
Análisis de estrategias de inversión en hedge fund

Thiago Ignacio Benitez
Diciembre 2025

Objetivo del Análisis:

Analizar las estrategias utilizadas por los hedge funds (1997-2018), determinando patrones, comparando y agrupando según rendimientos, a la vez que determinando los componentes principales que explican su fluctuación.

Base de datos utilizada:

- **Fuente:** Datos de Kaggle basados en los índices agregados de EDHEC-Risk. EDHEC publica índices compuestos de hedge funds para 13 estrategias típicas (p. ej. equity long/short, macro, event-driven, etc.)
- **Cobertura:** Retornos mensuales desde 1997 hasta 2021 para cada estrategia.
- **Variables clave:** Columna de fecha (fin de mes) y una columna de retorno mensual para cada estrategia de hedge funds. El dataset está en formato dataframe con valores netos de comisiones.

Temario:

1. Definiciones importantes antes de arrancar.
2. Análisis exploratorio de los datos, identificación de outliers y visualización.
3. Test de hipótesis en análisis de varianzas en los rendimiento según riesgo de las estrategias.
4. Clustering de las estrategias.
5. Análisis de los componentes principales e interpretación.
6. Conclusiones.

Introducción:

Los **hedge funds** son fondos de inversión privados y no regulados, que utilizan instrumentos o estrategias sofisticados, como valores derivados, posiciones cortas o apalancamiento, para generar rendimientos; para esto abarcan una amplia gama de estrategias con diferentes perfiles de riesgo y rentabilidad.

Este esquema establecido de autonomía e independencia de regulaciones, presenta desafíos para la presentación y alcance al la mayoría de los agentes involucrados en el mercado financiero. El presente trabajo aborda una breve definición, caracterización del comportamiento y del riesgo, para un conjunto de 13 estrategias financieras mayormente utilizadas por los hedge funds; mediante técnicas de reducción dimensional, análisis de variaciones y componentes principales, así como visualizaciones de clusters, sobre el periodo 1997-2018. Se barajaron diversas hipótesis, con el fin de explorar distintas incógnitas que pueden surgir ante el análisis de los rendimientos.

1. Estrategias de los hedge funds

Estrategias de Inversión y Nivel de Riesgo				
Estrategia	Definicion	Expectativa del Mercado	Riesgo	
Equity Market Neutral	Neutraliza exposición al mercado tomando posiciones opuestas.	= Neutral		
Fixed Income Arbitrage	Explota ineficiencias en renta fija (bonos, swaps, etc).	= Neutral		
Relative Value	Aprovecha diferencias entre instrumentos similares.	= Neutral		
Distressed Securities	Invierte en deuda de empresas en dificultades.	Bajista		
Short Selling	Apuesta a que un activo va a bajar de precio.	Bajista		
CTA Global	Sigue tendencias de mercado usando derivados globales.	Direccional		
Emerging Markets	Invierte en países emergentes, alto riesgo/retorno.	Direccional		
Global Macro	Apuesta macroeconómica a tasas, monedas, commodities.	Direccional		
Fund of Funds	Diversifica a través de múltiples hedge funds.	Direccional		
Convertible Arbitrage	Compra bonos convertibles y vende acciones subyacentes.	= Parcial		
Event Driven	Apuesta a eventos como fusiones, bancarrotas o spin-offs.	= Parcial		
Long/Short Equity	Posiciones largas y cortas en acciones.	= Parcial		
Merger Arbitrage	Captura ganancias en procesos de fusión/adquisición.	= Parcial		
= Parcial combina elementos de estrategias Bajistas y Alcistas , diferenciándose de las estrategias puramente Direccionales o Neutrales.				

2. Análisis exploratorio de los datos

- La base de datos dispone de 263 meses, desde Ene-1997 hasta Nov-2018, obteniendo observaciones para cada una de las 13 estrategias, es decir cada columna representa una estrategia, teniendo así 3419 observaciones en total.
- Las observaciones ya están representadas en términos porcentuales, ya sea de ganancia o pérdida según indique el signo.

2. Breve resumen estadístico

Resumen Estadístico por Estrategia				
Estrategia	Media	Maximo	Minimo	
Distressed Securities	0.69	5.04	-8.36	
Event Driven	0.63	4.42	-8.86	
Long/Short Equity	0.63	7.45	-6.75	
Emerging Markets	0.63	12.30	-19.22	
Relative Value	0.58	3.92	-6.92	
Convertible Arbitrage	0.55	6.11	-12.37	
Global Macro	0.54	7.38	-3.13	
Merger Arbitrage	0.54	2.72	-5.44	
Equity Market Neutral	0.45	2.53	-5.87	
Fixed Income Arbitrage	0.44	3.65	-8.67	
Funds Of Funds	0.43	6.66	-6.18	
CTA Global	0.41	6.91	-5.68	
Short Selling	-0.17	24.63	-13.40	

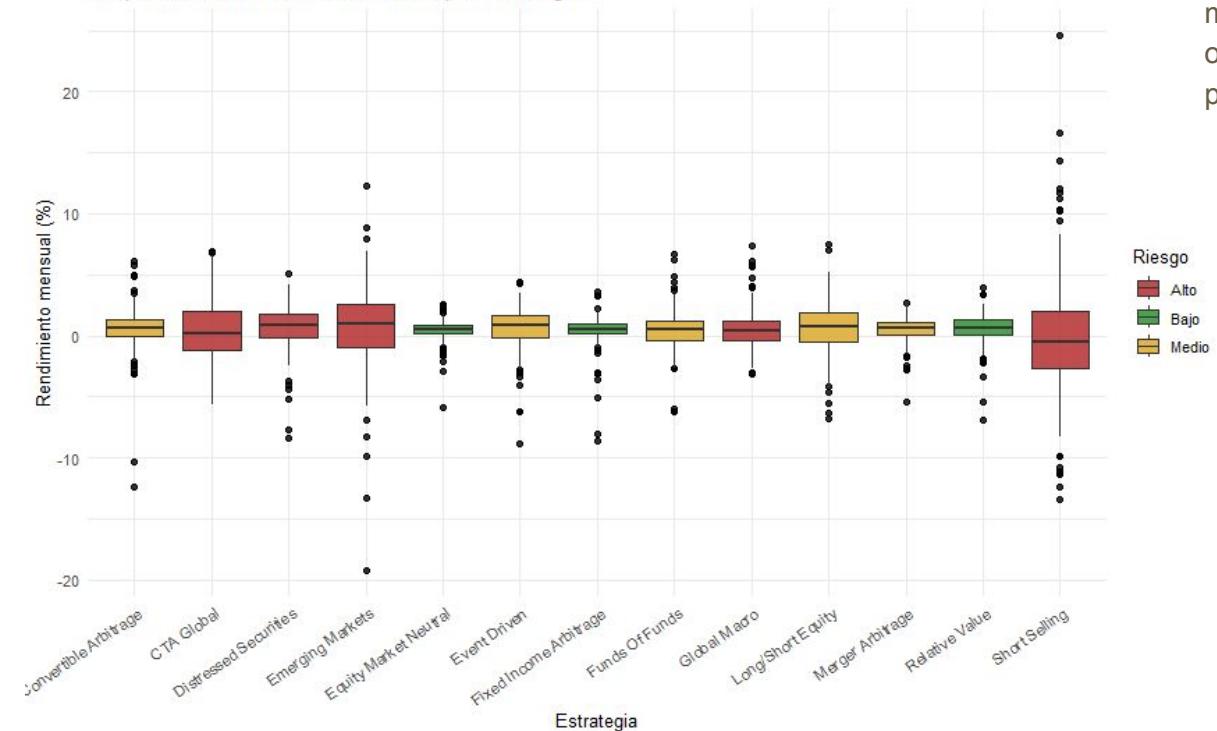
Podemos observar la media de cada estrategia ordenada de mayor a menor, donde también resulta relevante dimensionar los picos máximos y mínimos para identificar posibles outliers.

- **Short Selling** obtuvo en 1 mes una ganancia extraordinaria del 24.63%
- La mayor pérdida fue para **Emerging Markets** de un mes hacia otro en -19.22%

Ambas observaciones corresponden a la misma fecha, Agosto de 1998, posiblemente esté fuertemente relacionado a la crisis rusa del rublo, cuando el gobierno ruso entró en default, provocando fuga de capitales desde los mercados emergentes y venta masiva de activos riesgosos.

2. Visualización box-plot

Boxplot de Rendimientos Mensuales por Estrategia



Box-plot permite comparar rápidamente las medias, variaciones y distribución de las observaciones visualmente; a nivel general podemos destacar que:

- Las estrategias clasificadas como de **riesgo bajo** suelen presentar cajas más compactas, lo que refleja una menor variabilidad en sus rendimientos. Esto es consistente con su definición de riesgo más acotado.
- Por el contrario, las estrategias de **riesgo alto** tienden a mostrar mayor dispersión y presencia de valores extremos, lo que se traduce en cajas más extendidas y presencia visible de posibles outliers.

3. Test hipótesis (ANOVA)

En esta sección tendremos como hipótesis central que **las estrategias con distinto nivel de riesgo necesariamente tienen un rendimiento medio distinto.**

Para esto, el plan a ejecutar sería el siguiente:

1. Agrupar las estrategias según riesgo (bajo/medio/alto).
2. Comparación de las medias por grupo de riesgo, mediante ANOVA.
3. Visualizaciones.
4. Conclusión en base a los resultados.

Es importante resaltar que se decidió tratar los outliers encontrados por los métodos de Z-score, con las técnicas de Winsorización, reemplazando valores extremos por percentiles 2 y 98, con el objetivo de reducir la varianza y el impacto de los mismos sin disminuir la cantidad de observaciones.

