 2011; Shanmugham, 2011). De acuerdo con Anderberg (2014), la técnica ha sido ampliamente utilizada en distintos campos de la ciencia, específicamente en las áreas de la biología, diagnóstico clínico, criminología, geología, inteligencia artificial, políticas, economías, investigación de mercado, inversiones, entre otros.

En el ámbito de las inversiones, el análisis de conglomerado ha sido usado para clasificar las inversiones en acciones (Romesburg, 2004). Por ejemplo, un estudio realizado en el mercado de Turquía por Tekin y Gümüş (2017), identificaron que con la aplicación del análisis de

conglomerado, las 88 acciones del índice BIST 100 pueden ser clasificadas en 12 grupos diferentes, simplificando así la toma de decisiones a los inversionistas. Otro estudio realizado por Bjärkby y Grägg (2019), permitió construir un portafolio de diversificación cruzando diferentes sectores financieros. Su metodología consistía en aplicar un análisis de conglomerado a las 305 acciones de Suecia para clasificar estos productos en 9 grupos diferentes. Adicionalmente seleccionaron las 2 mejores acciones de los 4 grupos más representativos del mercado para establecer una estrategia de inversión de bajo riesgo y rendimiento estable.

Asimismo, Peachvanish (2016) empleó el método de análisis de conglomerado para identificar el grupo de acciones que tienen las mejores características en términos de los indicadores de tendencias e impulso, para así detectar aquellas acciones que son más probables de superar el rendimiento del mercado durante un periodo de corto tiempo. Sus resultados muestran que si se utiliza el método propuesto para seleccionar acciones, se obtiene un rendimiento superior al mercado a largo plazo. Recientemente, Medellu y Nugraha (2021) aplicaron el análisis de conglomerado de  $k$  medias para agrupar las acciones de acuerdo a su nivel de volatilidad, liquidez y la capitalización de mercado. Como resultado, logran desarrollar una herramienta para facilitarle a los inversionistas la comprensión de estas características financieras y la selección basada en los datos históricos.

De igual manera, el método de análisis de conglomerado ha sido muy utilizado a nivel internacional para categorizar los fondos de inversión. Tamara y Revina (2015), aplicaron este método para clasificar los 690 fondos de inversión del mercado de Indonesia con base en las variables del precio y rendimiento de diferentes periodos. Como resultado logran agrupar estos productos en 2 categorías conformados por productos de distintas industrias. Acharya y Sidana (2007) aplicaron el análisis de conglomerado a los 100 fondos de inversión de 6 diferentes sectores de la India, con el fin de facilitarle a los inversionistas con deseos de invertir pero desinformados, la construcción de un portafolio diversificado cruzando fondos de inversión de diferentes sectores y objetivos de inversión.

Seguidamente, Menardi y Lisi (2015), aplicaron este método para categorizar los 1436 fondos de inversión del mercado estadounidense mediante 18 indicadores financieros de retorno, riesgo y desempeño. Al final clasificaron los productos en 6 grupos diferentes para evaluar su nivel de consistencia con los objetivos declarados. Asimismo, Sakakibara et al., (2015) aplicaron el

análisis de conglomerado para clasificar los 551 fondos de inversión del mercado japonés en 2 categorías con base en las características de la industria, tasa de inversión, activos netos, dividendos netos y la tarifa de transacción. Los autores determinaron que el método propuesto fue más adecuado que las categorías proporcionadas por el portal financiero Morningstar Inc.

Dentro del análisis de conglomerado existe gran variedad de métodos, pero todos persiguen la misma finalidad de generar agrupaciones según el nivel de similitud entre los objetos. Para estudiar los fondos de inversión, se ha encontrado que la mayoría de los investigadores emplearon el método jerárquico y el método de  $k$  medias. (e.j Acharya y Sidana, 2007; Shanmugham, 2011; Menardi y Lisi, 2015; Tamara y Revina, 2015). Sin embargo, también se han usado otros métodos, aunque en menor medida, como mapa autoorganizado (Moreno et al., 2016) y agrupamiento espectral (Sakakibara et al., 2015).

Los resultados que se obtienen de un análisis de conglomerado requieren de la validación de propiedades como: densidad, tamaños y formas de conglomerados, separabilidad de conglomerados y la solidez de la clasificación. (Yatskiv, 2003, como se citó en Yatskiv y Gusarova, 2005). Por lo general, existen tres enfoques principales para la validación de conglomerado (Gordon, 1999, como se citó en Yatskiv y Gusarova, 2005). En primer lugar, se encuentra la validación externa, este método fue utilizado por la mayoría de los investigadores del mercado internacional para los estudios de fondos de inversión (e.j Marathe y Shawky, 1999; Sakakibara et al., 2015).

Sin embargo, como en el mercado costarricense no posee una agrupación existente similar, la validación externa no es posible para el presente estudio y se optará por la validación interna y la validación relativa con diferentes parámetros de  $k$ . En la validación interna sólo se utilizan los datos de entrada para evaluar la calidad de la clasificación, mientras que en la validación relativa se comparan varias clasificaciones diferentes de un conjunto de datos utilizando el mismo algoritmo de clasificación pero con diferentes parámetros. En cuanto a la validación interna, los estudios de conglomerado en inversiones han utilizado el índice de Silueta, índice de Dunn, índice de RMSSTD, índice de  $r$  cuadrado, índice de la validez SD, índice de Frey, índice de Davies-Boldin, índice de McClain, índice C y estadística de Hubert modificada, (Das, 2003; Momeni et al., 2015; Isakov, 2019).



Finalmente, en cuanto a la evaluación de los fondos de inversión, se ha encontrado que las variables como: rendimiento, desviación estándar, ratio de Treynor, ratio de Sharpe, coeficiente de beta y alfa de Jensen son las más utilizadas a nivel internacional por los investigadores, tanto para la evaluación financiera como para el análisis de conglomerado (e.j Acharya y Sidana, 2007; Kumar y Rama, 2011; Shanmugham, 2011; Menardi y Lisi, 2015; Sunnetha y Latha, 2020). También se ha destacado el uso de otros indicadores en menor medida, caso del índice de Omega, exposición de ganancia potencial, ganancia de cola, pérdida de cola esperada y ratio de sortino (e.j Menardi y Lisi, 2015), r cuadrado (e.j. Acharya y Sidana, 2007; Shanmugham, 2011) y curtosis (e.j. Moreno et al., 2006). Por otra parte, se ha encontrado que el mercado costarricense posee cierta similitud ya que por lo general se utiliza el rendimiento, desviación estándar, ratio de Sharpe, ratio de Treynor, Alfa de Jensen, CAPM y coeficiente de beta para la evaluación financiera de los productos (e.j Bolaños et al., 2017; Gamboa et al., 2019).

## **Metodología**

### **Datos**

Este estudio se realizó con base en los datos proporcionados por la Superintendencia General de Valores (SUGEVAL), específicamente de las estadísticas diarias emitidas sobre la industria de los fondos de inversión, para el período comprendido entre el 1 de agosto del 2018 al 31 de julio del 2021. El tamaño de la muestra utilizada fue de 92 fondos de inversión que cuentan con todas las informaciones disponibles sobre el rendimiento, desviación estándar y la calificación del riesgo que conforman parte de las variables del estudio. Asimismo, se complementará con otras variables financieras como el ratio de Sharpe, Treynor, el modelo de valuación de activos de capital (CAPM) y la aceptación de CAPM que fueron calculados utilizando los datos disponibles. Finalmente, el software R Studio fue utilizado para procesar con el análisis de los datos.

