

**Universidad de Costa Rica**

Facultad de Matemáticas

Estadística actuarial 1: Bitácora 2

Impacto en los rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses, consecuente de la crisis sanitaria del Covid-19.

Integrantes:

Edwin Amador Abarca | B80429

Francisco José Cambronero Quirós | B71494

Johan Castaño Bustamante | B71744

05 de Octubre de 2023

# Índice

<b>1. Parte de Planificación</b>	<b>3</b>
1.1. Pregunta de investigación . . . . .	3
1.1.1. Definición de la Idea . . . . .	3
1.1.2. Conceptualización de la Idea . . . . .	3
1.1.3. Identificación de Tensiones . . . . .	3
1.1.4. Reformulación de la Idea en modo pregunta . . . . .	3
1.1.5. Argumentación de la pregunta . . . . .	4
1.1.6. Argumentación a través de datos . . . . .	7
1.2. Revisión Bibliográfica . . . . .	9
1.2.1. Construcción de fichas de literatura . . . . .	9
1.2.2. Teorías y Principios . . . . .	18
<b>2. Parte escrita</b>	<b>18</b>
<b>3. Ordenamiento de la Literatura</b>	<b>20</b>
<b>4. Enlaces de la literatura</b>	<b>21</b>
<b>5. Manipulación de los datos</b>	<b>25</b>
5.1. Análisis descriptivo . . . . .	25
5.2. Propuesta metodológica . . . . .	28
5.3. Fichas de resultados . . . . .	30
<b>Referencias</b>	<b>32</b>
<b>A. UVE heurística</b>	<b>33</b>
A.1. Código . . . . .	34

# 1. Parte de Planificación

## 1.1. Pregunta de investigación

### 1.1.1. Definición de la Idea

Impacto en los rendimientos anuales de los Fondos de Pensión Obligatoria Complementaria Costarricenses, consecuente de la crisis sanitaria del Covid-19.

### 1.1.2. Conceptualización de la Idea

- Impacto: “Efecto producido en la opinión pública por un acontecimiento, una disposición de la autoridad, una noticia, una catástrofe, etc.”(RAE, 2023)
- Rendimiento de Capital: “Rendimiento consistente en las utilidades o contraprestaciones, cualquiera que sea su denominación o naturaleza, dinerarias o en especie, que provengan directa o indirectamente de elementos patrimoniales, bienes o derechos, cuya titularidad corresponda al contribuyente y no se hallen afectos a actividades económicas.”(RAE, 2023)
- Fondo de Pension Obligatoria Complementaria: Se conforma de los aportes obligatorios que el trabajador (afiliado al sistema nacional de pensiones) realiza y es de capitalización individual. Estos aportes los registra y controla la CCSS y los administra la operadora de pensiones elegida por el trabajador (ROP). Los montos que recibirá el pensionado por concepto de este pilar, dependen del aporte a lo largo del tiempo (que se calcula con base en el salario reportado ante la CCSS) y otros detalles como los rendimientos de las inversiones que haya logrado la operadora de pensiones (SUPEN, 2022)
- Fondo de Pensiones: “Los fondos de pensiones son patrimonios sin personalidad jurídica que pertenecen a una pluralidad de partícipes, cuyo derecho de propiedad se justifica mediante certificados de participación. Los fondos de pensiones se constituyen con objeto exclusivo de dar cumplimiento a los planes de pensiones. Su constitución debe inscribirse en el Registro Mercantil del domicilio de la Entidad gestora y la denominación del mismo, irá seguida por las siglas F.P.”.(INE, 2020)
- Consecuente: “Una proposición que se deduce de otra que se llama antecedente.” (RAE, 2023)
- Crisis: “Un cambio profundo y de consecuencias importantes en un proceso o una situación, o en la manera en que estos son apreciados.”(RAE, 2023)
- Sanitaria: “Pertenciente o relativo a la sanidad.”(RAE, 2023)
- Pensión: “Cantidad periódica, temporal o vitalicia, que la seguridad social paga por razón de jubilación, viudedad, orfandad o incapacidad.”(RAE, 2023)

### 1.1.3. Identificación de Tensiones

- Se requiere implementar un modelo con suficiente precisión dada la baja frecuencia de las observaciones a lo largo de los períodos estudiados.
- La conclusión obtenida dará una idea general sobre el comportamiento de los Fondos de Pension Obligatoria Complementaria Costarricenses, sin embargo, no mostrará la situación general de todos los Fondos de Pensiones del país, en especial los del sector privado.
- Implementar un método para proyectar los datos de los rendimientos en el fondo de pensiones complementario costarricense considerando los datos prepandémicos.

### 1.1.4. Reformulación de la Idea en modo pregunta

- ¿Cuál es el impacto que tuvo la crisis sanitaria del Covid-19 sobre los rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses?
- ¿Cómo se puede observar el impacto de la crisis sanitaria del Covid-19 sobre los rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses?
- ¿Por qué es importante conocer el impacto de la crisis sanitaria del Covid-19 sobre los Rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses??
- ¿Qué modelos permiten una visualización más precisa acerca del impacto de la crisis sanitaria del Covid-19 sobre los Rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses??

#### 1.1.5. Argumentación de la pregunta

¿Cuál es el impacto que tuvo la crisis sanitaria del Covid-19 sobre los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses?

##### Contraargumentos:

- **Lógica:** Los movimientos en los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses durante la crisis sanitaria del Covid-19 no tienen que estar necesariamente relacionados con dicha crisis.
- **Ética:** Asumir que los movimientos en los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses ocurridos durante la crisis sanitaria del Covid-19 podría generar conclusiones falsas ya que no está asegurada la relación entre dichos movimientos y la crisis sanitaria..
- **Emocional:** Considerar información no relacionada entre los movimientos de los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses y la crisis sanitaria del Covid-19 puede provocar un sesgo en la manera en la que las personas pensarán que en épocas de pandemia, los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses se van a comportar de una manera específica, lo cuál va a afectar su toma de desiciones.

##### Argumentos:

- **Lógica:** La crisis sanitaria trajo consigo varias muertes y una cuarentena que afectó directamente a la economía de las personas en el país de manera negativa, incluidas las personas contribuyentes en el sistema del Fondo de Pension Obligatoria Complementaría Costarricenses, por lo que es notorio que el período de pandemia afecta directamente a los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses.
- **Ética:** Una vez asegurada la existencia de una relación entre la crisis sanitaria del Covid-19 y los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses, se puede realizar una selección de los datos de interés para obtener dicha relación con presición y exactitud.
- **Emocional:** Al obtener la relación entre la crisis sanitaria del Covid-19 y los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses y conocer las variables que tienen relevancia en dicha relación, se puede informar y compartir con el público de manera segura, de modo que las personas podrán conocer y entender el funcionamiento del mismo, y a su vez estar preparados pára posibles eventualidades en el futuro.

**Conclusion:** Es importante poder observar que existió un impacto en los rendimientos anuales del los Fondos de Pension Obligatoria Complementaría Costarricenses debido a la crisis sanitaria del Covid-19, y una vez comprendida dicha existencia, buscar los datos pertinentes que revelen cuál es dicha relación, y ls misma deberá ser utilizada para facilitar el proceso de respuesta y seguridad para las personas contribuyentes y para el Fondo en sí.

## ¿Cómo se puede observar el impacto de la crisis sanitaria del Covid-19 sobre los Rendimientos anuales de los Fondos de Pension Obligatoria Complementaria Costarricenses?

### Contraargumentos:

- **Lógica:** Existen muchas variables ajenas a la crisis sanitaria del covid-19 que pueden generar impacto en los fondos de pensiones, lo que puede generar un sesgo al observar la información de la época determinada.
- **Ética:** El enfoque en pocos o insuficientes fondos de pensiones pueden generar un sesgo en los resultados al no contemplar un panorama completo de los datos de todos los fondos de pensiones.
- **Emocional:** Entender el concepto de impacto solo con cantidades monetarias abre la posibilidad de que solo se contemple un aspecto económico en dicho impacto cuando el mismo puede conformarse por más aspectos.

### Argumentos:

- **Lógica:** Al notar grandes cambios entre los períodos anteriores y posteriores a la pandemia, al igual que datos que mantengan relación durante el período de la crisis sanitaria, se puede concluir que el mayor factor que afectó a la obtención de estos resultados es la crisis sanitaria del covid-19.
- **Ética:** Los fondos de pensiones poseen una funcionalidad general, comparten muchos aspectos en común, por lo que el estudio de pocos fondos de pensiones nos permite formar una idea general o resumen sobre el comportamiento general de los fondos en el país..
- **Emocional:** Los fondos de pensiones trabajan en su mayor parte con movimientos de dinero y la inversión del mismo, de modo que se puede entender y derivar los otros posibles resultados de interés en base a los datos monetarios que sean utilizados.

**Conclusion:** A la hora de realizar observaciones con respecto a datos referentes a los fondos de pensiones durante la época de la crisis sanitaria del Covid-19, se debe de tomar en cuenta factores significativos que revelen cambios o efectos relevantes en los rendimientos de los mismos, al igual que tomar en cuenta a suficientes fondos de pensiones en base a la exactitud o resultado que se desee obtener, ya sea un estudio de un fondo en específico o de un grupo de estos, al igual que es necesario relacionar la información financiera o económica recopilada con las otras áreas en las que estas se conecten, como la calidad de vida de las personas, contribuciones en el sistema, entre otros.

## ¿Por qué es importante conocer el impacto de la crisis sanitaria del Covid-19 sobre los Rendimientos anuales del los Fondos de Pension Obligatoria Complementaria Costarricenses?