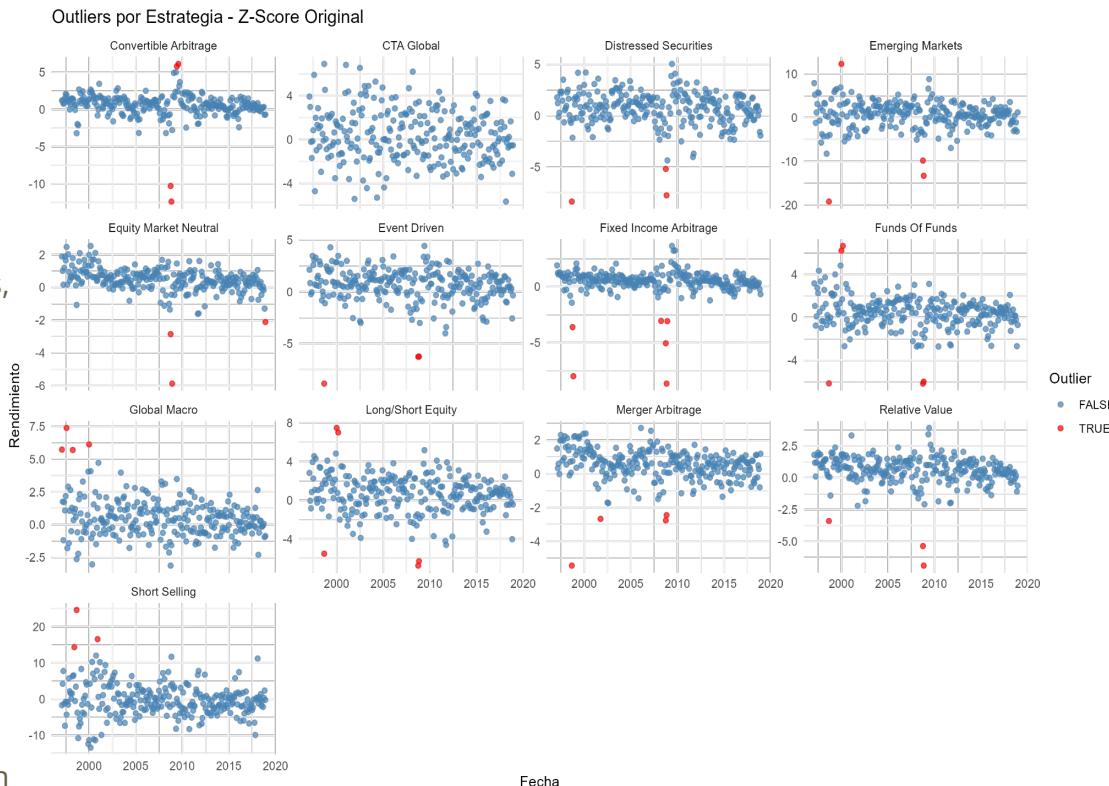
3. Supuestos y correcciones (normalidad de los residuos)

Vale la pena aclarar algunos **supuestos** asumidos acerca de las observaciones, con el fin de aportar robustez a las inferencias:

1. Debido a la gran cantidad de observaciones por grupo de riesgo (+200), podemos argumentar que debido al **Teorema Central del Límite**, existe tendencia a la normalidad en los promedios.
2. Aun así, se optó por **winsorizar** la base de datos, reemplazando los **outliers extremos** (47) por los percentiles 2 y 98.
3. Además se contrarrestó la base con el **test Shapiro Wilk**, que reflejó **P values < 0.05** para los 3 grupos de riesgo, rechazando H0 y argumentando la no normalidad de los residuos.

Esto es esperable en series financieras: los retornos rara vez son realmente normales, incluso después de transformaciones.

De esta manera, sería correcto realizar **inferencias estadísticas válidas** por más que los residuos no sigan una distribución normal.



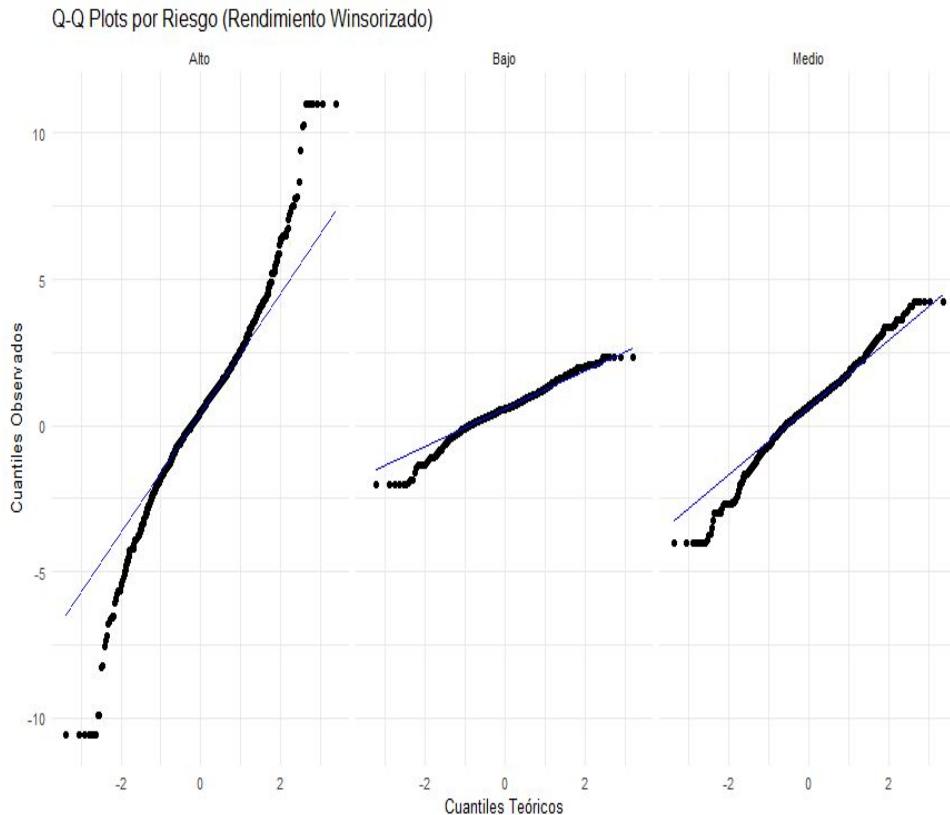
3. Supuestos y correcciones (Homogeneidad de varianza)

El **test de Levene** se usó para contrarrestar si la base cumplía con la homogeneidad de la varianza.

El mismo reflejo un **P value muy bajo < 0.05**, por lo cual rechazamos H₀ y se argumenta que la varianza de los 3 grupos no es homogénea.

Esto tiene lógica y es consistente si intuimos que:

- El riesgo Alto tiene rendimientos más volátiles
- El riesgo Bajo tiene rendimientos más estables
- El riesgo Medio está en el medio.



3. Test Welch ANOVA & Kruskal Wallis

Debido a la violación del supuesto de homogeneidad de varianzas para los diversos grupos de riesgo y la “debilidad” en la normalidad de los residuos, se decidió con el fin de robustecer el análisis, realizar los siguientes tests:

1. **Welch ANOVA** (no asume homogeneidad de las varianzas)
2. **Kruskal-Wallis** (no asume normalidad de la distribución de los residuos, ni homogeneidad de las varianzas)

Los resultados obtenidos fueron que:

- Para **Welch ANOVA**: H_0 No rechazada ($p = 0.1455$), no hay diferencia significativa de medias de rendimiento entre grupos de riesgo.
- **Kruskal-Wallis**: H_0 no rechazada al 5%, pero sí al 10% ($p = 0.0686$), hay indicios débiles de que los rendimientos podrían diferir entre grupos.

El **resultado** indica que **las medias no son estadísticamente diferentes** aunque las formas de las distribuciones sí podrían serlo (más colas, asimetrías...).

3. Análisis post hoc (GH - Dunn)

Finalmente, se realizaron tests post hoc (Games–Howell y Dunn con corrección por comparaciones múltiples) a pesar de no tener significancia clara. Como resultado, **ninguna comparación par a par resultó estadísticamente significativa tras ajuste**.

Recordando nuestra hipótesis disparadora inicial y preguntandonos, **¿Las estrategias con distinto nivel de riesgo tienen distinto rendimiento medio?**

Podemos concluir que si bien existen diferencias visuales en la forma de las distribuciones entre grupos (especialmente en el grupo de riesgo "Alto"), **no se evidencian diferencias significativas en los rendimientos promedio o mediano entre los niveles de riesgo bajo los criterios estadísticos tradicionales.**

4. Clustering

En esta sección observaremos el agrupamiento de las estrategias de los hedge funds según su rendimiento, intentando determinar una lógica de “mercado”, es decir si es posible agruparlas por según su correlación, descubriendo cuantos clusters obtendremos y si estos tienen una lógica “ex-ante”.

Para esto, se transformó la base de datos con los siguientes pasos

1. Estandarización de rendimiento
2. Cálculos de Clusters
3. Visualizaciones
4. Conclusión en base a los resultados

Cabe resaltar que a pesar de que los rendimientos están en la misma escala, la base fue estandarizada con el fin de obviar la magnitud de las distancias como fuente de distorsión y priorizar el enfoque del análisis sobre la forma y dirección de los rendimientos.