### **Variables**

Las variables utilizadas para el presente estudio fueron enfocadas principalmente para la evaluación de los fondos de inversión en términos del rendimiento, riesgo y desempeño. Tomando en cuenta que el rendimiento, el CAPM y la aceptación de CAPM son utilizados para evaluar el aspecto del rendimiento, la desviación estándar y la calificación del riesgo son utilizados para

evaluar el aspecto del riesgo y finalmente, el ratio de Sharpe y Traynor son utilizados para evaluar el desempeño de los fondos de inversión. A continuación, se mostrará detalladamente cada una de estas.

### ***Rendimiento***

Se define como el nivel de beneficios producto de una inversión. Se obtiene sumando la ganancia de capital y los ingresos corrientes (Gitman y Joehnk, 2009).

### ***Desviación Estándar***

Mide la dispersión de los rendimientos en torno al rendimiento promedio o esperado de un activo. (Gitman y Joehnk, 2009). Su respectivo cálculo se mostrará en la ecuación (1).

$$Desviación\ estándar(s) = \sqrt{\sum_{j=1}^n \frac{(r_j - \bar{r})^2}{n - 1}}$$
(1)

En la ecuación (1),  $r_j$  representa el rendimiento del resultado,  $\bar{r}$  es el rendimiento promedio o esperado y  $n$  el número total de resultados.

### ***Ratio de Sharpe***

Mide el rendimiento esperado por cada unidad de riesgo soportado, en comparación con un instrumento de inversión libre de riesgo (Sharpe 1994). La fórmula de cálculo del ratio de Sharpe se mostrará en la ecuación (2).

$$Ratio\ de\ Sharpe = \frac{r_p - R_f}{S_p}$$
(2)

En la ecuación (2),  $r_p$  representa el rendimiento total de la cartera,  $R_f$  es la tasa libre de riesgo y  $S_p$  representa la desviación estándar del rendimiento de la cartera.

### ***Ratio de Treynor***

Mide el exceso de rendimiento por cada unidad de riesgo sistemático asumido mediante la beta de la cartera (Verma y Hirpara, 2016). En la ecuación (3) se mostrará su fórmula de cálculo.

$$\text{Ratio de Treynor} = \frac{r_p - R_f}{\beta_p} \quad (3)$$

En la ecuación (3),  $r_p$  representa el rendimiento total de la cartera,  $R_f$  es la tasa libre de riesgo y  $\beta_p$  representa la beta de la cartera.

### ***Modelo de Valuación de Activos de Capital (CAPM)***

Mide la tasa de rendimiento teóricamente requerida para un activo mediante una relación lineal y positiva entre la rentabilidad de un activo y su beta con el mercado (Nieto y Rubio, 2002; Bolaños et al., 2017). Para su estimación, se puede dividirse en dos partes: la tasa de rendimiento libre de riesgo y la prima de riesgo (Gitman y Zutter, 2016). En la ecuación (4) se mostrará su fórmula de cálculo.

$$CAPM(r_i) = R_f + [\beta_i(r_m - R_f)] \quad (4)$$

En la ecuación (4),  $r_i$  representa el rendimiento requerido del activo  $i$ ,  $R_f$  es la tasa libre de riesgo,  $\beta_i$  representa la beta del activo  $i$  y  $r_m$  es el rendimiento del mercado.

### ***Aceptación de CAPM***

Si el rendimiento ofrecido de un activo no alcanza o supera el rendimiento requerido, no es recomendable rezzalizar la inversión (Dayaratne et al., 2010). Por esta razón, se ha agregado esta variable para medir la eficiencia de los fondos de inversión. En la ecuación 5 se mostrará su fórmula de cálculo.

$$\text{Aceptación de CAPM} = r_p - CAPM \quad (5)$$

En la ecuación (5),  $r_p$  representa el rendimiento total de la cartera y CAPM la tasa de rendimiento teóricamente requerida del activo. Si la aceptación de CAPM da un valor igual o superior al 0, significa que el activo es aceptado por ofrecer un rendimiento igual o superior a lo requerido y estará codificado con un valor “1”, al contrario con un valor “0”.

### ***Calificación del Riesgo***

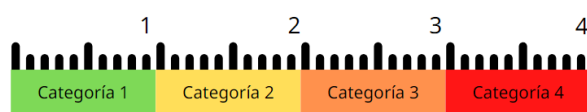
Según las estadísticas de SUGEVAL, los fondos de inversión del mercado costarricense son evaluados por la sociedad Calificadora de Riesgo Centroamericana S.A. (SCRiesgo). Según SCRiesgo (s.f), la calificación de los fondos de inversión está dividida en dos partes:

- Calificación de riesgo de mercado: es una evaluación de la sensibilidad del fondo de inversión a condiciones cambiantes en el mercado y se encuentra clasificado en 4 categorías, considerando que el nivel 1 es de baja sensibilidad hasta el nivel 4 que es muy alta sensibilidad.
- Calificación de fondos de inversión: es una evaluación sobre la calidad y diversificación de los activos del fondo, la capacidad para la generación de flujos, las fortalezas y debilidades de la administración. Los niveles están clasificados consecutivamente desde AAA, AA, A, BBB, BB, B, C, D, considerando que AAA es el nivel excelente que presenta la más alta probabilidad de cumplir con sus objetivos de inversión y una muy buena gestión para limitar su exposición al riesgo por factores inherentes.

Debido a que esta variable evalúa dos aspectos de riesgo y está clasificada en categorías diferentes, se ha aplicado una codificación para el presente estudio, como se muestran en la figura 1 y 2.

### **Figura 1**

#### *Codificación de la variable riesgo de mercado*

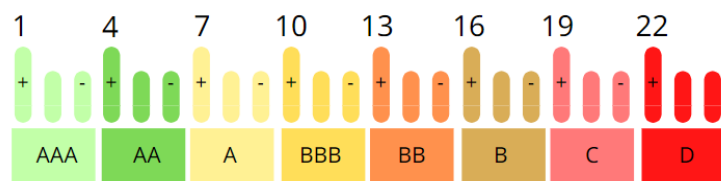


*Nota.* Para conocer las definiciones de cada categoría, consulte a la página:

<https://www.scriesgo.com/ScaleSummary>

**Figura 2**

*Codificación de la variable calificación de fondo de inversión*



*Nota. Para conocer las definiciones de cada tipo, consulte a la página:*

<https://www.sciesgo.com/ScaleSummary>

### **Método de Conglomerado**

Para el análisis de conglomerado aplicado a los fondos de inversión, se suelen emplear el método de conglomerado jerárquico y no jerárquico de  $k$  medias. De acuerdo con Menardi y Lisi (2015), el enfoque jerárquico comienza con una asignación de grupo diferente para cada objeto del análisis, luego, entre las agrupaciones que cuentan con una distancia más cercana se unirá hasta que quede un solo grupo final para formar una estructura jerárquica llamada dendrograma. Para detectar el nivel de similitud entre dos observaciones dentro de un análisis de conglomerado, se requiere calcular una medida de distancia. Según Madhulatha (2012), la medida más común es la distancia euclidiana que mide la distancia entre un punto  $x$  ( $x_1, x_2, \text{etc.}$ ) y un punto  $y$  ( $y_1, y_2, \text{etc.}$ ) en una línea recta. En la ecuación (6) se mostrará con mayor detalle la fórmula del cálculo.

$$\text{Distancia euclidiana} = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}$$

(6)

No obstante, cabe destacar que antes de procesar con la estimación de la distancia, todas las variables deben estar estandarizadas para no dar más importancia y peso a aquellas variables que cuentan con una escala de medida más amplia (Elguera, 2018).