### Contraargumentos:

- **Lógica:** la importancia del impacto de la crisis sanitaria del Covid-19 es meramente de utilidad y de relevancia para aquellas personas que tengan alguna relacion con los fondos de pensiones, pero para el resto de la poblacion que no esta inmiscuida no es de importancia.
- **Ética:** Conocer el impacto de la crisis sanitaria en los distintos fondos de pensiones puede generar filtraciones de información acerca de la situación personal de las personas colaboradoras o integrantes del sistema de pensiones, la cuál puede utilizarse para tomar desiciones relacionadas con estos integrantes, abusando de sus situaciones personales.
- **Emocional:** Las personas podrían no estar realmente interesadas en los movimientos de rendimientos de los fondos de pensiones, de modo que su falta de interés o comprensión puede generar conflictos en el sistema al que pueden colaborar.

### Argumentos:

- **Lógica:** El area de pensiones tiene una estrecha relación con la economía del país, la cuál afecta y colabora con el bienestar y la calidad de vida de las personas en el mismo, de modo que es importante conocer los efectos que ocurren en esta area para poder asociarlos con otros aspectos que modifican el día a día de las personas.
- **Ética:** Al conocer el impacto que tiene la crisis sanitaria en los fondos de pensiones puede ayudar a una mejor respuesta y toma de desiciones por parte de los fondos de pensiones en respuesta a la crisis y problemas en la economía que enfrentaran.
- **Emocional:** Si se utilizan métodos para llamar la atención de las personas y que éstas aprendan del funcionamiento del sistema de fondos de pensiones fomentará la participación y la contribución en el mismo, de modo que fortalecerá la eficacia de dicho sistema.

**Conclusion:** El área de fondos de pensiones en el país tiene una estrecha relación con el vivir de las personas, en áreas como el empleo, salario, seguridad y confort, de modo que eventos que puedan afectar el funcionamiento de los fondos de pensión van a afectar directa o indirectamente a estos otros aspectos, de modo que la información recopilada referente a esta area puede mostrar datos relevantes acerca de la vida de las personas relacionadas (como los contribuyentes), la cuál deberá ser respetada y no utilizarse para procesos cuyo fin sea aprovecharse de la situación de algunas personas y a su vez poder informar a las personas y llamar su atención para poder encontrar brechas o problemas en el sistema de fondos de pensiones, de modo que dicho sistema pueda obtener mejorías gracias a la colaboración de las personas interesadas.

## ¿Qué modelos permiten una visualización más precisa acerca del impacto de la crisis sanitaria del Covid-19 sobre los rendimientos anuales del los Fondos de Pension Obligatoria Complementaria Costarricenses?

### Contraargumentos:

- **Lógica:** Una baja precisión en las observaciones obtenidas referentes al rendimientos anuales del los Fondos de Pension Obligatoria Complementaria Costarricenses durante la época estudiada puede conllevar limitaciones al utilizar e implementar un modelo que no contrarreste dicha imprecisión.
- **Ética:** El modelo elegido podría utilizar técnicas que no se adapten de manera correcta a la realidad para obtener dicha precisión, lo que causaran aproximaciones inexactas.
- **Emocional:** El hecho de que el modelo que se utilice pueda dar resultados inexactos, que no se verán reflejados en la realidad va a disminuir la confianza de las personas en dicho modelo, cuando es posible que este se haya utilizado en un contexto diferente para que el fue diseñado.

### Argumentos:

- **Lógica:** Mediante una investigación rigurosa es posible encontrar un modelo útil para sobrellevar las limitaciones causadas por la pérdida de precisión causada por una posible baja frecuencia de las observaciones.
- **Ética:** Con una debida elección de los modelos utilizados se puede evitar sesgar los datos y así mantener la exactitud del estudio.
- **Emocional:** Una implementación correcta del modelo puede contemplar y ponderar la realidad de los rendimientos del periodo estudiado, haciendo justicia al escenario real y no afectar a la confianza de las personas involucradas.

**Conclusion:** El modelo implementado debe combinar teoría estadística y computacional de modo que se logren sobrellevar las limitaciones dada la posibilidad de que existan sesgos debido a la baja frecuencia de los datos, de igual manera es importante tener en consideración la ponderación que puedan tener los distintos sectores a los rendimientos totales por periodo. Cabe mencionar que a lo largo del proyecto se debe tener en mente que el estudio realizado se apegue a la realidad, para obtener resultados que hagan justicia a esta y, de esta manera, no afectar a la confianza o seguridad de las personas contribuyentes.

### 1.1.6. Argumentación a través de datos

Para el presente proyecto se trabajará con una base de datos que consta de los rendimientos por mes y año de los Fondos de Pension Obligatoria Complementaria Costarricenses durante el período 2008 - 2023.

La base de datos es de fácil acceso y gratuita, se puede encontrar en formato csv en el siguiente enlace: <https://webapps.supen.fi.cr/contenedor/2>.

La población estudiada corresponde a aquella que participa o aporta al Fondos de Pension Obligatoria Complementaria Costarricenses en el período de 2008 - 2023 en Costa Rica.

La unidad estadística de los datos corresponde al monto de dinero en millones de colones y dólares.

**Las variables de interés son las siguientes (SUPEN, 2023):**

- MES: Numero del mes y año correspondiente
- BAC SJ PENSIONES Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del BAC SJ PENSIONES
- BAC SJ PENSIONES Rentabilidad Anual: Rentabilidad Porcentual Anual de Fondo de pensiones del BAC SJ PENSIONES
- BCR-PENSION Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del BCR-PENSION
- BCR-PENSION Rentabilidad Anual: Rentabilidad Porcentual Anual de Fondo de pensiones del BCR-PENSION
- BN-VITAL Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del BN-VITAL
- BN-VITAL Rentabilidad Anual: Rentabilidad Porcentual anual de Fondo de pensiones del BN-VITAL
- CCSS-OPC Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del CCSS-OPC
- CCSS-OPC Rentabilidad Anual: Rentabilidad Porcentual Anual de Fondo de pensiones del CCSS-OPC
- POPULAR PENSIONES Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del POPULAR PENSIONES
- POPULAR PENSIONES Rentabilidad Anual: Rentabilidad Porcentual Anual de Fondo de pensiones del POPULAR PENSIONES
- VIDA PLENA OPC Rentabilidad Historica: Rentabilidad Porcentual Historica de Fondo de pensiones del VIDA PLENA OPC
- VIDA PLENA OPC Rentabilidad Anual: Rentabilidad Porcentual Anual de Fondo de pensiones del VIDA PLENA OPC



## 1.2. Revisión Bibliográfica

### 1.2.1. Construcción de fichas de literatura

Tabla 1.  
**Ficha de Global Economic Prospects**

Título:	Global Economic Prospects
Autor(es):	Banco Mundial
Año:	2020
Nombre del tema:	COVID19 y el paro economico en las economias avanzadas y otras partes del globo
Cronológico:	2020
Metodológico:	Proceso de pronóstico, utilizando modelos econométricos que permiten la combinación de juicio y coherencia con conocimientos basados en modelos
Temático:	Método de pronóstico y combinación de juicio.
Teoría:	Implicaciones macroeconómicas del Covid-19
Resumen en una oración:	Informe del pronóstico de la economía global posterior al Covid-19
Argumento central:	El informe abarca las consecuencias del Covid-19 en detalle considerando tanto regiones desarrolladas como regiones en vías de desarrollo.
Problemas con el argumento o el tema:	En el informe se indica una limitación de los datos, dado que en períodos anteriores se encuentra como única fuente estadísticos sesgados (Pues no dan una idea clara del panorama completo).
Resumen en un párrafo:	Se analiza el ambiente económico en el 2020 viendo tanto las tendencias globales como los cambios politicos. Se intenta hacer una estimación de lo que pueda seguir a este período partiendo del impacto inicial. Se ve el efecto en el desarrollo económico a corto y largo plazo. Se hace una observación de como el bajo precio del petroleo durante la pandemia puede llegar a afectar. Menciona como gracias al Covid-19 podría generar una de las más grandes recesiones económicas, así como también cómo este impacto no solamente se centra en las zonas económicas con mucho ingreso, sino a nivel global

Tabla 2.