5. Estudio de sensibilidad

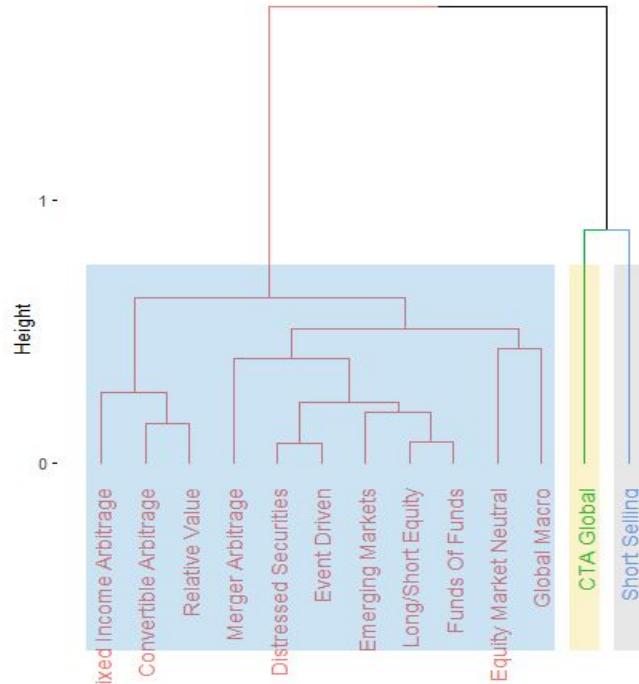
Para los análisis posteriores se decidió **trabajar con la base de datos original, sin tratamiento de outliers (sin winsorización)**.

Se realizó previamente un análisis de sensibilidad utilizando una base con datos winsorizados, a fin de mitigar el impacto de valores extremos.

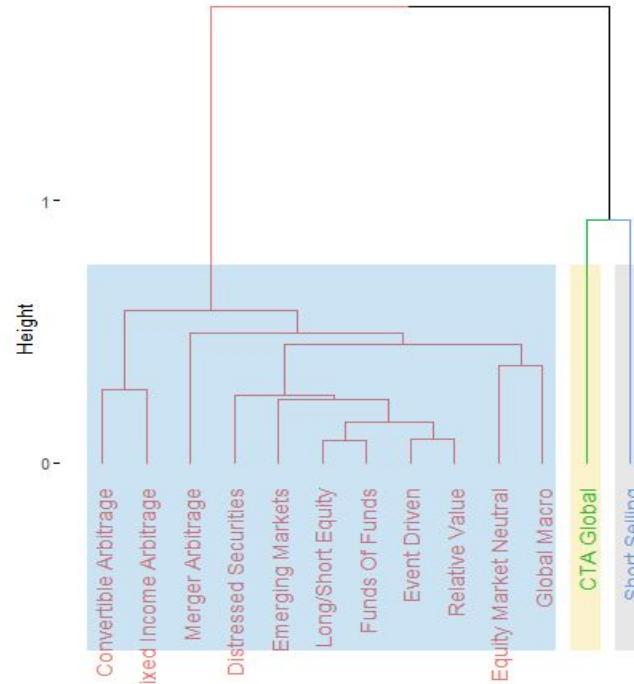
Si bien podemos ver alguna modificación leve en la segmentación de **clústeres** vía **dendrograma**, la estructura general y los patrones de correlación se mantuvieron.

Se optó por preservar la variabilidad original de los datos para captar mejor las diferencias entre estrategias.

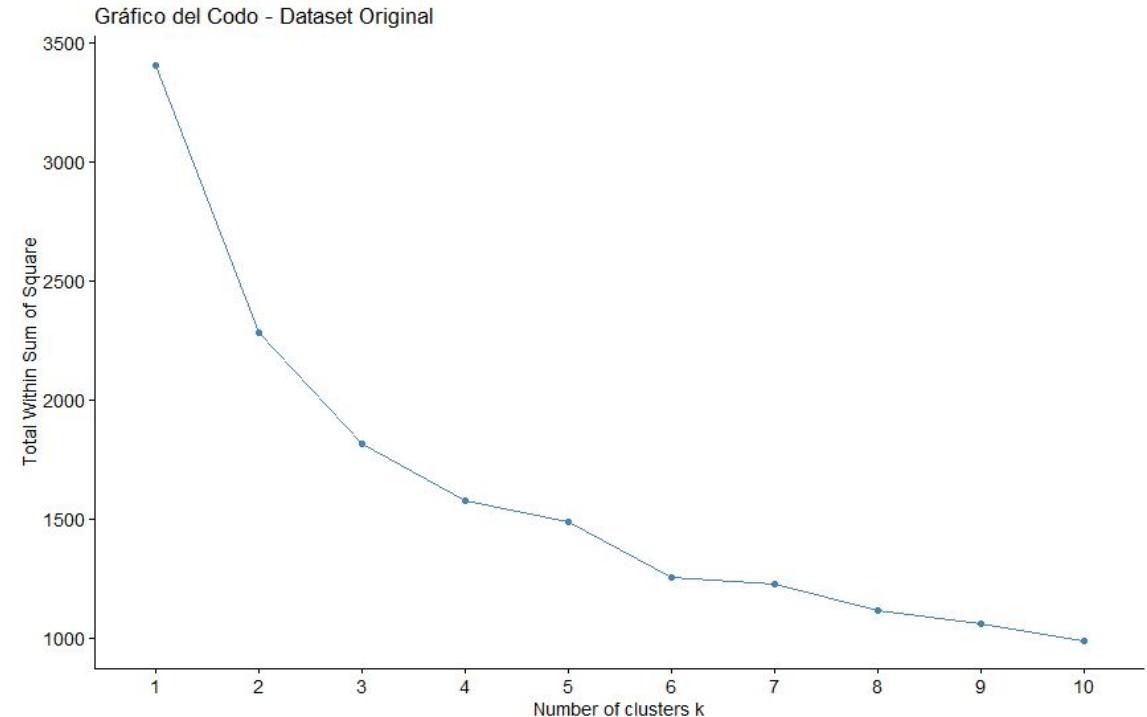
Dendrograma - Dataset Original



Dendrograma Outliers tratados



5. Visualización gráfico tipo Codo



El gráfico de codo muestra la relación entre el número de clusters (k) y la suma de cuadrados intra-cluster (WSS).

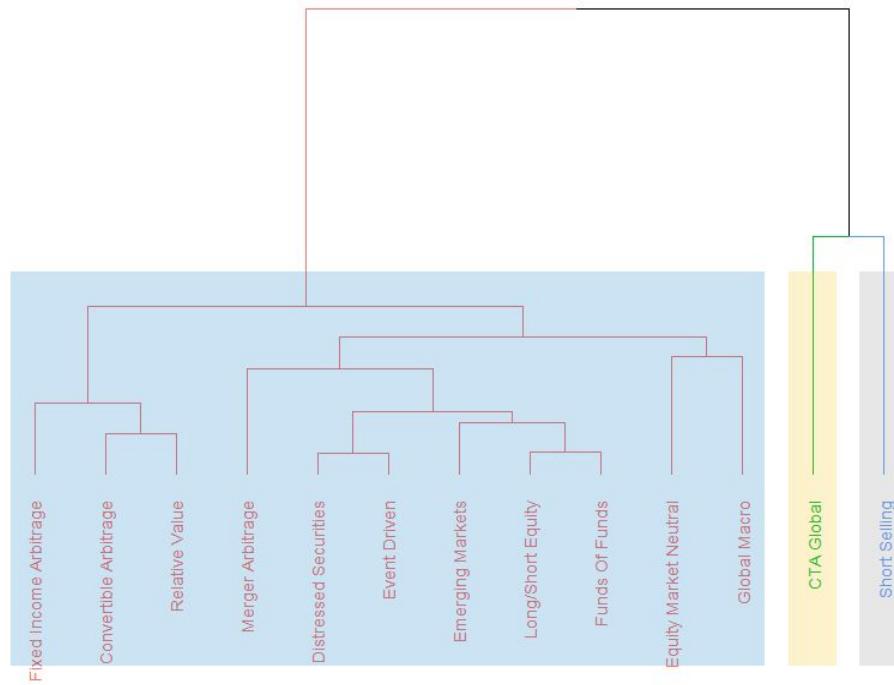
Un WSS **más bajo** implica que las observaciones dentro de cada cluster están **más juntas entre sí** (más compactas), por ende con más grupos, es más fácil que cada uno sea "**homogéneo**".

- Se observa un punto de inflexión claro en **k = 3 o 4**, quizás **5** puede ser un número de clusters interesante a tener en cuenta.
- A partir de ese valor, las mejoras en compactación son ,
- Justifica la elección de **3 o 4 grupos** como estructura natural de los datos, alineado con los dendrogramas obtenidos.

4. Dendrograma

El dendrograma permite visualizar cómo se agrupan las estrategias de hedge funds en función de la **similitud en la evolución de sus rendimientos mensuales**, para esto se utilizó como métrica la **distancia basada en 1 - correlación de Pearson** y el método de agrupamiento de **enlace completo**. El resultado evidencia una estructura jerárquica clara con **tres agrupamientos principales**:

Dendrograma - Dataset Original



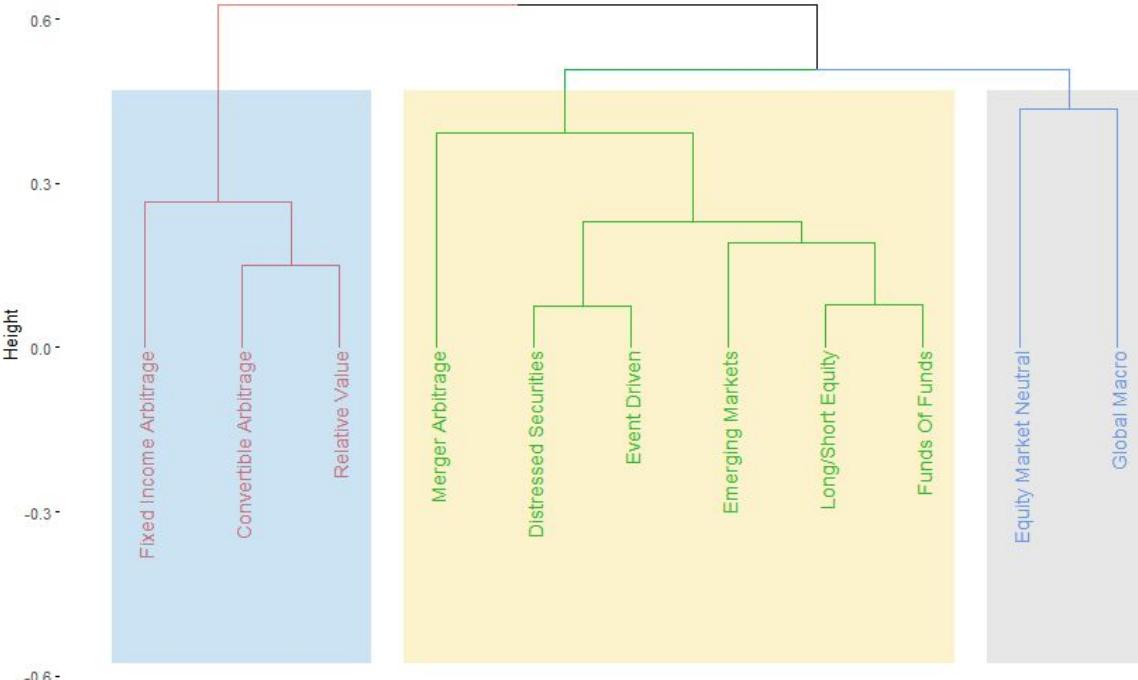
1. **Principal (azul)**: incluye la mayoría de las estrategias tradicionales, que tienden a moverse en línea positiva con el mercado.
2. **Cluster amarillo**: *CTA Global*, cuya dinámica de rendimientos difiere del grupo general (estrategia sistemática).
3. **Cluster gris**: *Short Selling*, estrategia con perfil inverso al mercado.

La segmentación captura **diferencias estructurales en la lógica de inversión**: estrategias como *Short Selling* o *CTA Global* tienen comportamientos contracíclicos o sistemáticos, a diferencia del resto.

La separación observada justifica su exclusión temporal en algunos análisis posteriores (como PCA restringido).

5. Dendrograma relevante

Dendrograma base original (sin CTA Global y Short Selling)



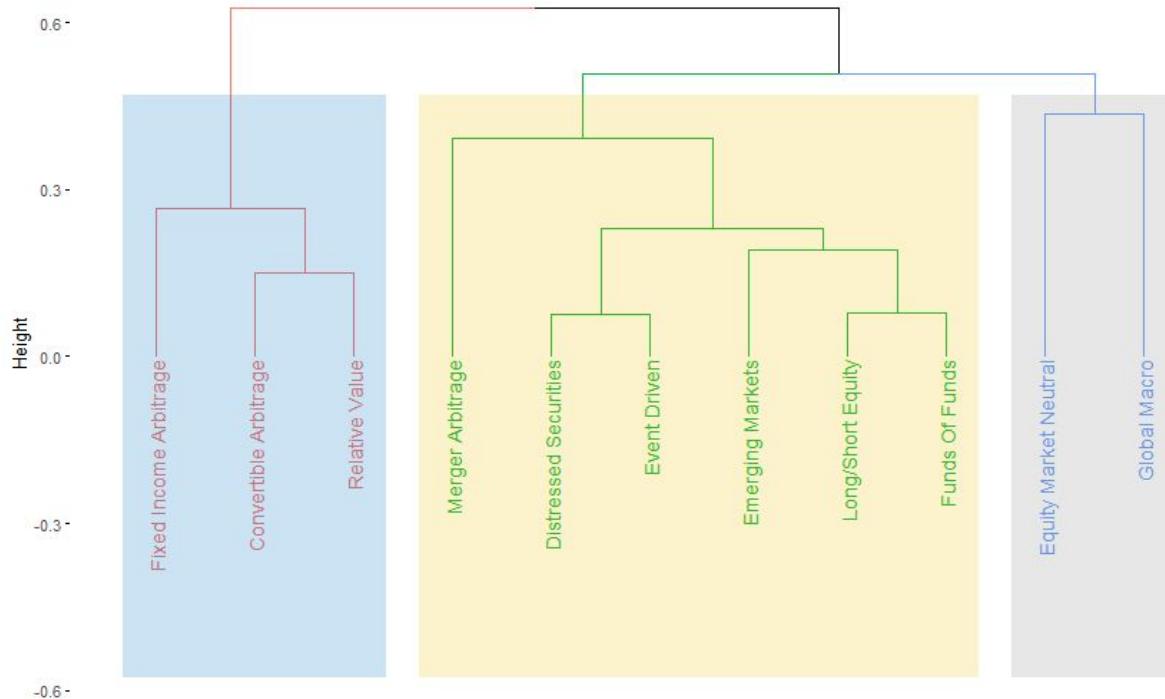
Podemos observar 3 clusters claramente, para entender estas agrupaciones, es necesario examinar los fundamentos financieros de cada estrategia, algunos de los cuales pueden ser:

1. Su exposición al mercado (beta)
2. Factor al riesgo
3. Naturaleza direccional vs. neutral
4. Sensibilidad a tasas de interés
5. Liquidez de sus inversiones
6. Otros factores.

Cabe resaltar que **Equity Market Neutral** (EMN) y **Global Macro** (GM) se distinguen por su rol de diversificadores, **EMN** minimiza el beta de mercado y **GM** reparte su riesgo en múltiples factores globales (ninguno dominante todo el tiempo), incluso pudiendo beneficiarse de dislocaciones que perjudican a otras estrategias. Por esta razón es visible su separación con el resto del cluster.

5. Dendrograma - Cluster alto beta mercado

Dendrograma base original (sin CTA Global y Short Selling)

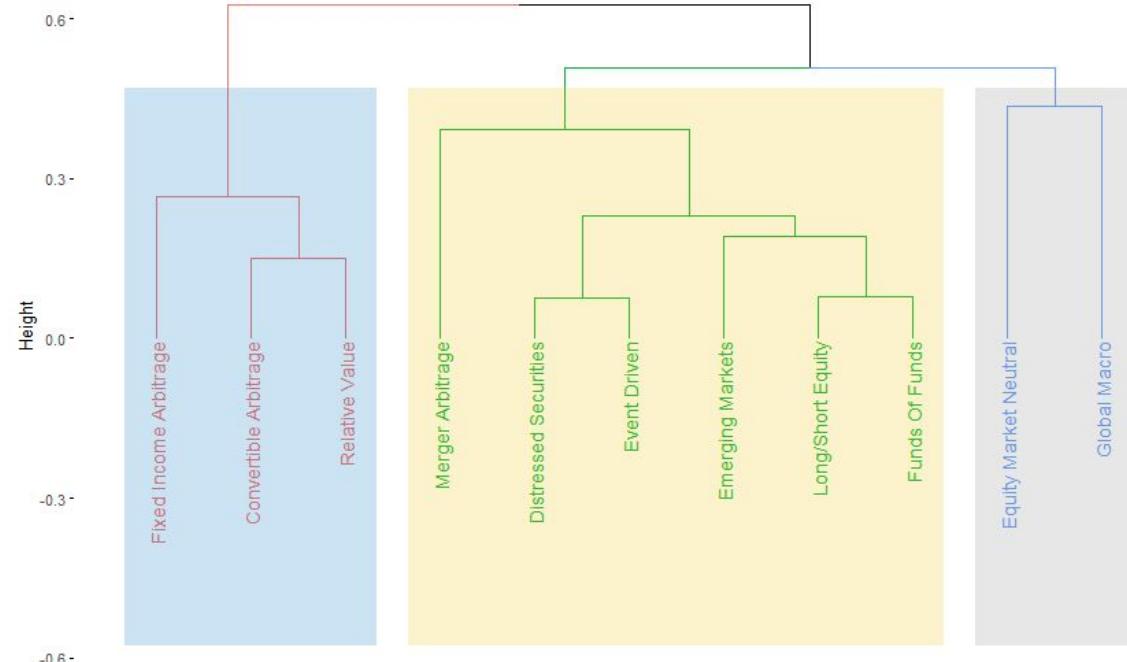


En primer lugar tenemos el cluster "verde" que abarca 46% de las estrategias.

Estas estrategias tienden a moverse en conjunto, compartiendo una **elevada exposición al mercado accionario global**. En periodos alcistas comparten la tendencia positiva de las bolsas, mientras que en crisis suelen caer simultáneamente. el factor común es la **beta de mercado**.

5. Dendrograma - Cluster de Relative Value

Dendrograma base original (sin CTA Global y Short Selling)

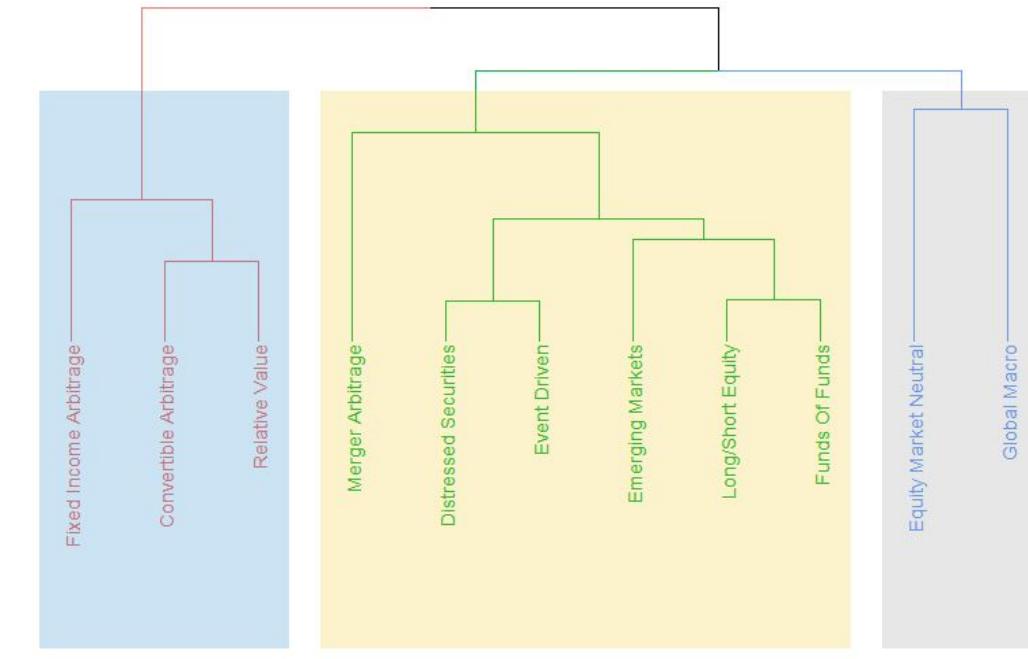


El cluster “rojo”, relacionado al arbitraje suele tener por característica común, una aproximación casi neutral al mercado, buscando ganancias en desequilibrios de precios entre valores relacionados, usualmente empleando alta apalancamiento para amplificar pequeños diferenciales.