Seguidamente, de acuerdo con Das (2003), el enfoque no jerárquico de  $k$  medias se trata de dividir una muestra de  $n$  entidades en  $k$  cantidad de grupos basado en una medida de distancia, que

podría ser euclidiana. El algoritmo asigna cada elemento al grupo que tiene el centroide más cercano mediante tres pasos:

1. Para empezar, se divide los elementos en  $k$  grupos iniciales. Asimismo, se podría especificar  $k$  centroides iniciales.
2. Seguidamente, se asigna el elemento al conglomerado cuyo centroide es más cercano utilizando una medida de distancia, generalmente la euclidiana. Después, volver a calcular el centroide para cada uno de los conglomerados.
3. Por último, repetir el paso 2 hasta que no realicen más asignaciones. Es decir, hasta que la solución no cambie más.

Para el presente estudio, el método jerárquico y no jerárquico de  $k$  medias fueron empleados para categorizar los fondos de inversión. Sin embargo, durante el proceso, se ha encontrado que las observaciones cuentan con gran cantidad de valores atípicos que causan dificultades utilizando los dos métodos mencionados. Por esta razón, para la verificación, se ha aplicado el análisis de valores atípicos multivariantes mediante la técnica LOF.

De acuerdo con Babaei et al., (2019), el LOF es un algoritmo basado en la densidad que otorga una puntuación a cada punto de datos basado en el grado en que este punto se desvía de su vecindario local. Por lo general, cuando el LOF es superior a 1, se sugiere considerarlo como un valor atípico, sin embargo, la decisión depende de cada caso del análisis. Para el presente estudio, se tomará en consideración aquellos fondos de inversión que cuentan con un LOF superior a 1 y que al mismo tiempo presentan una distancia alejada de las demás observaciones.

De esta manera, el método de partición alrededor de medoids (PAM) fue complementado para solucionar el problema. De acuerdo con Lopes y Gosling (2020), esta técnica de agrupamiento es similar a la  $k$  medias, pero más robusto a la presencia de valores atípicos; debido a que su técnica de partición está basada en el medoid. Según Elguera (2018), un medoid es el objeto de un grupo cuya disimilaridad media es mínima a los demás objetos del grupo, en otras palabras, es el punto más cercano al centro de todo el grupo.

Continuando con lo anterior, Elguera (2018) menciona que el algoritmo PAM clasifica las observaciones bajo 6 pasos:

1. Para empezar, se selecciona  $k$  objetos de los  $n$  puntos como el medoid inicial.

2. Seguidamente, se calcula la distancia media de disimilaridad:  $d_{ij} = d(O_j, O_i)$ . Tomando en consideración que  $O_j$  sea un objeto no seleccionado y  $O_i$  como un medoid (objeto seleccionado).
3. Después, se indica que  $O_j$  pertenece al conglomerado representado por  $O_i$  cuando la observación es la mínima sobre todos los medoids  $O_e$  basado en :  $d(O_j, O_i) = \text{Min\_medoids}(O_j, O_e)$ .
4. Posteriormente, se calcula la ganancia total obtenida seleccionando el objeto j, basado en la ecuación:  $GT_j = \sum_i d_{ji}$ .
5. Si el costo total de la configuración aumentó, es necesario deshacer el intercambio e ir al paso 2. Si el resultado es contrario, continúa determinando los nuevos medoids.
6. Finalmente, se termina cuando no hay más objetos que agrupar.

### **Métricas de Validación**

Para el presente estudio, se aplicó una validación interna basada en las medidas internas y de estabilidad para evaluar el desempeño de los conglomerados. En la tabla 1 se mostrará detalladamente cada una de estas medidas. Las fórmulas matemáticas fueron tomadas en Datta y Datta (2003); Brock et al., (2008); Liu et al., (2010).

**Tabla 1***Medidas de validación para el análisis de conglomerado*

Medida	Descripción	Fórmula matemática	Valor óptimo
Conectividad	Es la medida en que los elementos se colocan en el mismo grupo que sus vecinos más cercanos.	$\sum_{i=1}^N \sum_{j=1}^L X_{i,nn_{i(j)}}$	Minimizar
Coefficiente de silueta	Mide qué tan bien está agrupada una observación y estima la distancia media entre agrupaciones.	$\frac{1}{NC} \sum_i \left\{ \frac{1}{n_i} \sum_{x \in C_i} \frac{b(x) - a(x)}{\max[b(x), a(x)]} \right\}$ $a(x) = \frac{1}{n_i - 1} \sum_{y \in C_i, y \neq x} d(x, y)$ $b(x) = \min_{j, j \neq i} \left[ \frac{1}{n_j} \sum_{y \in C_j} d(x, y) \right]$	Maximizar
Índice de Dunn	Mide la distancia entre observaciones del mismo conglomerado y observaciones de otros conglomerados.	$\min_i \left\{ \min_j \left( \frac{\min_{x \in C_i, y \in C_j} d(x, y)}{\max_k \{ \max_{x, y \in C_k} d(x, y) \}} \right) \right\}$	Maximizar
Distancia promedio (AD)	Mide la distancia promedio entre las observaciones ubicadas en el mismo grupo, en función de los datos completos y de los datos con una sola columna eliminada.	$\frac{1}{Ml} \sum_{g=1}^M \sum_{i=1}^l \frac{1}{n(C^{g,0})n(C^{g,i})} \sum_{g \in C^{g,0}, g' \in C^{g,i}} d(x_g, x_{g'})$	Minimizar

*Nota.* Componentes de las fórmulas:  $nn_{i(j)}$  = j-ésimo del vecino más cercano de la observación  $i$ ;  $X_{i,nn_{i(j)}}$  es cero si  $i$  y  $nn_{i(j)}$  están en el mismo grupo y 1 en caso contrario;  $N$  = observaciones en  $k$  grupos disjuntos;  $L$  = parámetro que determina el número de vecinos que contribuyen al medida de conectividad;  $D$  = conjunto de datos;  $n$  = número de objetos en  $D$ ;  $NC$  = número de conglomerado;  $C_i$  =  $i$ -ésimo de conglomerado;  $n_i$  = número de objetos en  $C_i$ ;  $d(x, y)$  = distancia entre  $x$  y  $y$ ;  $\bar{x}_{Cg,0}$  = perfil de expresión promedio para genes a través del conglomerado  $Cg, 0$ ;  $\bar{x}_{Cg,i}$  = perfil de expresión promedio de genes a través del conglomerado  $Cg, i$ .