**Ficha acerca de los Efectos económicos de la pandemia del Covid-19, en los principales fondos de pensiones privado de Colombia**

Título:	EFFECTOS ECONÓMICOS DE LA PANDEMIA COVID 19, EN LOS PRINCIPALES FONDOS DE PENSIONES PRIVADOS DE COLOMBIA
Autor(es):	Darío Ernesto Sánchez Manosalva, Cristian Fabían Díaz León
Año:	2022
Nombre del tema:	Implicaciones del Covid-19 en fondos de pensiones colombianos
Cronológico:	2020 - 2021
Metodológico:	Análisis cuantitativo relacional, causal y descriptivo. Así como estimaciones mediante el modelo VAR
Temático:	Análisis descriptivo y VAR
Teoría:	Implicaciones macroeconómicas en fondos de pensiones del COVID19
Resumen en una oración:	Implicaciones financieras del Covid-19 en los fondos de pensiones en Colombia.
Argumento central:	Se analizan los rendimientos de los fondos de las ATP colombianas marcadas por el Covid-19 utilizando el VAR
Problemas con el argumento o el tema:	Al usarse el VAR (value at risk) se tiene que no es coherente como medida de riesgo pues no cumple la propiedad de subaditividad.
Resumen en un párrafo:	Se evalúan los rendimientos de los fondos de pensiones durante 2020 - 2021 Se hace uso del modelo VAR para ver la recuperación posterior a la caída. Se observa que el fondo moderado, que tiene una mayor participación en portafolios, presenta un impacto mayor en la rentabilidad total Se menciona que el fondo de menor riesgo presenta una recuperación más rápida sufriendo una menor caída Se tiene que el fondo moderado enfrenta mejor las posibles inestabilidades del mercado

Tabla 3.  
**Ficha de Probability and Statistics**

Título:	Probability and Statistics
Autor(es):	Morris H. DeGroot, Mark J. Schervish.
Año:	2002
Nombre del tema:	Prueba de hipótesis Kolmogorov-Smirnov
Cronológico:	1933
Metodológico:	Presentación de conceptos y teoría fundamental referente a los temas de probabilidad y la estadística, con énfasis en la comprensión conceptual y la aplicación práctica para la enseñanza y aprendizaje.
Temático:	Enseñanza y Educación
Teoría:	Probabilidad y Estadística
Resumen en una oración:	Teoría y ejercicios referentes a la materia de Probabilidad y Estadística.
Argumento central:	Teoría de probabilidad de Kolmogorov
Problemas con el argumento o el tema:	Se deben especificar (y conocer de antemano) varios parámetros de ubicación, escala y forma para poder aplicar el modelo de interés.
Resumen en un párrafo:	El libro abarca una amplia gama de temas para la enseñanza, entre ellos: probabilidad, variables aleatorias, distribuciones de probabilidad, estadísticas descriptivas, estimación de parámetros, pruebas de hipótesis (aquí aparece el modelo de Kolmogorov-Smirnov) y análisis de regresión.

Tabla 4.

**Ficha impacto de la crisis económica por Covid-19 en el Sistema de pensiones Mexicano**

Título:	Impacto de la crisis económica por COVID-19 en el sistema de pensiones mexicano y perspectivas ante el proyecto de su reforma
Autor(es):	Marissa R. Martínez-Preece, Carlos Zubieta-Badillo
Año:	2020
Nombre del tema:	COVID-19 y su efecto en el sistema de pensiones de México
Cronológico:	2020
Metodológico:	Investigación acerca del funcionamiento de distintas medidas tomadas por diferentes países en relación con cambios en los sistemas de pensiones y sobre las repercusiones en el área financiera relacionadas con la crisis sanitaria del Covid-19.
Temático:	Análisis descriptivo
Teoría:	Implicaciones de la crisis sanitaria del Covid-19 en el área financiera y en los sistemas de pensiones.
Resumen en una oración:	Consecuencias de la crisis sanitaria del Covid-19 en el sistema de pensiones mexicano.
Argumento central:	Funcionamiento y relación del sistema de pensiones mexicano con la economía, y los cambios que ocurren en ella.
Problemas con el argumento o el tema:	Se refiere al sistema de pensiones mexicano, el cuál tiene algunas diferencias con el costarricense.
Resumen en un párrafo:	<p>Se realizó un estudio relacionado con el sistema de pensiones mexicano y las consecuencias que provocó la crisis sanitaria del Covid-19 en el mismo, utilizando distintos supuestos con respecto a los efectos causados por la pandemia en la economía del país, afectando tanto económica como socialmente a la vida de las personas.</p> <p>Estos efectos que sucedieron inmediatamente a la crisis son analizados para corroborar qué tan mayor o menor es el impacto, los cuáles se van a relacionar también con las reformas que recibió (recibirá) el sistema de pensiones mexicano, al ser comparado con reformas tomadas por otros países y las consecuencias que tuvieron al tomar dichas reformas.</p>

Tabla 5.  
**Ficha de Los planes de Pensiones en la Crisis del Covid-19**

Título:	Los planes de pensiones en la crisis de la COVID-19: Impacto y consideraciones en materia de políticas
Autor(es):	Csaba Feher, Ignatius de Bidegain
Año:	2020
Nombre del tema:	Impacto y consideraciones en los planes de pensiones debido al Covid-19
Cronológico:	2020
Metodológico:	Descripción de consecuencias principales relacionadas con los planes de pensiones públicos producto del Covid-19
Temático:	Análisis Descriptivo
Teoría:	Implicaciones inmediatas del COVID19 en el ámbito de las pensiones públicas
Resumen en una oración:	Consecuencias inmediatas del Covid-19 en los planes de pensiones públicos
Argumento central:	Estas notas abarcan los efectos inmediatos del Covid-19 a los planes de pensiones públicos, esto principalmente en tres vías, entre las cuales se abarca la recesión de la economía, el impacto al ámbito laboral especialmente las personas mayores con posibilidad de pensiones anticipadas y el shock negativo en los activos.
Problemas con el argumento o el tema:	El enfoque que se le da en la nota es en un ámbito considerablemente general del ámbito público, con lo cual no necesariamente se ajusta a la situación Costarricense, además de que es una nota que se desarrolló durante la primera parte de la pandemia, con lo cual los datos utilizados pueden reflejar la situación del momento, sin embargo no visualiza el ámbito general de la pandemia, con lo cual algunos de los factores o impactos pueden haber disminuido de impacto.
Resumen en un párrafo:	<p>Se analiza el período del 2020 apartir de las consecuencias inmediatas que tuvo el Covid-19 con respecto a los planes de pensiones públicos.</p> <p>Se abarca la consecuencia de la recesión de la economía y como esta afecta en la reducción de los rendimientos, lo cual produce un flujo de efectivo menor y por tanto no se alcanza a cubrir las prestaciones, esto genera, que los fondos de pensiones estatales con respecto a sus flujos y activos disminuyeron en un promedio de 5%.</p> <p>Disminución en la demanda de mano de obra, lo cual disminuyó los ingresos percibidos por los planes de pensiones públicos que dejaron de percibir ingresos de las personas desempleadas, además hay que considerar la gran cantidad de pensionados anticipadas debido a que en el sistema de pensiones público se permite las pensiones anticipadas si cumple con unos requerimientos, esto provocó una disminución en los flujos de efectivo esperado.</p>

Tabla 6.

**Ficha de Retiro de Fondos de Pensiones por el Covid-19 y su Impacto en la Rentabilidad AFP**

Título:	El retiro de fondos de pensiones por el COVID-19 y su impacto en la rentabilidad de las AFP, Perú-2020
Autor(es):	Giulianna Pilar Soto Escobar
Año:	2021
Nombre del tema:	retiro de fondos de pensiones por el COVID-19 y su impacto en la rentabilidad de las AFP Perú-2020
Cronológico:	2019-2020
Metodológico:	El método cuantitativo o método tradicional
Temático:	Enfoque cuantitativo
Teoría:	Disminucion Rentabilidad de los fondos de Pensiones de Peru
Resumen en una oración:	impacto negativo en la rentabilidad de la AFPs debido al retiro de fondos por el Covid-19
Argumento central:	Impacto en la rentabilidad de los fondos de pensiones de AFPs debido al retiro de fondos
Problemas con el argumento o el tema:	No se usa un metodo de correlacion o comparacion estadística, aparte de la comparacion de graficos con los valores de rentabilidad
Resumen en un párrafo:	<p>Se realiza una investigacion con respecto a la rentabilidad de los fondos de pensiones relacionados con las AFPs y se presentan en graficos en los cuales se evidencia una clara caida en las rentabilidades nominales en el periodo de pandemia.</p> <p>Se desarrolla un contraste entre los diferentes rendimientos de los fondos de pensiones que conforman las AFPs, así como un contraste con respecto no solo a los rendimientos nominales, sino también sobre el patrimonio y los activos.</p> <p>Se desarrolla el metodo de prueba de hipotesis para poder determinar el impacto en los rendimientos nominales debido a el retiro de fondos, en el cual se escoje como hipotesis nula el echo de que el retiro de fondos no tuviera efectos negativos, con lo cual se logra determinar que si hubo impactos negativos en los rendimientos.</p>

Tabla 7.  
**Ficha de Guía del Alumno del curso MA145 de la UPC**

Título:	Guía del Alumno
Autor(es):	Cuadros G, Tarazona E, Cardenas C, Ramirez R
Año:	2013
Nombre del tema:	Estadística Aplicada 2
Cronológico:	2013
Metodológico:	Presentación de materia referente al curso MA145 (Estadística Aplicada 2) de la Universidad Peruana de Ciencias Aplicadas (UPC)
Temático:	Enseñanza y Educación
Teoría:	Estadística
Resumen en una oración:	Teoría referente al muestreo, prueba de hipótesis, Distribución Ji cuadrada, Diseños Experimentales, Regresión Lineal Simple, Regresión Múltiple, Series de Tiempo y el Método de Atenuación Exponencial
Argumento central:	Herramientas e Instrumentos para la realización de procesos estadísticos
Problemas con el argumento o el tema:	Se refiere a la materia completa de un curso impartido por la UPC, de modo que no todo el material del documento es de nuestro interés, habrá que hacer una separación y lectura primero para poder separar los temas de interés para el trabajo.
Resumen en un párrafo:	Se refiere a la materia de un repositorio de la Universidad Peruana de Ciencias Aplicadas, separado por distintos temas vistos en el curso de MA145 (Estadística Aplicada 2), entre ellos se encuentran: conceptualización y técnicas de muestreo, pruebas de hipótesis (tema de interés para esta investigación), usos de la distribución Ji cuadrada, diseños experimentales, regresión lineal simple, Regresión múltiple, series de tiempo, y métodos de atenuación exponencial.