Estas estrategias comparten sensibilidades a factores de renta fija: **tasas de interés, diferenciales de crédito y volatilidad implícita**, manteniendo el **beta** de mercado bajo.

5. Clusters y conclusiones

Dendrograma base original (sin CTA Global y Short Selling)



Entonces es posible agrupar las estrategias en 3 clusters distintos:

1. **Clúster** de valor relativo (arbitraje de bonos/ convertibles) de **rojo**.
2. **Clúster** de eventos y equity (situaciones especiales y exposición accionaria) de color **verde**.
3. estrategias macro/market neutral, más ajenas al resto de situaciones de color **azul**.

Este ejercicio, además de validar la consistencia interna de la base de datos, ofrece una clara visión: los **clusters** de **correlación** efectivamente capturan estilos de inversión análogos en su mayoría según el rendimiento observador en el periodo 1997 - 2018.

6. PCA - Análisis de los componentes principales

El **PCA** reduce la dimensionalidad al construir nuevos ejes (**componentes**) que capturan la mayor parte de la variación total . Esto permite resumir la información de múltiples estrategias en unos pocos **factores**, atribuyéndole mayor importancia a algunos factores “**main driven**”.

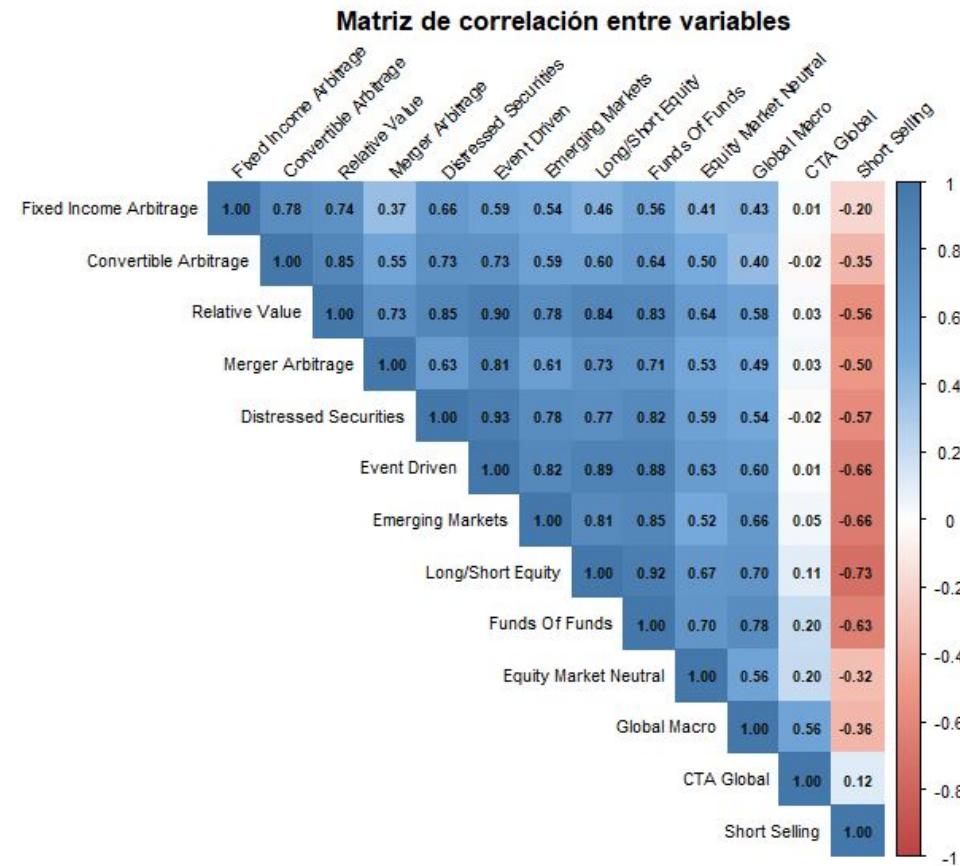
Se extraen componentes ordenados por **varianza** explicada. El primer componente (PC1) es la combinación lineal de estrategias que maximiza la varianza , y cuantifica el “factor común” dominante.

El **objetivo** será analizar los PC del dataset en general y determinar los mismos de cada cluster.

Para esto, el plan a ejecutar sería el siguiente:

1. Setear semilla para garantizar reproducibilidad y transformar dataset.
2. Análisis exploratorio inicial.
3. Verificación de supuestos a asumir para el cálculo del PCA.
4. Cálculo de PCA, criterios de números de componentes a retener y visualizaciones
5. Interpretación de los componentes.

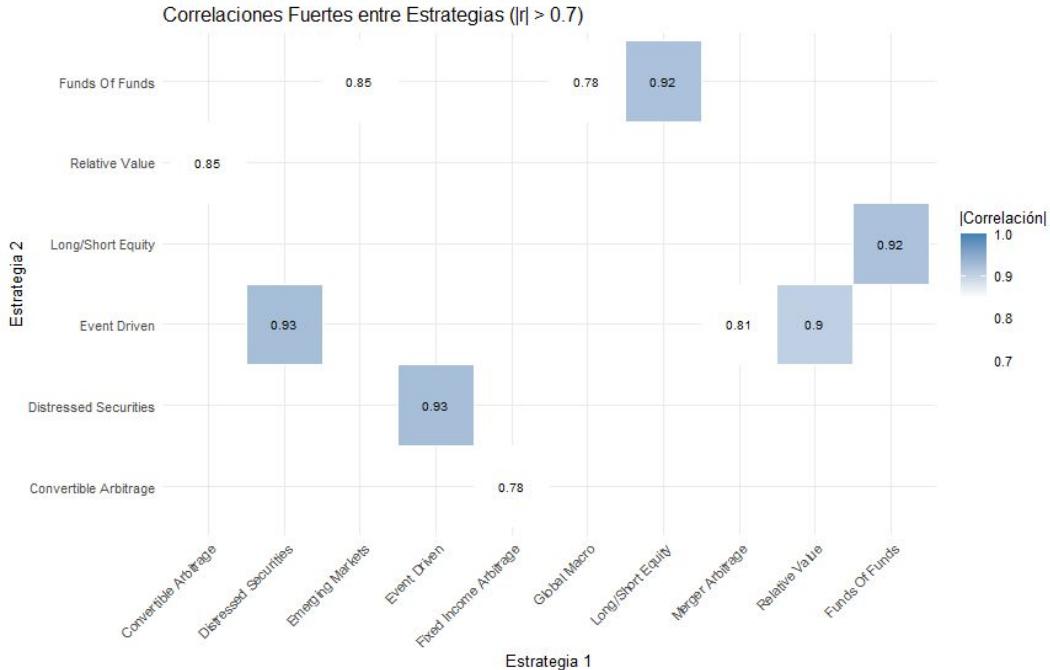
6. PCA - Matriz de correlación entre estrategias



La matriz de correlación permite visualizar fácilmente **qué estrategias tienden a moverse juntas** (alta correlación) y **cuáles ofrecen potencial de diversificación** (baja o negativa correlación).

- La mayoría de las estrategias muestran **correlaciones positivas moderadas a altas** (>0.6), lo que sugiere que comparten ciertos drivers comunes de mercado.
- **Short Selling** se presenta como una **estrategia atípica** con **correlaciones negativas** frente a casi todas las demás. Esto es esperable: su exposición corta al mercado accionario implica que gana cuando otras estrategias pierden, y viceversa.
- **CTA Global** (estrategia de futuros gestionados) aparece **aislada** con correlaciones muy bajas. Su naturaleza direccional y sistemática, basada en tendencias macroeconómicas, le otorga un perfil de retorno diferente al del resto de las estrategias más tradicionales.

6. PCA - Correlaciones fuertes



El heatmap que vemos a continuación nos permite visualizar correlaciones fuertes >0.7 , lo que indica que tienden a moverse de forma muy similar a lo largo del tiempo.

- **Relaciones altamente redundantes:** estrategias como *Distressed Securities* y *Event Driven*

Lo más importante quizá sea reforzar la idea de una **Aplicación práctica** para la preparación de una cartera diversificada, **no conviene combinar estrategias altamente correlacionadas**, ya que no contribuyen a reducir el riesgo total. Por el contrario, estrategias con baja o nula correlación (como **Short Selling** o **CTA Global**, vistas en otras figuras) sí podrían aportar beneficios en términos de **diversificación**.