## Resultados y Discusiones

### Análisis de Correlación y Estadística Descriptiva de las Variables

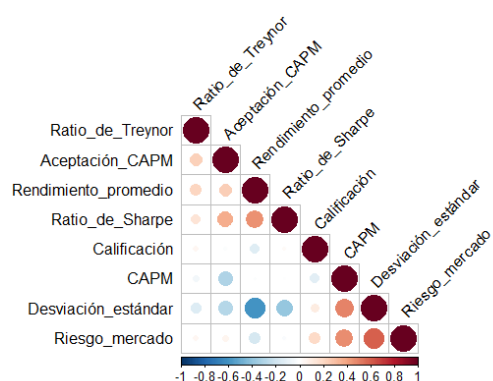
Antes de proceder con el análisis de conglomerado, es necesario revisar las características que conforman cada uno de los indicadores financieros seleccionados. Para el presente estudio, se



realizó un análisis de correlación para determinar si dentro de las variables estudiadas existe alguna que cuenta con un nivel de correlación muy elevada con otras variables. La razón de realizar este análisis es debido a que cuando existen variables altamente correlacionadas, se podría estar asignando mayor peso hacia algún aspecto en específico. A continuación, en la figura 3 se mostrará la matriz de correlación de las variables.

**Figura 3**

*Matriz de correlación de las variables*



*Nota.* Calificación se refiere a la calificación de fondos de inversión.

Como se observa en la figura 1, todas las variables analizadas del presente estudio poseen un coeficiente de correlación debajo de 0.6, por lo tanto, se puede decir que no hay variables que compartan una alta similaridad con alguna de las restantes. Sin embargo, cabe destacar que se intentó incluir otras variables que han sido empleadas frecuentemente para la evaluación de los fondos de inversión, como el coeficiente de beta y el Alfa de Jensen, pero finalmente han sido eliminadas debido a que estos indicadores poseen un nivel de correlación muy elevado con otras variables del estudio.

**Tabla 2***Estadística descriptiva de las variables*

<b>Variab</b> les	<b>Mínimo</b>	<b>Primer cuartil</b>	<b>Mediana</b>	<b>Media</b>	<b>Tercer cuartil</b>	<b>Máximo</b>
Rendimiento promedio	-20.59	1.80	2.87	2.65	5.08	14.69
Desviación estándar	0.09	0.78	1.54	4.27	2.99	31.17
Ratio de Sharpe	-7.24	0.86	1.39	1.81	2.25	11.15
Ratio de Treynor	-14.74	1.12	2.24	4.19	5.23	144.44
CAPM	0.46	1.91	2.97	4.27	5.41	16.28
Calificación de riesgo de mercado	2.00	2.00	3.00	2.80	3.00	4.00
Calificación de fondos de inversión	4.00	4.00	5.00	5.36	6.00	21.00
Aceptación de CAPM	0.00	0.00	0.00	0.48	1.00	1.00

Una vez determinadas las variables del estudio, se realizó un análisis descriptivo para visualizar el comportamiento de los fondos de inversión. En la tabla 2 se encuentra el resultado del análisis. Se observa que los fondos de inversión del mercado costarricense tienden a ser de bajo rendimiento, con una media de aproximadamente 2.65% durante los últimos 3 años y con un nivel de riesgo moderado. Por lo tanto, los indicadores de desempeño como el ratio de Sharpe y Treynor muestran que en promedio los productos ofrecen 1,81% de rendimiento por cada unidad de desviación estándar asumido y 4,19% de rendimiento por cada unidad beta asumido.

Finalmente, dentro de los datos analizados, se encuentra un fondo de muy bajo desempeño que llegó a una pérdida de aproximadamente -20,59% durante los últimos 3 años y otro que llegó a obtener un rendimiento sumamente elevado de hasta 14,69%. Sin embargo, debido a que algunos productos son de alta volatilidad, el nivel de rendimiento teóricamente requerido muestra una variación desde 0.46% hasta 16,28%. También cabe resaltar que al menos un 50% de los fondos de inversión costarricenses proporcionan un nivel de rendimiento inferior a lo requerido, lo cual muestra una falta de eficiencia del mercado en la administración de estos productos.

### **Análisis de Conglomerado**

A continuación, se mostrará detalladamente el análisis de conglomerado aplicado a los fondos de inversión costarricenses. Para empezar, se aplicó una validación interna a las diferentes

técnicas de agrupamiento y su respectiva partición óptima con el fin de identificar cuál es el método más eficiente para separar las observaciones bajo el escenario del análisis. Cabe destacar que durante este proceso, todas las variables fueron estandarizadas para no dar más importancia y peso a aquellas variables que cuentan con una escala de medida más amplia. En la tabla 3 se presentará el resultado sobre los 3 métodos de agrupamiento: jerárquico, k medias y PAM.

**Tabla 3**

*Validación interna de diferentes técnicas de agrupamiento*

<b>Número de conglomerado</b>		<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
Jerárquica	Conectividad	2.93	5.86	11.33	16.76	19.36	22.16	25.63	27.46	36.52
	Dunn	0.93	0.53	0.37	0.20	0.28	0.28	0.28	0.28	0.28
	Silueta	0.66	0.60	0.54	0.46	0.44	0.40	0.35	0.29	0.26
	Distancia promedio	3.26	3.15	2.83	2.70	2.61	2.54	2.45	2.41	2.27
k medias	Conectividad	2.93	11.98	14.55	20.83	37.80	32.31	38.47	40.57	42.90
	Dunn	0.93	0.31	0.31	0.25	0.08	0.18	0.20	0.20	0.23
	Silueta	0.66	0.55	0.54	0.46	0.29	0.28	0.33	0.33	0.32
	Distancia promedio	3.26	2.84	2.72	2.64	2.59	2.35	2.08	1.96	2.00
PAM	Conectividad	11.12	20.10	23.14	31.41	30.56	32.46	33.28	36.11	38.48
	Dunn	0.17	0.11	0.11	0.10	0.10	0.10	0.12	0.14	0.14
	Silueta	0.21	0.28	0.30	0.32	0.35	0.36	0.38	0.42	0.44
	Distancia promedio	3.05	2.74	2.48	2.35	2.13	1.93	1.77	1.69	1.50

Como se muestra en la tabla 3, tanto el método jerárquico como k medias, técnicas más utilizadas para clasificar los fondos de inversión según la revisión de literatura realizada, indican que la mejor solución es una partición en 2 grupos. Esto se debe a que logran minimizar la conectividad, maximizar los índices de Dunn y Silueta; excepto la distancia promedio que no logra su mejor eficiencia, aunque no se observa un cambio significativo en comparación con otras particiones de 3 a 10 grupos.

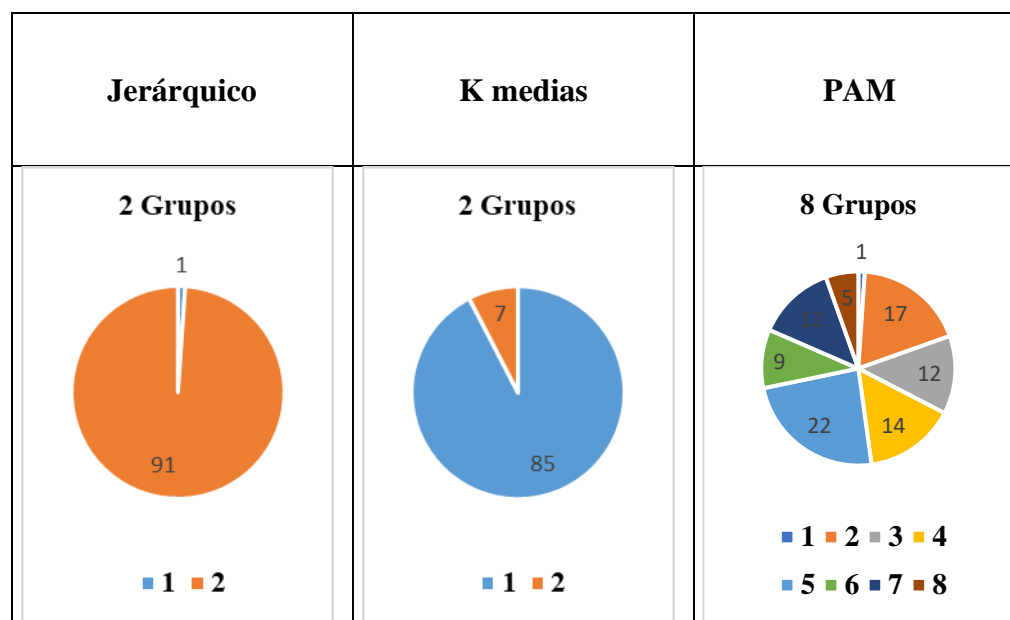
En la figura 4 se presenta una comparación de la cantidad de casos que quedaron en cada grupo con su mejor solución de las tres técnicas mencionadas. Como se aprecia en el resultado, aunque la técnica jerárquica cuenta con mejores resultados en la validación interna, no logra