Tabla 8.  
**Ficha de Prueba de Hipótesis**

Título:	Prueba de Hipótesis
Autor(es):	Fallas J
Año:	2012
Nombre del tema:	Prueba de significancia, Prueba de hipótesis o contrastes
Cronológico:	2012
Metodológico:	Presentación de supuestos y procedimientos estadísticos relacionados con la prueba de hipótesis.
Temático:	Enseñanza y Educación
Teoría:	Estadística
Resumen en una oración:	Teoría referente a la materia de la prueba de hipótesis.
Argumento central:	Prueba de hipótesis
Problemas con el argumento o el tema:	Se requiere de cierto nivel de conocimiento en temas estadísticos (como el intervalo de confianza o nivel de significancia) que es posible que no se hayan visto en el curso para el momento del inicio de la lectura acerca del tema.
Resumen en un párrafo:	En el documento podemos hallar varias definiciones y teoría, empezando con datos experimentales y las distribuciones de frecuencias teóricas, luego de detallar las mismas, obtenemos las definiciones de hipótesis nula ( $H_0$ ), hipótesis alternativa ( $H_a$ ) y cómo la $H_0$ es aquella que se somete a prueba (Rechazo o No Rechazo), también se define el "efecto de confusión", el "nivel de significancia", errores tipo I y II, "potencia de la prueba", "error marginal" se utilizan en ejemplos prácticos para poder desarrollar la prueba de hipótesis, ya sea direccionada o no direccionada, y cómo esta prueba es utilizada en la comparación de medias y/o varianzas



Tabla 9.  
**Ficha de Modelos de Regresión Lineal Múltiple**

Título:	Modelos de regresión lineal múltiple
Autor(es):	Roberto Montero Granados
Año:	2016
Nombre del tema:	Modelos de regresión lineal múltiples
Cronológico:	2016
Metodológico:	Presentación de conceptos y teoría relacionada a los procesos de regresión lineal, enfocado en en la comprensión conceptual y la aplicación práctica de la misma.
Temático:	Enseñanza y Educación
Teoría:	Regresión lineal múltiple
Resumen en una oración:	Curso introductorio para el desarrollo de modelos de regresión lineal múltiple.
Argumento central:	Regresión lineal múltiple
Problemas con el argumento o el tema:	Explcia exclusivamente sobre modelos de regresión lineal, cuando no necesariamente será el método que mejor se acople a nuestros objetivos.
Resumen en un párrafo:	Se explica y detalla la teoría detrás de la creación de modelo de regresión lineal múltiple, explicando distintos conceptos que deben conocerse con antelación y realizando pruebas o experimentos de referencia, esto con el fin de entender los modelos que se muestran en el documento y para la futura comprensión de modelos menos intuitivos.

Tabla 10.  
**Ficha de Regresión Lineal**

Título:	Regresión lineal simple y polinomial: teoría y práctica
Autor(es):	Pablo Vinuesa
Año:	2016
Nombre del tema:	Modelos de regresión lineal simple y polinomial
Cronológico:	2016
Metodológico:	Presentación de conceptos y teoría relacionada a los procesos de regresión lineal simple y polinomial, enfocado en en la comprensión conceptual y la aplicación práctica en lenguaje de programación..
Temático:	Enseñanza y Educación
Teoría:	Regresión lineal simple y polinomial
Resumen en una oración:	Capítulo relacionado a la enseñanza y ejercicios sobre modelación estadística de regresión lineal y polinomial aplicado en el lenguaje de programación R.
Argumento central:	Regresión lineal y polinomial en R
Problemas con el argumento o el tema:	Da una explicación sobre modelos de regresión lineal y polinomial, cuando no necesariamente será el método que mejor se acople a nuestros objetivos.
Resumen en un párrafo:	Se explica y detalla la teoría detrás de la modelación de regresión lineal y polinomial en el ámbito estadístico, realizando varios ejemplos y ejercicios en código en R, modificando y depurando tablas de datos, graficación, detallando tanto cómo hacerlo de manera manual o explicando cómo lo realiza R con los paquetes respectivos.

Tabla 11.  
**Ficha ADF**

Título:	TESTING TIME SERIES DATA FOR STATIONARITY
Autor(es):	Rizwan Mushtaq
Año:	2011
Nombre del tema:	TIME SERIES DATA FOR STATIONARITY
Cronológico:	2011
Metodológico:	Descripción de estacionariedad y diferentes métodos para calcularla, usando Dickey and Fuller
Temático:	Educativo y Explicativo
Teoría:	Pruebas para estacionariedad
Resumen en una oración:	Descripción detallada del funcionamiento de Dickey and Fuller para la determinación de estacionariedad
Argumento central:	Dickey and Fuller y la estacionariedad
Problemas con el argumento o el tema:	La utilización solamente de Dicky fuller para el Analisis..
Resumen en un párrafo:	Descripción de los diferentes métodos de Dickey and Fuller para la determinación de estacionariedad, como es el caso formal de la prueba de estacionariedad en donde $Y_t$ y $Y_{t-1}$ representan la serie temporal en esos tiempos el $p$ representa la dependencia de la serie de tiempo con respecto al tiempo anterior y el $\epsilon$ el error, además se plantea la hipótesis nula igual a 0, y la alternativa menor que 0. además se plantea la evaluación de hipótesis nula a partir de la fórmula $t$ , en donde si $t$ es más grande que el valor crítico no se descarta la hipótesis nula

### 1.2.2. Teorías y Principios

El impacto en los rendimientos de los Fondos de Pension Obligatoria Complementaria Costarricenses consecuente a la crisis sanitaria del Covid-19 se puede especular en lo mencionado por WB (2020), en donde se evidencia cómo en el período del 2020, en particular el primer semestre de dicho período, hubo una caída en los precios a nivel mundial, así como también una especulación de una posible recesión de gran impacto, además, también se evidencia con un impacto económico no solo de las principales potencias a nivel mundial, sino que incluso los países con poca cantidad de ingresos se vieron afectados, todo esto lleva a que, considerando que los fondos de inversión se vieran afectados por esta caída de la economía, los rendimientos generados por estos hayan disminuido.

Del mismo modo, como menciona Dario Sanchez (2022) con respecto a los fondos de pensiones de Colombia, los rendimientos de dichos sufrieron una gran disminución, al punto en que se dieron rendimientos negativos, considerando esta información, es común teorizar que en el caso de Costa Rica, en particular el Fondo Popular de Pensiones, también se vio afectado en sus rendimientos.

## 2. Parte escrita

Es normal asumir o pensar directamente que, al darse un evento catastrófico en el mundo que amenace la comodidad de las personas y la manera en la que viven su vida de manera negativa, ya sea con un restricciones para salir de su hogar, muertes en grandes cantidades o miedo por una enfermedad, como lo fue con el covid-19, esto genere problemas o caídas en los sistemas económico del país y y del mundo en general , ya que las personas se ven limitadas o afectadas directamente en sus capacidades para trabajar y, por ende, contribuir en dicho sistema.

Este pensamiento es usual debido a las noticias que ocurren en el mundo, especificando como la economía posee una decadencia durante estos eventos, por ejemplo, según WB (2020), .El impacto súbito y generalizado de la pandemia del coronavirus y las medidas de suspensión de las actividades que se adoptaron para contenerla han ocasionado una drástica contracción de la economía mundial, que, según las previsiones del Banco Mundial, se reducirá un 5,2% este año”.

Del mismo modo, podemos rescatar una idea de la misma noticia, “Los efectos están siendo particularmente profundos en los países más afectados por la pandemia y en aquellos que dependen en gran medida del comercio internacional, el turismo, las exportaciones de productos básicos y el financiamiento externo” WB (2020), lo que nos hace preguntarnos cuál es el nivel del impacto que generó la pandemia en Costa Rica, siendo este un país en el cual parte importante de su economía consiste en el área de turismo y el comercio de productos.

Se han hecho varios estudios en los cuáles se analiza los efectos que ha tenido la pandemia y crisis sanitaria del covid-19 en la economía del mundo, y cómo cada país ha recibido cierto impacto de la misma, hemos notado que el efecto es diferente en base a las medidas que se tomaron o al sistema económico que se maneje en el país de estudio.

Podemos ver como en Colombia, según Dario Sanchez, mediante un modelo VAR (Value at Risk) se utilizaron “los promedios de rendimientos de los diferentes fondos de manera conjunta para analizar los impactos que tendrían los fondos en relación al total de la rentabilidad total de los fondos”, de modo que se obtuvieron resultados en donde “los fondos de mayores riesgos sufrieron en promedio una mayor caída que sus demás contrapartes” (Dario Sanchez, 2022).

De esta manera, se considera la opción de realizar un estudio similar para conocer dicho impacto en Costa Rica y así poder comparar los resultados con los de otros países, al igual que se pueden comparar las medidas tomadas para ver si estas fueron factibles o si existieron mejores maneras de atender la crisis sanitaria.

Conociendo la existencia de distintos modelos para la obtención de esta clase de impactos que puedan ocasionarse debido a la pandemia del covid-19, es de nuestro interés utilizar algún otro modelo conocido, en este caso hemos elegido utilizar la prueba de Kolmogorov-Smirnov para hipótesis simples DeGroot (2002), la cuál es de nuestro interés debido a la relación que posee con el curso de Estadística Actuarial, el cuál hará énfasis sobre dicho modelo durante el transcurso del semestre, pero es importante tener una idea básica del objetivo de este modelo, el cuál es, principalmente, evaluar si una muestra de datos proviene de una distribución específica, de esta manera se pueden comparar distintas colecciones de datos en base a sus distribuciones.