5. Supuestos PCA

Se realizó una breve validación de los supuestos necesarios para el cálculo del PCA. Los resultados son muy favorables:

Proporción de muestras variables (n/p)

- Obs: 263, Var: 13 → n/p = 20.23
- Ratio ideal: estimaciones estables y confiables

Test de Kaiser-Meyer-Olkin (KMO) Global

- KMO global > 0.8: MERITORIO (muy bueno para PCA)

Test de Kaiser-Meyer-Olkin (KMO) por variable

- CTA Global tiene un KMO individual < 0.5 (0.447), lo que sugiere bajo nivel de correlación con el resto de variables
- Resto de variables con KMO > 0.8

CTA Global podría revisarse o interpretarse con mayor cautela.

Test de Esfericidad de Bartlett:

- $\chi^2 = 4328.15$, gl = 78, p < 0.001
- Se rechaza H_0 (matriz de correlaciones ≠ identidad): PCA es pertinente

Determinante de la matriz de correlaciones

- Determinante = 4.8e-08
- Refleja alta colinealidad, las variables contienen redundancia, por ende es ideal para PCA.

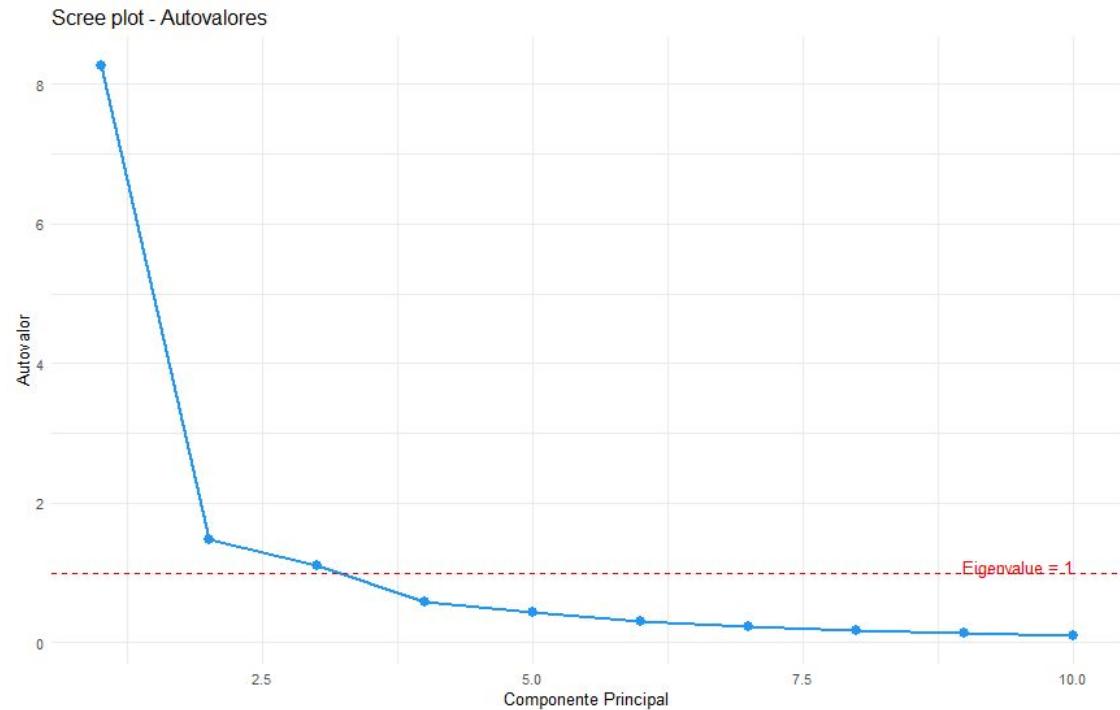
5. PCA Evaluación de la varianza explicada

¿Cuántos componentes retener?

Resultados del PCA:

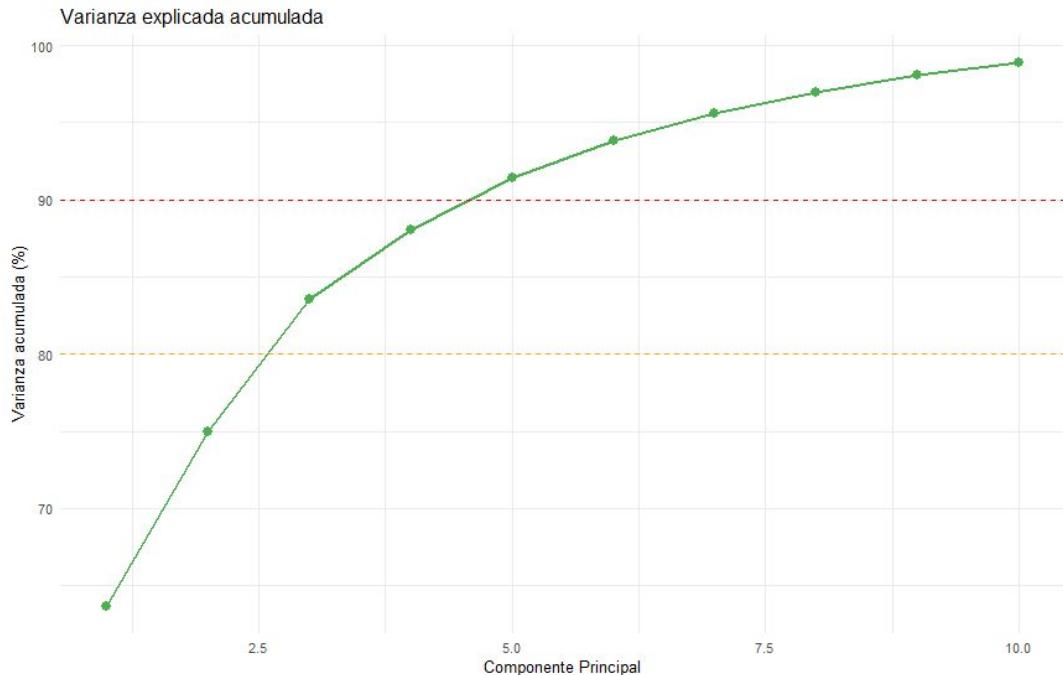
PC1: 63.66% (acumulado: 63.66%)
PC2: 11.32% (acumulado: 74.98%)
PC3: 8.56% (acumulado: 83.55%)
PC4: 4.52% (acumulado: 88.07%)
PC5: 3.38% (acumulado: 91.45%)
PC6: 2.35% (acumulado: 93.8%)
PC7: 1.8% (acumulado: 95.61%)
PC8: 1.37% (acumulado: 96.98%)
PC9: 1.13% (acumulado: 98.11%)
PC10: 0.81% (acumulado: 98.92%)

Con **3 componentes** alcanzamos el **83.55%** de la varianza acumulada.



6. PCA - Varianza explicada acumulada

¿Qué tan representativa es la reducción obtenida?



Este gráfico continua lo mencionado en la anterior slide

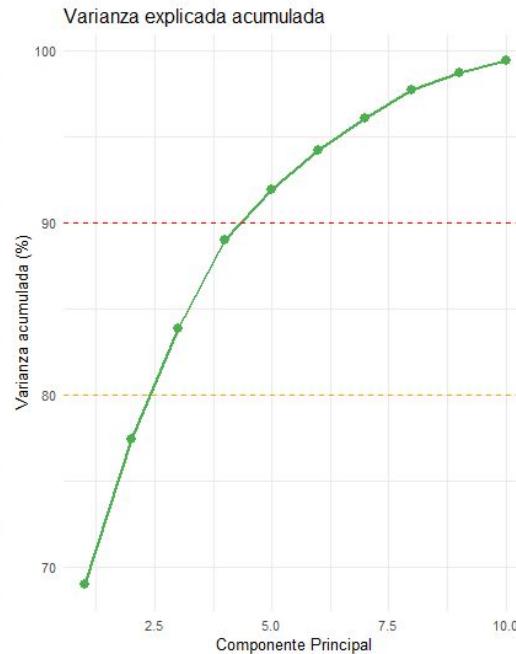
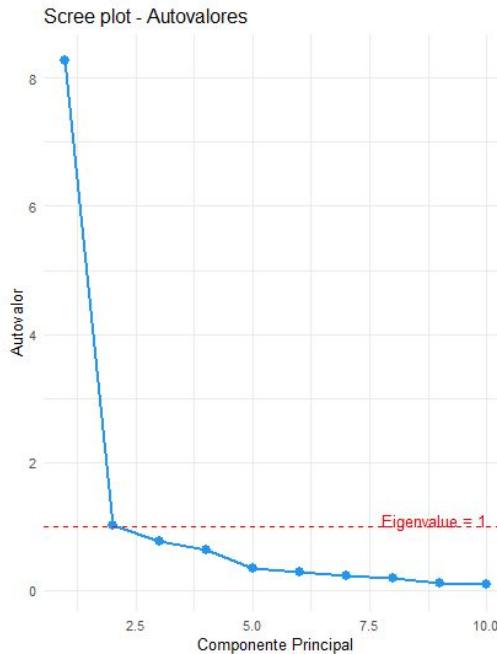
A partir del cuarto componente, cada nueva dimensión agrega muy poca información marginal

Algunos criterios más comunes a la hora de elegir el número de componentes a retener son:

- **Kaiser ($\lambda > 1$)**: dando 3 componentes retenidos.
- **$\geq 80\%$ varianza acumulada**: también se alcanzan con 3 componentes.
- Para una visión más exhaustiva ($\geq 90\%$), podrían considerarse 5 componentes.

6. PCA - Nueva evaluación?

¿Qué sucede si decidimos no tener en cuenta la estrategia CTA Global para el cálculo del PCA?