clasificar de manera eficiente los fondos de inversión, debido a que concentra casi la totalidad de las observaciones en el grupo 2 y únicamente logra separar el fondo más extremo.

Por su parte el método k medias muestra un resultado similar, debido a que la solución óptima de 2 grupos clasifica casi la mayoría de los fondos de inversión en el grupo 1, con 85 observaciones de las 92 existentes. La razón por la cual estos métodos concentran los casos en pocos grupos puede estar relacionada con la presencia de valores extremos, por ser sensibles a este tipo de valores. (Lopes y Gosling, 2020). En la figura 5 se observa un análisis de valores atípicos multivariados mediante la técnica LOF. Como se observa en el resultado de la figura 6, los datos del presente estudio contienen valores atípicos, de los 92 fondos de inversión, 9 de ellos cumplen con las condiciones establecidas.

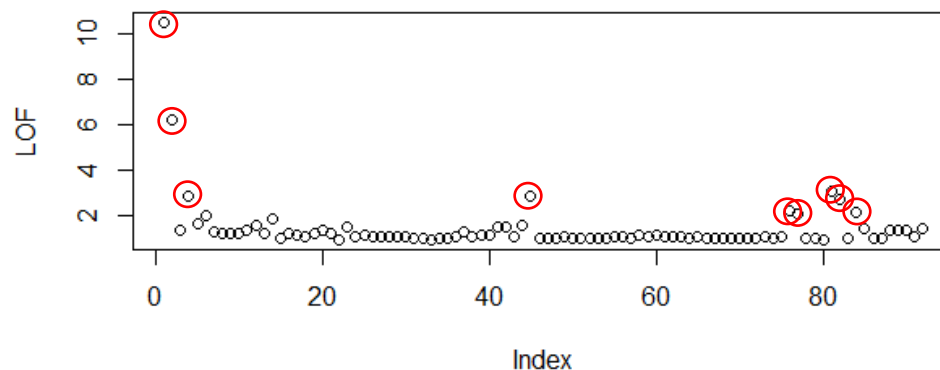
#### Figura 4

*Cantidad de casos en cada agrupamiento con la mejor solución de las tres técnicas mencionadas*



**Figura 5**

*Resultado del análisis de LOF*



*Nota.* El eje y representa el valor de LOF y el eje x es el orden en que aparecen las observaciones. Los círculos rojos representan observaciones que contienen un LOF elevado y además, distancian a los demás fondos de inversión.

Como los datos cuentan con valores atípicos, se optó por elegir el agrupamiento de inversiones con base en la técnica PAM, ya que este método es más robusto con la presencia de estos valores (Lopes y Gosling, 2020) y por ende permite mantener las 92 inversiones iniciales. Para este método, se seleccionó una partición de 8 grupos en lugar de la solución óptima de 10 grupos. La razón es debido a que durante el análisis, se ha encontrado que al segregar más grupos (9 a 10), los resultados no cambian significativamente y además, generan diversos grupos que son muy similares entre sí. Por lo tanto, para simplificar la solución, se optó por renunciar una pequeña parte de la eficiencia de separación buscando mayor conveniencia a la toma de decisión financiera.

### **Interpretación y Análisis de Cada Conglomerado**

A continuación, en este apartado se presenta el resultado final de la clasificación de los fondos de inversión en 8 grupos mediante el método PAM. En la tabla 5 se muestra un resumen para cada uno de los grupos. Cabe destacar que el valor mostrado representa el promedio de todos los fondos de inversión que conforman un mismo conglomerado para cada variable en específico.

Por su parte, las figuras 6, 7 y 8 muestran una comparación de los gráficos de caja para los conglomerados, basado en los indicadores de rendimiento, riesgo y desempeño.

**Tabla 5**

*Valor promedio de las variables para los 8 grupos de fondos de inversión*

Grupo	Rendimiento promedio	Desviación estándar	Ratio de Sharpe	Ratio de Treynor	CAPM	Riesgo de mercado	Calificación del fondo	Aceptación de CAPM	Cantidad de fondos
1	1.45	3.56	0.20	144.44 <sup>a</sup>	0.79	3.00	7.00	1.00	1
2	2.19	1.20	0.75	1.00	2.48	2.00	4.53	0.00	17
3	5.63	0.93	5.69	8.86	4.18	3.00	6.42	1.00	12
4	0.79	7.03	0.57	-1.10	2.90	3.21	6.21	0.00	14
5	3.55	3.28	1.38	3.53	2.90	3.09	4.91	1.00	22
6	2.57	1.07	1.82	1.98	2.29	2.00	4.67	1.00	9
7	7.41	3.96	2.74	5.21	9.26	2.92	5.92	0.00	12
8	-12.74	25.86	-0.53	-5.06	12.78	3.80	4.80	0.00	5

*Nota.* Para conocer los fondos de inversión de cada grupo y sus respectivas características individuales, consulte al siguiente enlace: [shorturl.at/mBJ49](http://shorturl.at/mBJ49)

<sup>a</sup> El ratio de Treynor de este fondo está alterado por un coeficiente de beta muy cercano a 0, su beta es de 0.0048.

Con base en la tabla 5 y los gráficos de caja, se puede concluir que los grupos 1 y 4 están conformados principalmente por fondos de inversión que tienden a ofrecer un nivel de rendimiento muy bajo, inferiores a 2%. Se destaca que el grupo 1 cuenta con un único fondo extremo que posee un ratio de Treynor alterado por un coeficiente de beta muy cercano a 0. No obstante, midiendo el nivel de desempeño con el ratio de Sharpe, se observa que este producto no es eficiente en la proporción del rendimiento por cada unidad de riesgo asumido. Sin embargo, el grupo 4 es aún menos eficiente, ya que posee un nivel de riesgo sumamente elevado (de acuerdo con la desviación estándar) para un producto de muy bajo rendimiento y por lo tanto, el nivel de rendimiento teóricamente requerido es significativamente más elevado que el grupo 1. Como estos dos grupos poseen muy poca eficiencia en la administración de riesgo-rendimiento de acuerdo con el ratio de Sharpe y Treynor, muestran una calificación muy riesgosa, con una categoría de riesgo mayor e

igual a 3 y calificaciones inferiores al tipo AA- (en el gráfico de caja se observa que el grupo 4 cuenta con una observación extrema, con una calificación del tipo C-).