Debido a esta relación observable entre la economía de algún lugar y la pandemia, nos interesa realizar este estudio para poder utilizar dicha información para estar preparados ante posibles eventualidades en el futuro y poder manejarlas de una manera más adecuada.-

### 3. Ordenamiento de la Literatura

Tabla 12: Fuentes Descriptivas

Organización			Literatura		
Tipo	Tema General	Tema específico	Titulo	Año	Autor
Descriptivo	Covid-19 y el paro económico en las economías avanzadas y otras partes del globo	Consecuencias del Covid-19 en detalle, considerando tanto regiones desarrolladas como regiones en vías de desarrollo	Global Economic Prospects	2020	Banco Mundial
Descriptivo	Covid-19 y su efecto en el sistema de pensiones de México	Funcionamiento y relación del sistema de pensiones mexicano con la economía, y los cambios que ocurren en ella.	Impacto de la crisis económica por Covid-19 en el sistema de pensiones mexicano y perspectivas ante el proyecto de su reforma	2020	Marissa R. Martínez-Preece, Carlos Zubieta-Badillo
Descriptivo	Planes de pensiones en la crisis del Covid-19: Impacto y consideraciones en materia de políticas	Impacto y consideraciones en los planes de pensiones debido al Covid-19	Los planes de pensiones en la crisis del Covid-19: Impacto y consideraciones en materia de políticas	2020	Csaba Feher, Ignatius de Bidegain

Tabla 13: Fuentes miden el impacto en los rendimientos con diferente metodos

Organización			Literatura		
Tipo	Tema General	Tema específico	Titulo	Año	Autor
Contraste y prueba de Hipotesis	Retiro de fondos de pensiones por el covid-19 y su impacto en la rentabilidad de las AFP Perú-2020	Impacto en la rentabilidad de los fondos de pensiones de AFPs debido al retiro de fondos utilizando Contrastes y prueba de Hipotesis	El retiro de fondos de pensiones por el Covid-19 y su impacto en la rentabilidad de las AFP, Perú-2020	2021	Giulianna Pilar Soto Escobar
Análisis Descriptivo y Var	Implicaciones del Covid-19 en fondos de pensiones	Análisis de los rendimientos de los fondos de las ATP colombianas marcado por el Covid-19 utilizando el VAR	EFFECTOS ECONÓMICOS DE LA PANDEMIA del Covid-19, EN LOS PRINCIPALES FONDOS DE PENSIONES PRIVADOS DE COLOMBIA	2022	DARIO ERNESTO SANCHEZ MANO-SALVA, CRISTIAN FABIAN DIAZ LEON

Tabla 14: Fuentes Metodológicas

Organización			Literatura		
Tipo	Tema General	Tema específico	Titulo	Año	Autor
Metodológico	Probabilidad y Estadística	Prueba de hipótesis Kolmogorov-Smirnov	Probability and Statistics	2002	Morris H. DeGroot, Mark J. Schervish.
Metodológico	Prueba de Hipótesis	Estructura y desarrollo de prueba de hipótesis	Prueba de Hipótesis	2012	Fallas J
Metodológico	Estadística Aplicada 2	Prueba de Hipótesis	Guía del Alumno	2013	Cuadros G, Tarazona E, Cardenas C, Ramirez R
Metodológico	Modelos de regresión lineal simple y polinomial	Regresión Lineal	Regresión lineal simple y polinomial: teoría y práctica	2016	Pablo Vinuesa

## 4. Enlaces de la literatura

El Covid-19 generó un cambio en el ámbito general de la vida de las personas, ya que demostró lo vulnerable que pueden ser las personas ante una crisis repentina en este caso una pandemia, esto llevó a consecuencias sociales, culturales, sanitarias, económicas, entre otras. En el aspecto económico, el cual es de interés para el desarrollo de esta investigación se encuentra lo mencionado por WB (2020), en donde se abarcan aspectos de que el Covid-19 se podría generar una de las más grandes recesiones económicas, así como también como este impacto no solamente se centra en las economías con muchos ingresos, sino más bien es algo que cubre de manera global hasta el punto de que en el momento en que se desarrolló el Covid-19, el Banco Mundial consideró posible que fuera una de las peores recesiones desde la segunda guerra mundial, esto sin considerar la gran caída de precios producto de la epidemia, esto muestra el gran impacto económico que el Covid-19 trajo consigo para la economía ya que no solamente afectó el comercio, el cual se detuvo en gran medida debido a las medidas tomadas por los diferentes gobiernos para detener la propagación de la epidemia, sino que también el sector financiero se vio en gran medida, ya que debido a la caída de precios masiva, los precios de contratos, bonos, acciones y entre otros aspectos que componen el sistema financiero global se vieron sumamente afectados, esto lleva a que en referencia con el sistema de pensiones de los diferentes países, así como también los fondos que los componen, hayan desarrollado rentabilidades bajas o incluso negativas, esto debido a que los fondos de pensiones invierten los fondos recaudados de sus contribuyentes en diferentes áreas de la economía tanto nacional como internacional, por lo que los valores de estas inversiones, así como también los rendimientos generados se vieron sumamente afectados.

En relación con los Sistemas de Pensiones los fondos que los componen, se puede además considerar lo mencionado por Feher y de Bidegain (2020), en donde se evidencia las consecuencias inmediatas del Covid-19 con respecto a los planes de pensiones, además de que se menciona como a su vez la recesión o contracción de la economía presente disminuyó los flujos de efectivo y el valor de los activos, esto ya que al haber una contracción de la economía, los rendimientos con respecto a los fondos de pensiones llegaron a incluso números negativos y en promedio a una reducción del 5%, por último menciona como debido a esta recesión de la economía se generó una disminución en la demanda de mano de obra, lo que provocó a su vez una disminución de los flujos percibidos por los fondos de pensiones, ya que al haber un aumento de desempleo debido al Covid-19 generó que las personas que quedaron desempleadas disminuyeran o no aportaran al monto de pensión al cual estaban inscritos. Lo abarcado por el autor evidencia claramente como a nivel mundial durante el periodo del 2020 se evidencia una grave recesión económica que afectó no solo de manera superficial a los fondos de pensiones, sino que fue un impacto considerable al que fueron expuestas, además se evidencia de manera más clara como el impacto más significativo percibido por los fondos de pensiones fue la reducción de los rendimientos, así como también una disminución de los flujos de efectivo percibidos, esto evidencia claramente como el Covid-19 tuvo un impacto significativo en los fondos de pensiones a nivel mundial y más particularmente se puede hablar de un impacto negativo contundente en los rendimientos y en los valores de los activos.

Sin embargo es de considerar que los efectos mencionados anteriormente, son efectos que se produjeron a nivel global y de una manera muy centrada en los Estados Unidos, con lo cual es necesario considerar más específicamente regiones más cercanas al objeto de estudio para que se pueda analizar de manera más específica cuales de estos problemas se desarrollaron con mayor particularidad en las regiones cercanas a Costa Rica,

con lo cual en lo mencionado por Dario Sanchez (2022) se puede visualizar que en el caso de los Fondos de Pensiones Privados en Colombia durante los años 2020-2021, los fondos de pensiones privados, tuvieron una caída considerable al inicio de la pandemia 2020, además se evidencia como los fondos de pensiones que tienen mayor participación en portafolios de inversión, se vieron como los más afectados con respecto a la rentabilidad total, además de que se presenta como los fondos con menor riesgo son los menos afectados con respecto al Covid-19, esto ya que los fondos de pensiones privados en general manejan 3 variantes, las cuales van de la más riesgosa a la más conservadora, siendo la más riesgosa los que tienen más proporción invertida en la bolsa y sectores donde se genera más rentabilidad. Lo mencionado por el autor evidencia como los rendimientos de los fondos de pensiones privados de Colombia se vieron afectados, en especial aquellos con seleccionado con riesgo considerable, esto ya que las rentabilidades a inicios del periodo del 2020 se redujeron en gran medida, lo cual da indicios de que los fondos de pensiones relacionados con países latinoamericanos también se vieron afectados en gran medida, en especial se podría especular como los rendimiento durante el periodo a inicio del 2020 en Costa Rica, también se vieron sumamente afectados.

Además de lo mencionado anteriormente también se puede reflejar el impacto en la rentabilidad de los fondos de pensiones, no solamente en el sistema de pensión de Colombia, sino que también se puede reflejar en el de México, ya que como menciona Preece y Badillo (2020), la crisis generó un grande impacto en el sistema de pensiones de México, se plantea como la crisis económica producto del Covid-19 generó un aumento en el desempleo y de la demanda de trabajadores, lo cual influyó a que el dinero que los contribuyentes de un fondo de pensiones puedan reunir en su vida disminuya considerablemente, además se plantea como tanto el comercio internacional como la tuvieron fuertes caídas, con lo cual se puede visualizar que en el caso de México, el impacto en el comercio fue significativo. Lo mencionado por los autores refleja no solo la caída en las rentabilidades de los fondos de pensiones, sino que también refleja una disminución en los flujos percibidos por parte de los fondos de pensiones, esto al haber un aumento en el desempleo, lo cual provoca una disminución en los aportes en relación con las pensiones, lo cual significa también una disminución nominal de los rendimientos, este aspecto también se relaciona en lo que menciona Soto Escobar (2021), en donde se refleja el impacto negativo en las rentabilidades nominales producto del retiro de fondos debido al Covid-19, en donde se clarifica como se dio una disminución en la rentabilidad nominal de los diferentes fondos de pensiones de las AFPs a lo largo del periodo del 2020 debido a la reducción de la economía, así como también debido al retiro de fondos y reducción de las aportaciones de las personas inscritas.