Resultados de PC:

#PC	% Varianza	% Var Acum	Eigenvalue
PC1	68.97	68.97	8.276
PC2	8.43	77.39	1.011
PC3	6.43	83.82	0.771
PC4	5.18	89.00	0.622
PC5	2.91	91.90	0.349
PC6	2.29	94.19	0.275
PC7	1.86	96.06	0.224
PC8	1.64	97.69	0.196
PC9	0.99	98.68	0.119
PC10	0.71	99.39	0.085

Algunos criterios más comunes a la hora de elegir el número de componentes a retener son:

- **Kaiser ($\lambda > 1$):** dando 2 componentes retenidos.
- **$\geq 80\%$ varianza acumulada:** también se alcanzan con 3 componentes.
- Para una visión más exhaustiva ($\geq 90\%$), podrían considerarse 5 componentes.

6. PCA - Nueva evaluación?

Los resultados parecen ser bastante similares, pero el **PCA sin CTA Global** parece captar el objetivo de encontrar alguna “**fuerza de mercado**” que explique el movimiento en el rendimiento de las estrategias.

Además el hecho de que la estrategia **CTA Global** tenga una correlación tan baja con el resto nos induce a pensar que la misma no es afectada por los mismos factores, esto es resaltado en el **test KMO**, por ende se decidió continuar con el PCA “sin CTA Global” de ahora en más para la interpretación de sus componentes principales.

La elección de cuantos PC retener será completamente discrecional, el presente trabajo no busca más que confirmar y afirmar suposiciones que se adapten mejor a la realidad.

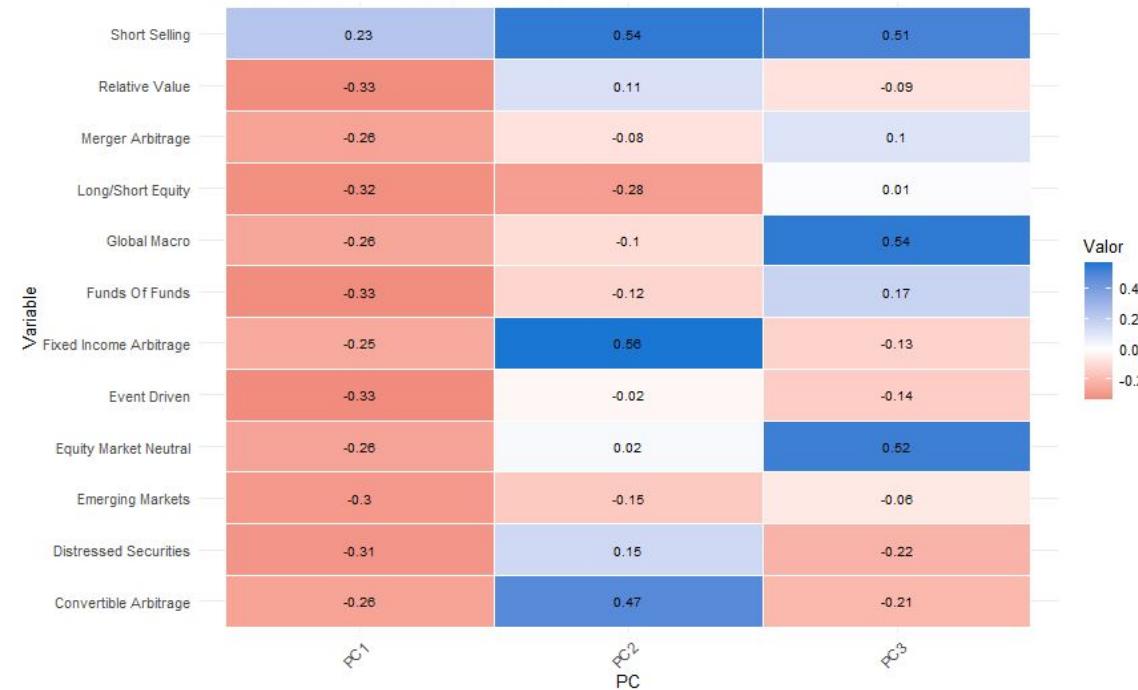
Además el hecho de interpretar los PC suele ser dificultoso por sus carácter meramente matemático.

Resultados de PC:

#PC	% Varianza	% Var Acum	Eigenvalue
PC1	68.97	68.97	8.276
PC2	8.43	77.39	1.011
PC3	6.43	83.82	0.771
PC4	5.18	89.00	0.622
PC5	2.91	91.90	0.349
PC6	2.29	94.19	0.275
PC7	1.86	96.06	0.224
PC8	1.64	97.69	0.196
PC9	0.99	98.68	0.119
PC10	0.71	99.39	0.085

6. PCA - Interpretación de los PC

Cargas de variables en componentes (PC1-PC3)

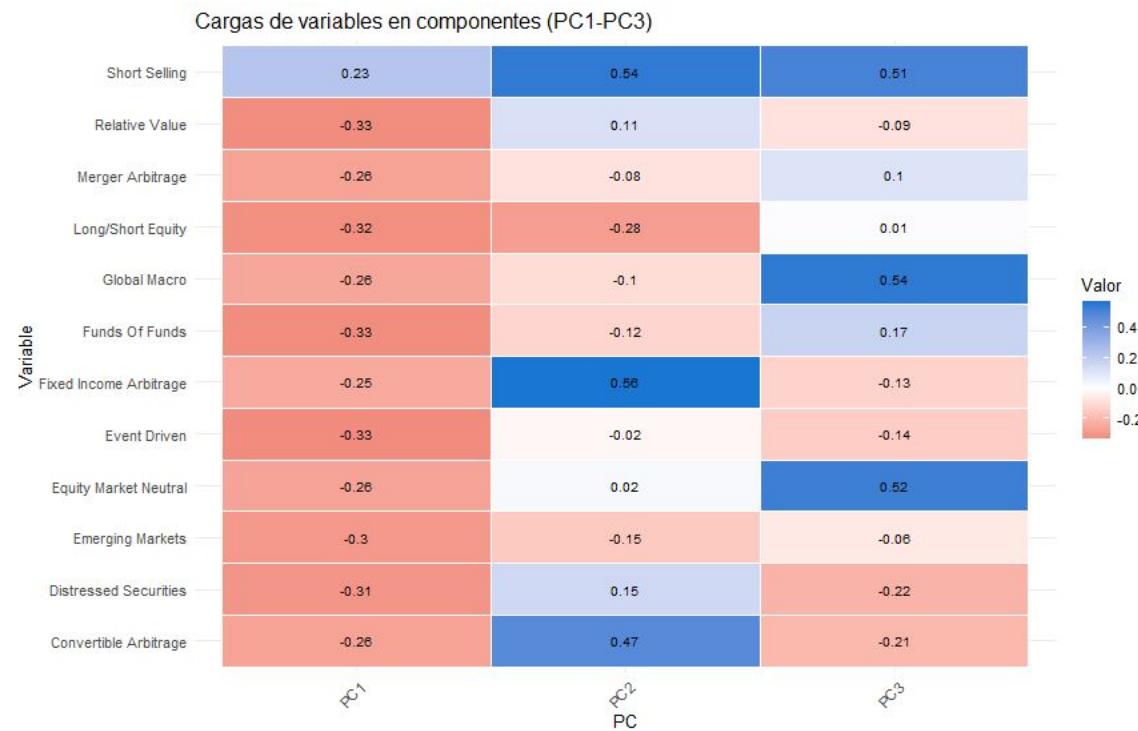


El PC1 presenta algunas características típicas:

- Parece estar dominado por estrategias direccionales (*Long/Short Equity, Event Driven, Distressed, Emerging Markets*), que son influenciadas por el rumbo general de los mercados.
- Explica el 68% de la varianza del conjunto (primer componente \approx “**beta**” común). Parece capturar el **riesgo sistémico de mercado**, similar a un índice de mercado amplio.
- En términos prácticos, es como un **factor** agregado del mercado de capitales en el que participan la mayoría de las estrategias.
- La relación general **negativa** para con las estrategias pero justamente **positiva** para con **Short Selling** reafirman las suposiciones anteriores.

PC1 podría ser catalogado como un **factor direccional de mercado**, justamente que explica la mayoría de movimientos de las estrategias, donde estas suelen ser direccionales.

6. PCA - Interpretación de los PC

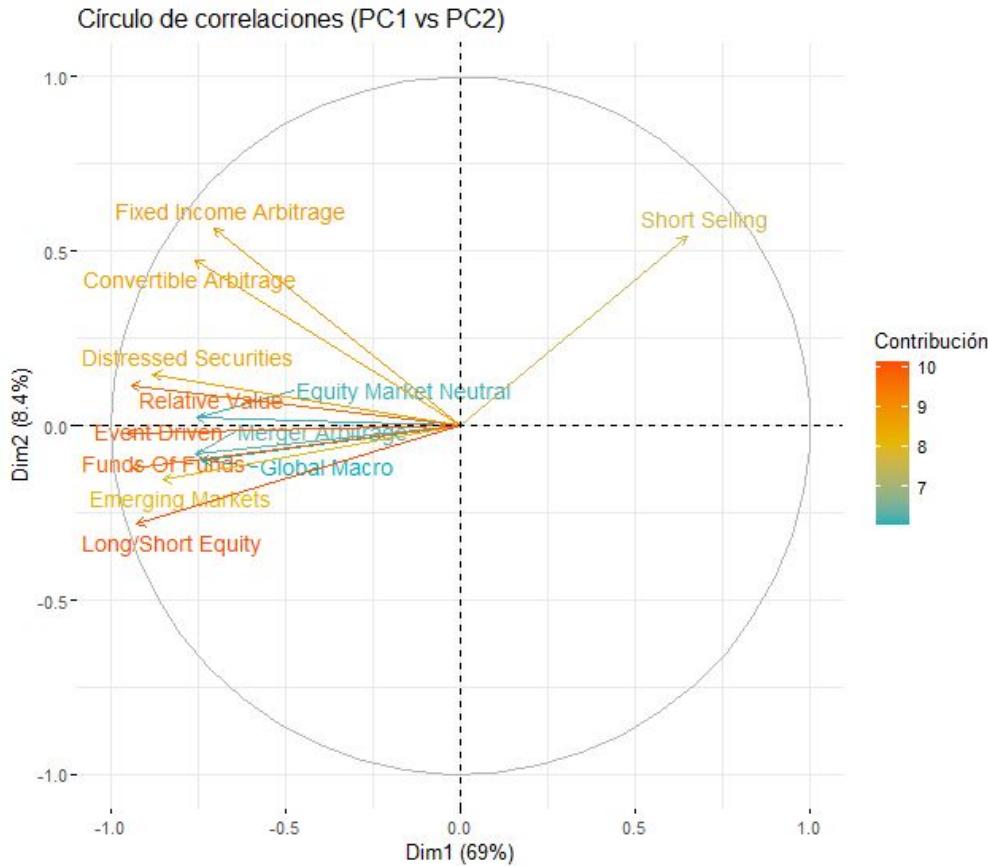


El análisis de **PC2** y **PC3** no resulta tan visible en cambio, además c/u explica **8.43%** y **6.43%** respectivamente (**83.82%** acum), por lo cual su interpretación es difícil. Aun así es posible sacar algunas características para su interpretación:

- Estrategias discrecionales y no neutrales suelen mostrar baja correlación o negativa, reflejando su baja relación con este **factor**.
- En el extremo opuesto destaca **Short Selling** (venta en corto) con cargas inversas así también como estrategias más neutrales al mercado.
- Aun así este patrón no se muestra continuamente para con el resto de estrategias neutrales como por ej **Relative Value**

En conjunto y usando la imaginación, podemos pensar que **PC2** y **PC3** diferencian estrategias con perfil **"short volatility"** (venden volatilidad y/o apuntan a un carry estable) vs. **estrategias direccionales** o long volatility (aprovechan movimientos de mercado y volatilidad).

6. PCA - Visualizaciones Círculo de correlaciones PC1 vs PC2

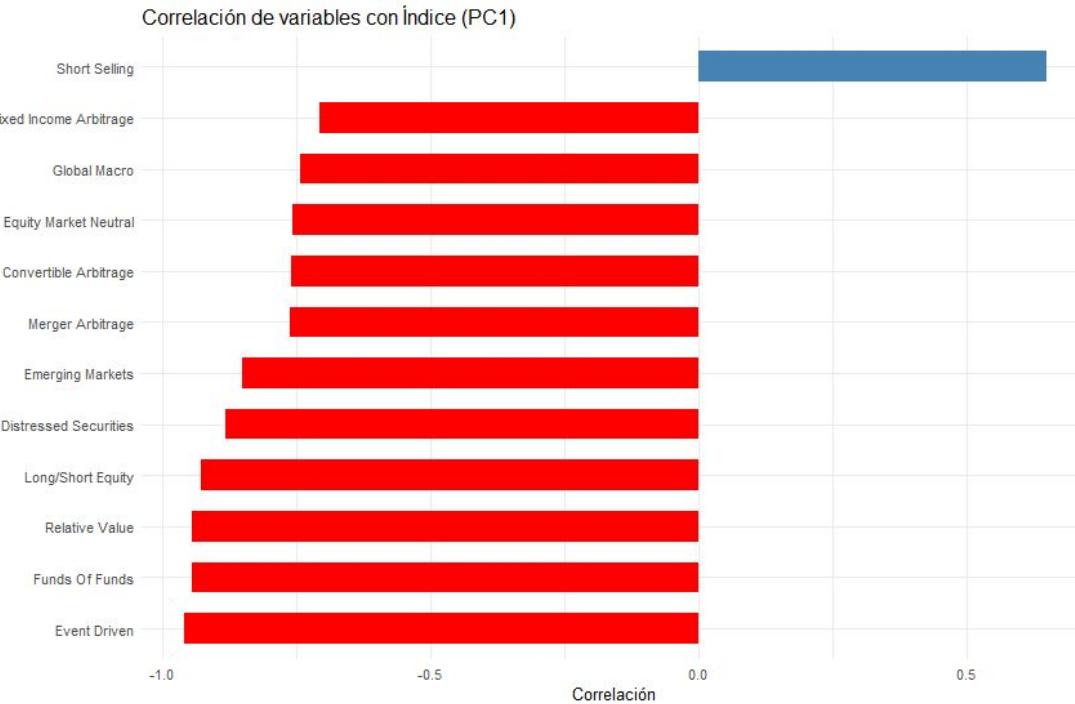


El gráfico de **Círculo de correlaciones** contiene en eje “X” el **PC1** y en eje “Y” el **PC2**, explicando así el 77% de la varianza acumulada.

Podemos identificar fácilmente según el color de las flechas, la contribución de cada variable para la formación de cada **PC**.

Se puede apreciar la “contrariedad” de **Short Selling** para con el resto de estrategias en **PC1**, además de la clara visualización de direccionalidad (beta mercado) por parte de las estrategias que “derivan” principalmente el **PC1**.

6. PCA - Índice basado en PC1



Un índice compuesto con base en el **primer componente principal (PC1)** se construye como una combinación lineal de las variables originales ponderadas por sus *loadings* (pesos) en PC1.

El índice actúa como un **proxy del factor común de mercado** que más varianza explica (~69%).

- Al ser PC1 negativo para casi todas las estrategias, un **aumento en el índice implica rendimientos por debajo del promedio** en la mayoría de ellas. Por el contrario, **Short Selling se mueve de forma inversa**, reflejando su rol como cobertura frente a shocks del mercado.

Correlaciones con el índice:

- Estrategias como **Event Driven, Funds of Funds** muestran correlaciones cercanas a **-1**, se alinean fuertemente con el **factor común**.
- **Short Selling**, en cambio, presenta **correlación positiva significativa (+0.65)** ⇒ es la única que se mueve en dirección opuesta, confirmando su valor diversificador.

6. PCA - Conclusiones

El análisis de **Componentes Principales (PCA)** permitió reducir la complejidad de las 13 estrategias en factores subyacentes, conservando más del 83% de la varianza explicada con tan solo **3 componentes**.

El **PC1**, que explica el 68.97% de la varianza, se interpreta como un **factor direccional de mercado**: las estrategias con alta exposición al beta del mercado presentan cargas negativas, mientras que **Short Selling** actúa como contracara, con una correlación positiva al índice derivado. Esto sugiere que **PC1** capta el “**riesgo sistemático agregado**”.

El **índice compuesto basado en PC1** permite visualizar este “factor común de mercado” de forma sintética. Su interpretación es práctica: cuando el índice cae, las estrategias tienden a rendir mejor, y viceversa. Esto aporta una herramienta útil para **análisis de riesgo agregado**.

En cuanto a **PC2** y **PC3**, su interpretación resulta más sutil: no parecen explicar una única dimensión clara, pero en conjunto podrían reflejar la dicotomía entre estrategias “**short volatility**” (**carry, estabilidad**) vs. “**long volatility**” (**direccionalidad y dislocaciones**), a partir de sus patrones de correlación cruzada.

La separación de **Short Selling** y estrategias de valor relativo frente a estrategias direccionales sugiere una **estructura latente en los retornos**, y valida la existencia y creencia “lógica” en el mercado de factores financieros comunes.

Conclusiones finales generales

El presente análisis tuvo como objetivo la exploración profunda de datos, sobre las **estrategias** que los **hedge funds** emplean, en el periodo 1997-2018. El propósito fue identificar **patrones comunes**, evaluar su exposición al **riesgo**, y detectar posibles factores latentes que puedan explicar su **movimiento conjunto** o diferencial de los rendimientos.

Para ello, el trabajo se estructuró en diversas etapas complementarias. En primer lugar, se realizaron **visualizaciones exploratorias** que permitieron observar el comportamiento general de los rendimientos. Luego, se aplicaron contrastes de hipótesis (**ANOVA**) para evaluar diferencias de medias según riesgo, sin evidencia concluyente de que las estrategias difieran significativamente en términos de rendimiento esperado.

Posteriormente, se avanzó hacia un **análisis de conglomerados (clustering)** que permitió agrupar estrategias con patrones similares. Algunos grupos resultaron particularmente **interesantes y consistentes**.

Finalmente se testeó una reducción de dimensiones por **análisis de los componentes principales (PCA)**, donde se identificó un **primer componente (PC1)** que explica cerca del **69%** de la variabilidad conjunta, lo cual sugiere la presencia de un “**factor de mercado dominante**” que tiende a arrastrar a la mayoría de las estrategias en la misma dirección. Este factor podría interpretarse como una especie de “beta sistémico” de los hedge funds. A su vez, se identificaron estrategias particularmente diferenciadas, como **Short Selling** o **CTA Global**, confirmando los clusters formados.

Conclusión

Finalmente, la combinación del análisis multivariado permitió validar empíricamente ciertos comportamientos esperados (correlaciones fuertes, direccionalidad compartida), además de también proponer nuevas hipótesis sobre el rol de las estrategias en distintos entornos de volatilidad. El trabajo logró cumplir con su objetivo inicial: **entender, clasificar y sintetizar el universo de estrategias de hedge funds**, revelando tanto su homogeneidad estructural como su heterogeneidad operativa.

Bibliografía y mas

Base de datos:

- Dataset - [EDHEC Hedge Fund Index Return](#)

Referencias

- CFA Institute - [Hedge Fund Strategies](#)
- CANDRIAM - [CTA Commodity Trading Advisors](#)
- HFR - [Strategy Classification System](#)

Consultas: thiagoigbenitez@gmail.com

**Thiago Benitez - FCE - UBA
Diciembre 2025**