Los grupos 2 y 6 están conformados por fondos de inversión que tienden a ofrecer un nivel rendimiento bajo, inferiores a 3% y con un nivel de riesgo sumamente bajo como lo refleja la desviación estándar y la calificación de riesgo. Estos dos grupos se caracterizan por ser productos de moderada sensibilidad a las condiciones cambiantes en el mercado y con una calificación entre el tipo AA+ y AA, de esta manera, reflejan una administración muy buena. Sin embargo, debido a que el ratio de Sharpe y Treynor del grupo 6 es el doble del grupo 2, se puede decir que este grupo es más eficiente para proporcionar rendimiento por cada unidad de riesgo asumido. Asimismo, debido a la mayor eficiencia en el grupo 6, el rendimiento ofrecido también es superior a lo requerido de acuerdo con el CAPM.

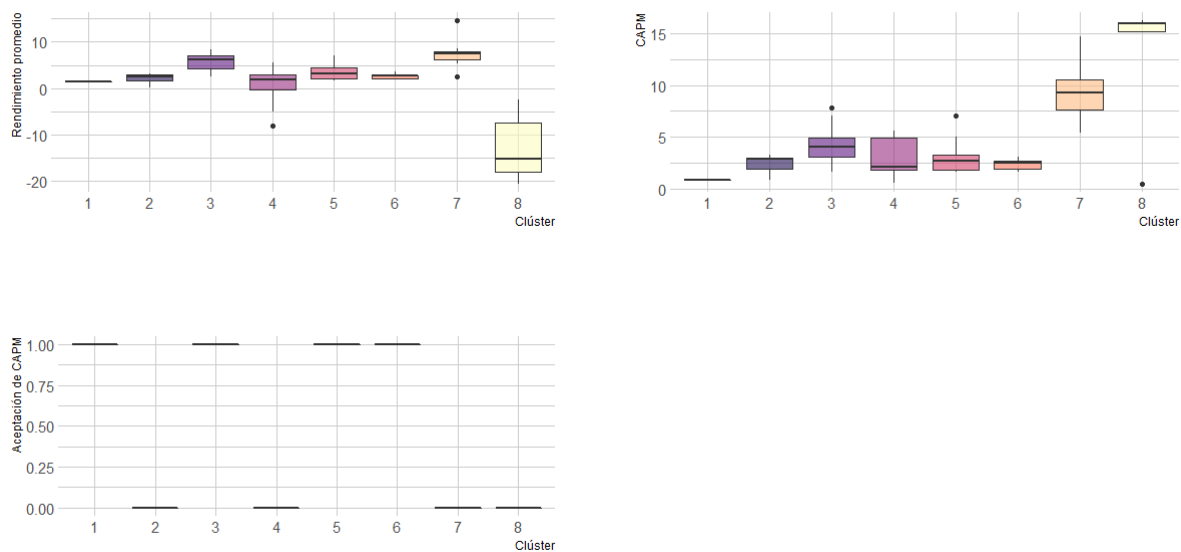
Los grupos 3 y 7 están conformados por fondos de inversión que ofrecen mayores rendimientos en el mercado. El grupo 3 se caracteriza por tener productos de alto rendimiento, pero con un nivel de riesgo muy bajo, incluso, menos riesgosos que los fondos de bajo rendimiento como se observa en la desviación estándar. Por lo tanto, son productos más eficientes en los indicadores de desempeño (sin tomar en cuenta el ratio de Treynor alterado del grupo 1). Además, como son más eficientes en la administración de riesgo-rendimiento, bajo los estimados generados por el CAPM, es aceptado, por ofrecer un nivel de rendimiento superior a lo requerido. Por otra parte, el grupo 7 se destaca por la tendencia de ofrecer rendimientos más elevados del mercado, con un nivel de riesgo moderado según la desviación estándar. Sin embargo, debido a que el nivel de riesgo agregado es superior al rendimiento agregado en comparación con el grupo 3, se considera que este grupo es menos eficiente en el desempeño y ofrece un rendimiento inferior a lo requerido de acuerdo con el CAPM.

El grupo 5 está conformado por fondos de inversión que tienden a ofrecer un moderado nivel de rendimiento y riesgo. Como se observa en los resultados, no se destaca en ningún indicador financiero, pero es el grupo que cuenta con mayor concentración de productos, está integrado por 22 fondos. Asimismo, debido a que los indicadores de rendimiento y riesgo están ubicados en un nivel intermedio, el desempeño de los fondos también se encuentra en un nivel moderado. No obstante, aunque el riesgo de este grupo es más elevado que el grupo 3, bajo las estimaciones del CAPM sigue siendo aceptado por ofrecer mayores rendimientos a lo requerido.

Por último, el grupo 8 está conformado por los fondos de inversión más riesgosos del mercado con una pérdida significativamente elevada durante los últimos 3 años. Como se observa en la desviación estándar, el valor es de 3.68 veces mayor que el grupo 4 de fondos de alto riesgo. Asimismo como lo refleja en la calificación de riesgo de mercado, son productos de alta a muy alta sensibilidad a las condiciones cambiantes del mercado. Por esta razón, es el grupo menos eficiente en el desempeño, con valores negativos en el ratio de Sharpe y Treynor, eso indica que los productos no ofrecen mayor e igual rentabilidad que un activo libre de riesgo. Sin embargo, aunque son productos de alto riesgo a las condiciones del mercado, en promedio cuentan con una calificación muy bueno del tipo AA, indicando que presentan una alta probabilidad de cumplir con sus objetivos de inversión y una buena gestión para limitar su exposición al riesgo por factores inherentes.

### Figura 6

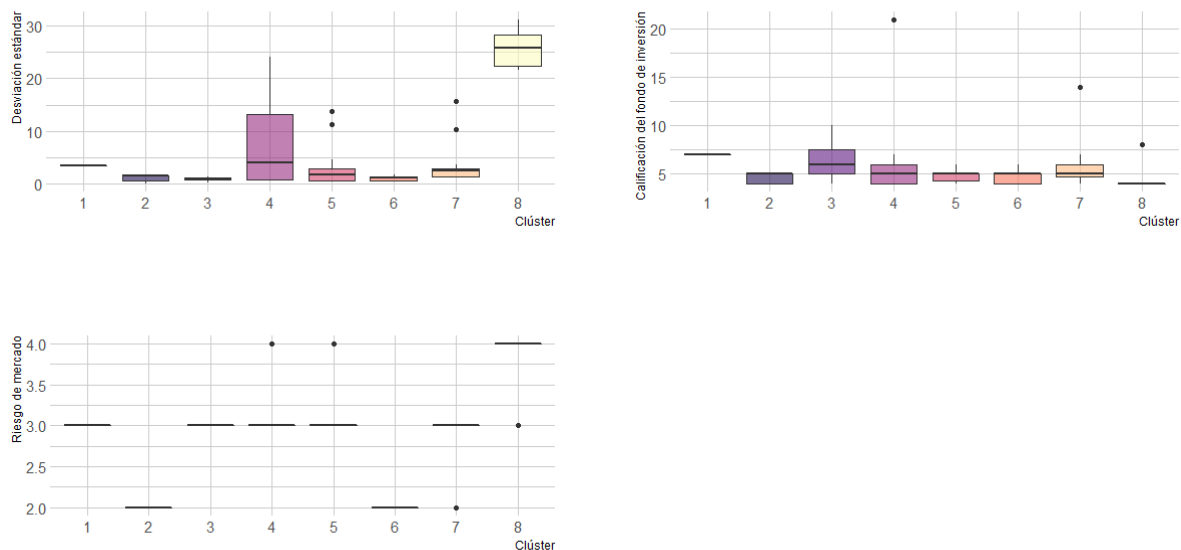
*Gráficos de caja de las variables de rendimiento según grupo del análisis de conglomerados*