En lo visto por lo anteriores autores, cada uno de estos utiliza algún método que les resulte útil para poder medir los impactos que tiene el Covid-19 con los diferentes sistemas de pensiones, ya sea para medir el impacto en los rendimientos debido a un factor, o para poder proyectar como se deberían haber comportado los datos para posteriormente poder medir este impacto, esto mediante diferentes métodos como lo son la prueba de hipótesis, contrastes, Var, entre otros, por lo cual para este trabajo de investigación se decidió que para poder medir el impacto en los rendimientos en el fondo de pensiones complementario costarricense se ha optado por realizar un estudio de “prueba de hipótesis”, el cual consiste en responder a la pregunta “¿existe una diferencia estadísticamente significativa entre un estimador y el parámetro de la población o entre dos estimadores?” (Fallas, 2012), de modo que podemos comparar la significancia de dos valores estadísticos.

Ya que el fin de este trabajo es obtener el impacto que tuvo la crisis sanitaria en los rendimientos de las entidades de pensiones, primero debemos tener una idea de cuál hubiera sido el rendimiento que hubiese existido si la crisis no hubiera sucedido, de modo que se planea utilizar un método para poder proyectar los rendimientos si estos se hubieran desarrollado de manera normal (sin una crisis sanitaria), para así poder comparar dichos rendimientos con los rendimientos realmente existentes y así ver el cambio o las diferencias en los mismos.

Para la proyección se utilizará un método de regresión lineal, el cual “trata de ajustar modelos lineales o linealizables entre una variable dependiente y más de una variables independientes”. (Granados, 2016)

“La regresión lineal tiene una versión simple que empareja dos variables, pero esta suele ser insuficiente para entender fenómenos mínimamente complejos en la que influyen más de dos variables, esta versión es la “múltiple”. En el modelo de regresión lineal múltiple suponemos que más de una variable tiene influencia o está correlacionada con el valor de una tercera variable.” (Granados, 2016)

Es importante utilizar una cantidad considerable de datos, mínimo 30, para poder realizar una regresión lineal confiable, sin embargo, no necesariamente los datos van a tener una tendencia lineal, por lo que el resultado de la regresión no daría un resultado confiable, así que es importante considerar esta posibilidad por si fuera necesario cambiar de método al ver el comportamiento de los datos.

Un factor importante en el proceso de regresión lineal es estimar la máxima verosimilitud, mediante el uso del método de mínimos cuadrados, se “garantiza encontrar el valor de máxima verosimilitud de los parámetros del modelo lineal para un conjunto de datos, al minimizar la magnitud de la diferencia entre los puntos y la recta”. (Vinuesa, 2016)

Se define como residuo a la distancia existente entre los puntos y la recta resultante de la regresión lineal, dichos “residuos describen la bondad de ajuste de la recta de regresión. Nuestro modelo de máxima verosimilitud se define como el modelo que minimiza la suma de los cuadrados de estos residuos”. (Vinuesa, 2016)

Una vez que tengamos la proyección, debemos definir los parámetros medios para considerar un nivel bajo,

alto o medios de rendimientos para las pensiones, y para definir estas medidas “se necesita un patrón o valor de referencia contra el cual se pueda comparar el set de datos. En el mundo de la estadística se conoce a dicho set de referencia como la distribución esperada de las variables en estudio” (Fallas, 2012)),

Ahora teniendo la distribución esperada de las variables, podremos ver las diferencias con los datos obtenidos de la base de datos, “Una vez obtenida la diferencia debe decidir si la misma es muy grande o muy pequeña (estadísticamente significativa) o si por el contrario las diferencias pueden atribuirse al efecto del azar o ruido” (Fallas, 2012).

El siguiente paso es definir la hipótesis nula ( $H_0$ ), la hipótesis alternativa ( $H_a$ ), nivel de significancia, el error, ya sea de tipo I o II y potencia de una prueba.

La hipótesis nula, “es la que se somete a prueba y sobre ella se hace la decisión. Para los propósitos de la prueba se asume como verdadera y se rechaza o no se rechaza como resultado del proceso de análisis” (Fallas, 2012).

La hipótesis alternativa se encuentra más relacionada con la razón por la cual deseamos hacer la prueba de hipótesis.

Las razones por las que se trabaja con la hipótesis nula y no la con la hipótesis alternativa son explicadas por Fallas:

“ $H_0$  es sometida a prueba en lugar de la hipótesis alternativa ( $H_a$ ) porque la serie estadística provee la información necesaria para estimar los parámetros de su distribución muestral; en tanto que  $H_a$  no ofrece esta ventaja. Por ejemplo, si sometemos a prueba la hipótesis  $H_0: \mu = 0,87 \text{ gr/cm}^3$ , asumimos que la distribución muestral de las medias está centrada en el valor 0,87. Conociendo esto podemos determinar si la media muestral corresponde o no a dicha distribución y además si el valor de la 3 media muestral es suficientemente raro (muy grande o muy pequeño) y por lo tanto poco probable como para que deba rechazarse la hipótesis nula. Por otra parte, si intentamos probar directamente la hipótesis alternativa  $H_a \neq 0,87$ , nos encontraríamos con el inconveniente de que no sabríamos donde se centra la distribución muestral de las medias, lo único que podríamos afirmar es que no se centra en el valor 0.87. Lo anterior imposibilita someter a prueba la hipótesis alternativa y a la vez justifica la necesidad de probar la validez de la hipótesis nula.” (Fallas, 2012)”

Según la UPN, se resume en una oración la relación entre ambas hipótesis: “partiendo de los resultados obtenidos de la muestra, o bien rechazamos la hipótesis nula a favor de la alternativa, o bien no rechazamos la hipótesis nula y suponemos que nuestra estimación inicial del parámetro poblacional podría ser correcto” (Fallas, 2012)) .

“El hecho de rechazar  $H_0$  significa que los datos muestrales brindan suficiente evidencia como para pensar que lo planteado por la hipótesis nula es estadísticamente improbable a un nivel de significancia dado. De la misma manera cuando no rechazamos  $H_0$  significa que los datos muestrales no brindan suficiente evidencia como para pensar que lo planteado por la hipótesis nula sea improbable a un nivel de significancia dado.”

El nivel de significancia es definido por Fallas como un indicador que indica “cuan rara (muy grande o muy pequeña) deber ser la diferencia con respecto a lo planteado por la hipótesis nula como para que sea rechazada dado que sea correcta.” (Fallas, 2012)

Con respecto a los errores, se detalla que, aunque  $H_0$  sea verdadera, se considera un porcentaje de error que define que algunas veces en las que se realice la prueba, esta sea rechazada por razones externas o azar. El error del tipo I se refiere a esta clase de errores (falso positivo).

Los errores del tipo II (falso negativo) es considerar  $H_0$  verdadera cuando no lo es.

Estos se pueden ver con el siguiente cuadro hecho por la UPN:

Tabla 15: Tipos de errores

		Información muestral	
		Aceptar $H_0$	Rechazar $H_0$
La Realidad	$H_0$ es cierta	No hay error	Error I
	$H_0$ es falsa	Error II	No hay error

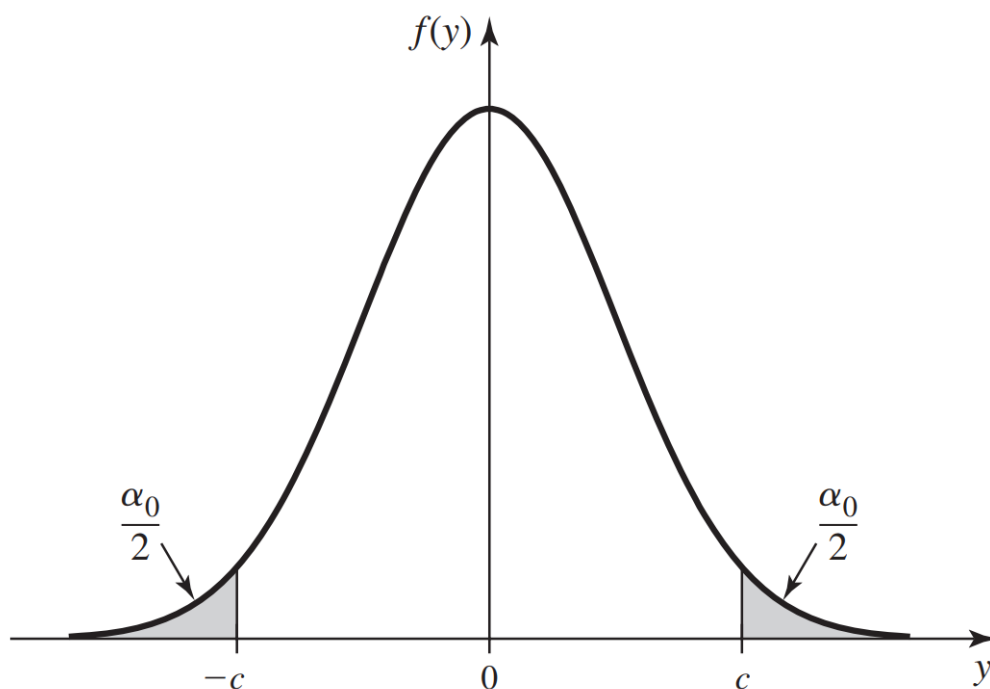
Fuente: Obtenido de (Cuadros y cols., s.f.)

La potencia de una prueba de hipótesis es definida, tanto por Fallas como por la UPN como “la probabilidad de que se rechace  $H_0$  cuando la misma es falsa” y es igual a uno menos la probabilidad del error de tipo II.

También es importante tomar en cuenta que hay dos maneras de realizar la prueba de hipótesis; direccionada o no direccionada.

La prueba direcccionada (una cola o unilateral) es aquella en la que “estamos interesados en determinar si el valor del estimador (muestra) es mayor o menor que el valor del parámetro” (Fallas, 2012)

La prueba no direcccionada es aquella en la que nos “interesa determinar si el valor del estimador es diferente al valor del parámetro, sin importar si es mayor o menor”. (Fallas, 2012)



Fuente: Obtenido de (DeGroot, 2002).

Gráfico 1: Prueba No Direcccionada



## 5. Manipulacion de los datos

### 5.1. Analisis descriptivo

Tabla 16: Cantidad de datos faltantes por fondo

Fondo	Datos Faltantes
IBP	153
INS	129

Fuente: Elaboracion Propia con datos SUPEN

Al observar la base de datos a estudiar, en la tabla 16 se nota que hay un faltante de 153 observaciones para el IBP y de 129 observaciones para INS, a diferencia del resto de variables que no presentan ningun faltante. De modo que se optó por omitir estas dos variables (tanto su version historica como su versión anual)

Naturalmente, la siguiente duda que se genera en torno a esta tabla es que tanto sus variables estan correlacionadas, lo cual se esperaria que presente una alta correlacion positiva ya que todas poseen una naturaleza similar

Tabla 17: Correlacion entre los fondos

	BAC	BCR	BN	CCSS	POP	VP
BAC	1.000	0.975	0.953	0.913	0.936	0.859
BCR	0.975	1.000	0.964	0.930	0.937	0.889
BN	0.953	0.964	1.000	0.911	0.923	0.926
CCSS	0.913	0.930	0.911	1.000	0.934	0.901
POP	0.936	0.937	0.923	0.934	1.000	0.872
VP	0.859	0.889	0.926	0.901	0.872	1.000

Fuente: Elaboracion Propia con datos SUPEN

Ahora, en la tabla 17 se ve la correlacion esperada. Luego, sigue ver el comportamiento de cada variable entonces analisamos los datos y obtenemos como resultado la tabla 18

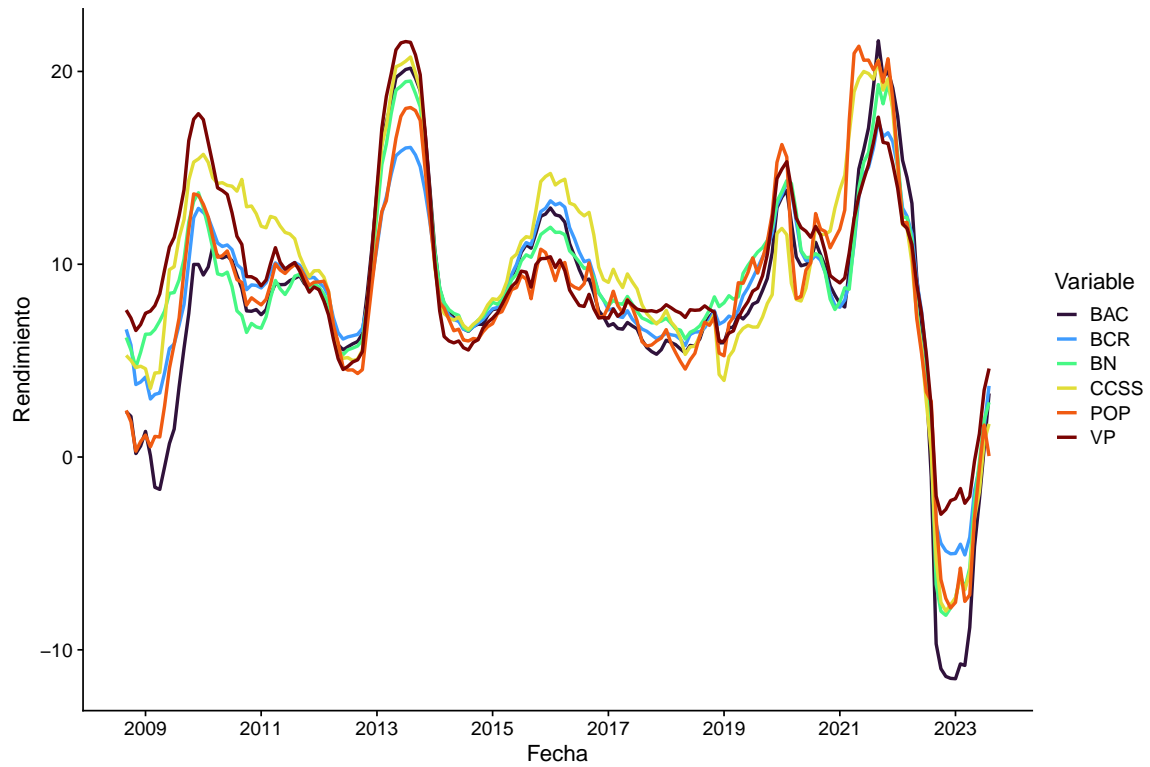
Tabla 18: Estadisticos por fondo

	BAC	BCR	BN	CCSS	POP	VP
Media	8.044	8.646	8.791	9.576	8.436	9.414
Mediana	8.010	8.800	8.505	9.580	8.695	8.785
Mínimo	-11.500	-5.080	-8.210	-7.990	-7.830	-2.970
Máximo	21.590	17.430	19.500	20.750	21.310	21.550
Desviación estándar	6.184	4.401	5.053	5.818	5.558	4.707

Fuente: Elaboracion Propia con datos SUPEN

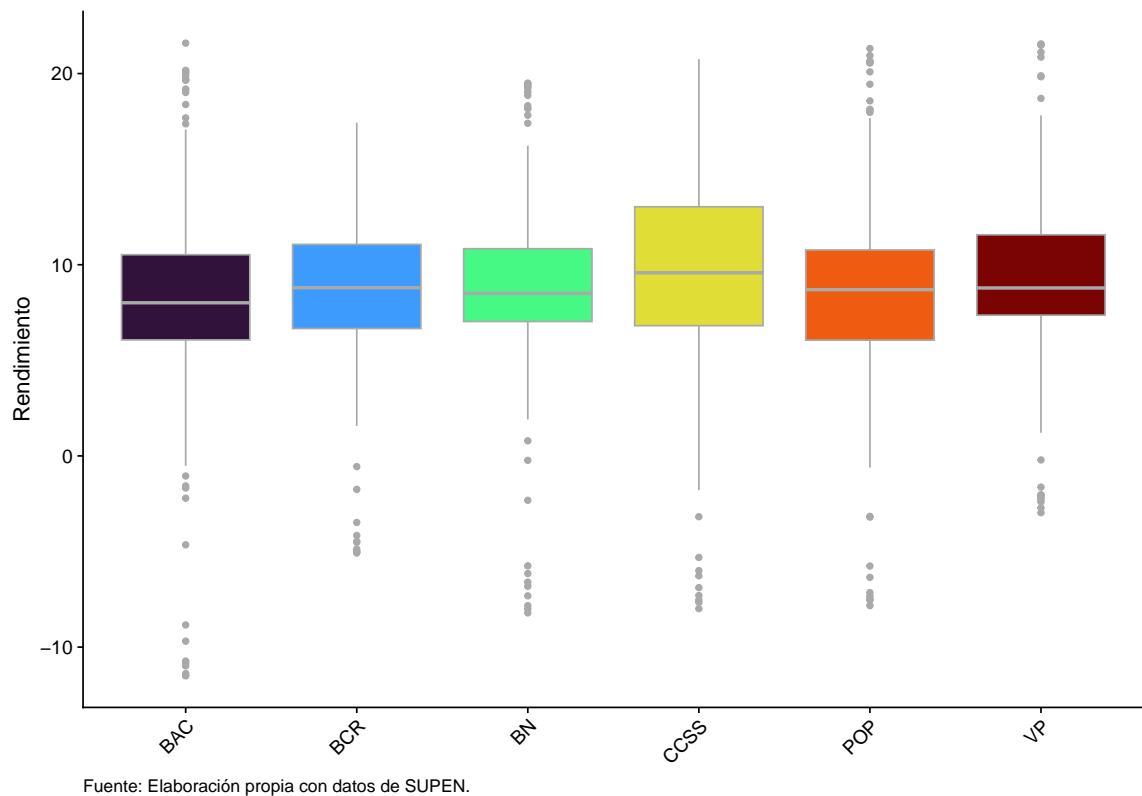
De la tabla 18 se ve que los rendimientos de los fondos presentan estadísticos similares, si graficamos estos vemos que su comportamiento es el siguiente

Gráfico 2: Rendimiento anual de los fondos



Ahora, viendo que su comportamiento medio es similar la siguiente duda a responder es la causa de la leve diferencia evidenciada previamente.