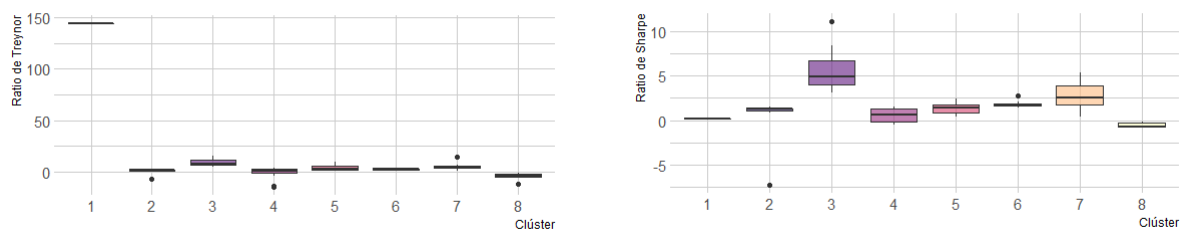
## Figura 7

Gráficos de caja de las variables de riesgo según grupo del análisis de conglomerados



## Figura 8

Gráficos de caja de las variables de desempeño según grupo del análisis de conglomerados



## Conclusiones

El propósito de este estudio fue desarrollar una propuesta de agrupamiento para los fondos de inversión costarricenses, con el fin de que los inversionistas y otros actores sociales puedan identificar cuáles productos en específicos son homogéneos entre sí desde la perspectiva de los indicadores financieros y así facilitarles una mejor toma de decisión según sus propias necesidades.

Entre las principales conclusiones se encuentra que, debido a la existencia de las observaciones extremas, se ha aplicado el método de agrupamiento PAM en lugar de otros métodos más utilizados para la clasificación de los fondos de inversión. Mediante esta técnica, se ha logrado clasificar los fondos de inversión en 8 grupos diferentes. En el presente documento se adjunta un enlace con la clasificación de cada uno de los fondos.

Con respecto al resultado de la clasificación, se ha encontrado que los grupos 1 y 4 tienden a ofrecer fondos de inversión de muy bajo nivel de rendimiento, muy poca eficiencia en el desempeño y cuentan con un nivel de calificación sumamente riesgosa. Sin embargo, el grupo 1 ofrece un mejor desempeño que el grupo 4 en la administración de riesgo.

Los grupos 2 y 6 tienden a ofrecer fondos de inversión de bajo nivel de rendimiento, bajo riesgo y con un nivel de desempeño moderado. Por lo tanto, es recomendable para aquellos inversionistas con un perfil conservador quienes no les gustan el riesgo y no les importan que las ganancias obtenidas sean bajas para minimizar las posibles pérdidas en el capital invertido. Podría resultar más recomendable seleccionar los productos del grupo 6, ya que este ofrece mejores desempeños y además, el nivel de rendimiento es superior a lo requerido de acuerdo con el CAPM.

Los grupos 3 y 7 tienden a ofrecer fondos de inversión de más altos niveles de rendimientos y mejores desempeños de acuerdo con los indicadores de ratio de Sharpe y Treynor. Sin embargo, el grupo 3 es el más destacado en los 8 grupos, porque ofrece mayores rendimientos por cada unidad de riesgo asumido y además, posee una desviación estándar más baja. De esta manera, podría ser el grupo más recomendable para los inversionistas tanto de perfil moderado como de conservador. Por otra parte, aquellos inversionistas con un perfil moderado y agresivo que desean buscar mayores rendimientos, aceptando un nivel de riesgo más elevado, pueden seleccionar los productos del grupo 7.

El grupo 5 tiende a ofrecer fondos de inversión con un nivel moderado de rendimiento, riesgo y desempeño. Es el grupo que cuenta con mayores opciones en el mercado, con 22 fondos disponibles para seleccionar. Resulta recomendable para aquellos inversionistas de perfil moderado que buscan un nivel de rendimiento superior a los grupos de bajo rendimiento como el 2 y 6, pero con la disposición de asumir un nivel de riesgo un poco más elevado.

Por último, en los resultados se ha encontrado que el grupo 8 tienen peores desempeños en todo el mercado de los fondos de inversión, con una pérdida significativamente elevada durante los últimos 3 años y asimismo son productos más riesgosos. Entonces, no son aptos para ningún tipo de inversionistas.

### **Limitaciones y Líneas Futuras**

Debido a la limitación de información disponible en el mercado costarricense, no fue posible incorporar los otros indicadores que se han sido utilizado en el mercado internacional para la evaluación de los fondos de inversión. Algunos de ellos son: índice de omega, ganancia de cola, ratio de Sortino, entre otros.

Una segunda limitación fue la estimación del coeficiente beta para los fondos de inversión, específicamente sobre el factor del rendimiento del mercado, debido a que actualmente, según la SUGEVAL y la sociedad calificadora de riesgo, el rendimiento del mercado de estos productos ha sido estimado bajo cada categoría de objetivo de inversión, es decir, se cuenta con un coeficiente de beta individual para la industria de mercado de dinero, accionario, inmobiliario, entre otros. Sin embargo, los resultados obtenidos bajo esta metodología de estimación preestablecida no se estarían tomando en cuenta el nivel de riesgo de todo el mercado de fondos de inversión completo.

Por esta razón, como una primera línea futura, se recomienda valorar la posibilidad de realizar una estimación del coeficiente beta incorporando toda la industria de fondos de inversión como un solo mercado, integrando productos de diferentes categorías en una misma regresión lineal. De esta manera, se puede evaluar más eficientemente el comportamiento de cada fondo de inversión ante los cambios en el mercado completo.

Una segunda línea futura está relacionada con la actualización del algoritmo, debido a que el desempeño de los fondos de inversión puede sufrir ciertos cambios a través del tiempo y la posibilidad de surgir nuevos productos en el mercado, es recomendable de realizar actualizaciones a la clasificación existente. Esta actualización puede realizarse bajo dos maneras: la primera se trata de replicar el mismo método con datos más recientes y la segunda es aplicar el algoritmo de predicción utilizando los resultados del presente estudio para clasificar la entrada de nuevos fondos de inversión.

Finalmente, otra línea futura puede estar enfocada en la aplicación de esta técnica de agrupamiento pero con otros productos financieros del mercado costarricense, como los fondos de pensiones que igualmente cuentan con gran cantidad de opciones disponibles (Gamboa et al., 2019) y mediante el análisis de conglomerado podría facilitarles a los inversionistas una mejor toma de decisión a la hora de seleccionar estos productos de ahorro.

### Referencias

- Acharya, D., & Sidana, G. (2007). Classifying mutual funds in India: Some results from clustering. *Indian Journal of Economics and Business*, 6(1), 71-79. <https://ssrn.com/abstract=999856>
- Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks* (Vol. 19). Academic press.
- Babaei, K., Chen, Z., & Maul, T. (2019). Detecting point outliers using prune-based outlier factor (plof). *arXiv preprint arXiv:1911.01654*.
- Björkby, S., & Grägg, S. (2019). A Cluster Analysis of Stocks to Define an Investment Strategy.
- Bolaños-Ulloa, A., Cascante-Porras, A., Rey-Vargas, J., y Vargas-Solano, R. (2017). *Valoración de la rentabilidad y riesgo de Fondos de Inversión de mercado monetario, en el caso de Costa Rica para el periodo 2012-2016*. [Tesis de licenciatura]. Repositorio TEC. <https://hdl.handle.net/2238/10057>
- Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). cIValid: An R package for cluster validation. *Journal of Statistical Software*, 25(1), 1-22. <https://doi.org/ggsmz5>
- Das, N. (2003, August). Hedge fund classification using k-means clustering method. *In 9th International Conference on Computing in Economics and Finance* (pp. 11-13).
- Datta, S., & Datta, S. (2003). Comparisons and validation of statistical clustering techniques for microarray gene expression data. *Bioinformatics*, 19(4), 459-466. <https://doi.org/10.1093/bioinformatics/btg025>

- Dayaratne, D., Dharmaratne, D., & Harris, S. (2010). Measuring the Risk and Performance in Plantation Sector Using CAPM Based Jensen's Alpha. *Sabaragamuwa University Journal*, 6, 68. <https://doi.org/d687pp>
- Elguera Vega, R. M. (2018). *Segmentación de clientes de un casino utilizando el algoritmo partición alrededor de medoides (PAM) con datos mixtos*. [Tesis de bachillerato]. Repositorio La Molina. <http://repositorio.lamolina.edu.pe/handle/UNALM/3312>
- Gamboa-Ulate, S., Gómez-Solano, K., Gutiérrez-Espinoza, L., Quirós-Jiménez, J., Ugalde-Alfaro, M., & Vega-Madriz, A. (2019). *Análisis financiero comparativo entre fondos de inversión y fondos de pensión complementaria voluntaria en Costa Rica*. [Tesis de licenciatura]. Repositorio TEC. <https://hdl.handle.net/2238/10938>
- Gitman, L. y Joehnk, M. (2009). *Fundamentos de Inversiones* (10a ed.). Pearson Educación.
- Gitman, L. y Zutter, C. (2016). *Principios de administración financiera* (14 ed.). Pearson Educación.
- González, R. M., y Álvarez, J. A. R. (2011). Caracterización de las PYMES en Costa Rica mediante el empleo del Análisis de Conglomerados o Cluster (Categorization of the Costa Rican SME using cluster analysis). *Tec Empresarial*, 5(1), 8-19. <https://hdl.handle.net/2238/5568>
- Isakov, V. (2019). *Performance appraisal of exchange-traded funds using clustering and data envelopment analysis (XETRA, Germany)*. [Master's Thesis]. LUTPub. <http://urn.fi/URN:NBN:fi-fe2019061720812>
- Kumar, N. L., & Rama Devi, V. (2011). Cluster analysis of mutual funds.
- Lopes, H. E. G., & Gosling, M. D. S. (2020). Cluster analysis in practice: Dealing with outliers in managerial research. *Revista de Administração Contemporânea*, 25(1). <https://doi.org/gzt8>
- Liu, Y., Li, Z., Xiong, H., Gao, X., & Wu, J. (2010). Understanding of Internal Clustering Validation Measures. *2010 IEEE International Conference on Data Mining*, 911-916. <https://doi.org/c39wqg>

- Madhulatha, T. S. (2012). An overview on clustering methods. *arXiv preprint arXiv:1205.1117*.  
<https://doi.org/gjw2cz>
- Marathe, A., & Shawky, H. A. (1999). Categorizing mutual funds using clusters. *Advances in Quantitative analysis of Finance and Accounting*, 7(1), 199-204.
- Medellu, J. V., & Nugraha, E. S. (2021, July). k-means and k-medoid algorithm application in clustering stock data in Indonesia. *In Proceeding of The Symposium on Data Science (SDS)* (Vol. 1).
- Menardi, G., & Lisi, F. (2015). Double clustering for rating mutual funds. *Electronic Journal of Applied Statistical Analysis*, 8(1), 44-56. <https://doi.org/gk39c5>
- Momeni, M., Mohseni, M., & Soofi, M. (2015). Clustering stock market companies via K-means algorithm. *Kuwait Chapter of the Arabian Journal of Business and Management Review*, 4(5), 1-10. <https://doi.org/gzt9>
- Moreno, D., Marco, P., & Olmeda, I. (2006). Self-organizing maps could improve the classification of Spanish mutual funds. *European Journal of Operational Research*, 174(2), 1039-1054. <https://doi.org/10.1016/j.ejor.2004.12.018>
- Nieto, B., y Rubio, G. (2002). El modelo de valoración con cartera de mercado: una nueva especificación del coeficiente beta. *Revista Española de Financiación y Contabilidad*, 31(113), 697-723. <https://doi.org/10.1080/02102412.2002.10779459>
- Peachavanish, R. (2016). Stock selection and trading based on cluster analysis of trend and momentum indicators. *In Proceedings of the International MultiConference of Engineers and Computer Scientists* (Vol. 1, pp. 317-321). IMECS.
- Romesburg, C. (2004). *Cluster analysis for researchers*. Lulu.com.
- Sakakibara, T., Matsui, T., Mutoh, A., & Inuzuka, N. (2015). Clustering mutual funds based on investment similarity. *Procedia Computer Science*, 60, 881-890.  
<https://doi.org/10.1016/j.procs.2015.08.251>
- SCRiesgo (s.f). *Escala de calificaciones*. <https://www.scriesgo.com/ScaleSummary>

- Sharpe, W. F. (1994). The sharpe ratio. *Journal of portfolio management*, 21(1), 49-58. <https://doi.org/10.3905/jpm.1994.409501>
- Shanmugham, D. (2011). Return-based performance analysis of selected equity mutual funds schemes in India—an empirical study. *CHIEF PATRON*. 1(1), 113-119.
- Suneetha, Y., & Latha, G. A. (2020). A study on Performance Evaluation of Selected Mutual Funds with special reference to Balanced Funds.
- Superintendencia General de Valores. (2021). *Boletín quincenal de fondos de inversión. Activos totales administrados y número de cuentas abiertas al 30 de julio de 2021*. <https://aplicaciones.sugeval.fi.cr/InformesEstadisticas/BoletinFondosInversion>
- Tamara, D., & Revina, S. (2015). Indonesian Mutual Funds Classification Using Clustering Method. *Advanced Science Letters*, 21(4), 826-829. <https://doi.org/10.1166/asl.2015.5892>
- Tekin, B., & Gümüş, F. B. (2017). The classification of stocks with basic financial indicators: An application of cluster analysis on the BIST 100 index. *International Journal of Academic Research in Business and Social Sciences*, 7(5). <https://doi.org/gzvd>
- Verma, M., & Hirpara, M. J. R. (2016). Performance evaluation of portfolio using the Sharpe, Jensen, and Treynor methods. *Scholars Journal of Economics, Business and Management*, 3(7), 382-390. <https://doi.org/gzvf>
- Yatskiv, I., & Gusarova, L. (2005). The methods of cluster analysis results validation. *In Proceedings of International Conference RelStat'04* (Vol. 6, pp. 75-80).