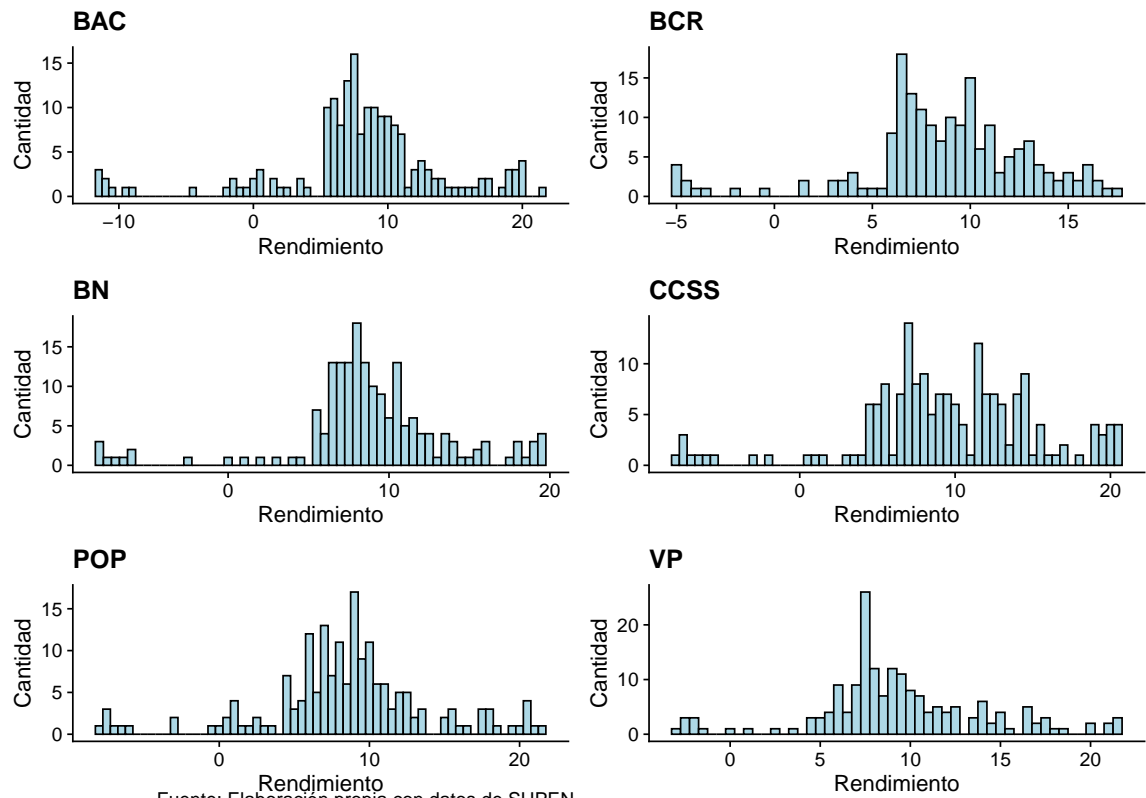
Gráfico 3: Diagrama de caja y bigotes de los fondos



El gráfico 3 señala que esta diferencia puede ser causada por valores extremos presentes en cada variable.

Ahora, si se quiere saber que tan bueno ha sido el rendimiento de estos fondos una opcion viable es analizar su distribucion

Gráfico 4: Distribucion de los fondos

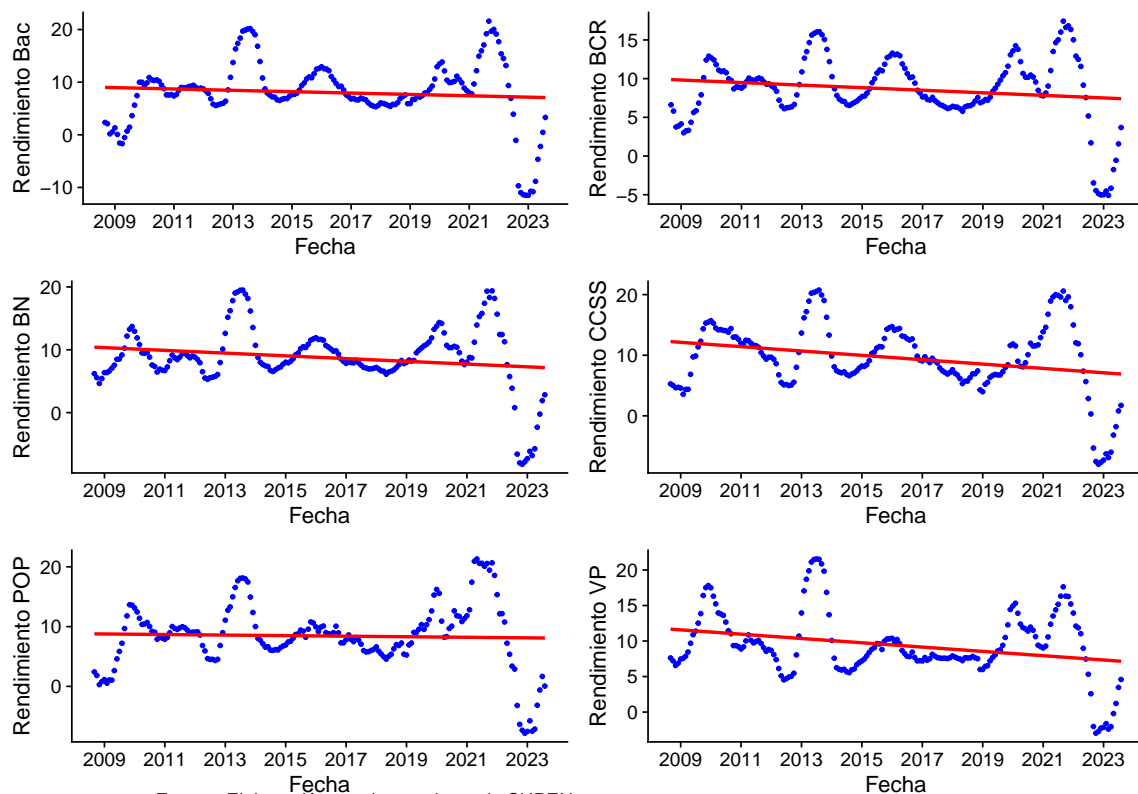


En el grafico 4 vemos que el BAC, el BCR, el BN y la CCSS presentan un mayor peso entre 5% y 15%, luego quedan POP y VP que presentan un mayor peso entre 5% y 10%. Lo que nos dice que, a pesar de la bajada evidenciada en el grafico 2 en total el rendimiento de estos fondos ha sido bueno.

## 5.2. Propuesta metodológica

El primer método que vamos a utilizar es la utilización de R para crear regresiones lineales y poder tener una idea para una proyección de los rendimientos guiandonos con la recta de mejor ajuste, para poder realizarla, se utilizará el método `lm()` de R, el cuál funciona para ajustar modelos de regresión lineal.

Gráfico 5: Regresión Lineal de los fondos



Fuente: Elaboración propia con datos de SUPEN.

El método funciona utilizando una fórmula que relaciona dos variables (la dependiente y la independiente), tomando la variable independiente como el tiempo y la dependiente como los rendimientos, el método se encarga de buscar una función que relacione ambas variables y obtenida dicha relación, se encarga de buscar un modelo de regresión lineal que se ajuste a los datos, el cuál es guardado y posteriormente utilizado para poder hacer predicciones.

Una vez tengamos un modelo ajustado y realicemos las predicciones para observar cómo se hubieran comportado los rendimientos si no hubiera existido una crisis sanitaria, podemos comparar los datos de dicha predicción con los datos recopilados durante la pandemia para validar si existe una diferencia notoria.

Una vez validada dicha pregunta, podemos iniciar con la prueba de hipótesis, en la cuál definiremos la hipótesis nula  $H_0$ , la hipótesis alternativa  $H_a$ , el nivel de significancia y los niveles de error aceptados.

Para poder ejemplificar mejor este método, se realizará una prueba de hipótesis de Kolmogorov-Smirnov, la cual según DeGroot (2002) se utiliza para probar si un conjunto de datos o muestra se acopla o sigue una distribución en particular (Teórica), esto mediante la definición de una Hipótesis Nula  $H_0$ , tal que se establece que la distribución desconocida  $f(X)$  sigue una distribución  $f^*(x)$  para todas las observaciones y en el caso de la Hipótesis Alternativa  $H_a$ , sigue la negación de la Hipótesis Nula, o sea que la distribución no sigue  $f^*(x)$ , además se establece un estadístico  $D_n$  que representa la máxima diferencia entre la función de distribución acumulada (c.d.f) de la muestra observada y la teórica, si  $D_n$  es un valor pequeño tiende a ser pequeño si  $H_0$  es verdadera y grande si se cumple  $H_1$ . Con respecto a la parte programada, se utiliza la librería `stats`, la cual contiene el `ks.test()`, lo cual ayuda a hacer la prueba de Kolmogorov recibiendo como parámetros vectores con los rendimientos de cada entidad, además se agrupan en un número par de variables para poder determinar si tienen la misma distribución.

Tabla 19: Resultados de la prueba KS para todas las combinaciones de pares de variables (columnas 2 a 7) de rentabilidadAnu.

Fondo 1	Fondo 2	Valor p KS
BAC	CCSS	0
CCSS	POP	0.001
BN	CCSS	0.002
BCR	CCSS	0.005
BAC	VP	0.013
CCSS	VP	0.035
POP	VP	0.047
BN	POP	0.082
BAC	BN	0.136
BAC	BCR	0.172
BCR	POP	0.172
BCR	VP	0.172
BN	VP	0.736
BAC	POP	0.819
BCR	BN	0.89

Como se observa en la tabla 19 los primeros 7 valores p menores a  $\alpha = 5$ , en estos casos, gracias a la prueba *Kolmogorov Smirnov* podemos inferir que, para los pares de fondos respectivos a estos valores, hay diferencias significativas entre las dos distribuciones de dichos fondos. De modo que se concluye que los ultimos 8 pares de fondos no presentan diferencias significativas entre sus distribuciones, en otras palabras, provienen de la misma distribucion.

También se realizó una prueba de Dickey-Fuller, la cual examina la hipótesis nula de un modelo autorregresivo integrado de media móvil (ARIMA) en comparación con la hipótesis de estacionariedad, con el fin de determinar si una serie temporal es estacionaria o no.

La estacionariedad se refiere a una propiedad deseada en la que las propiedades estadísticas de una serie temporal (media, varianza autocorrelación, entre otros) se mantienen constantes a lo largo del tiempo. Es decir, una serie temporal se considera estacionaria cuando las características estadísticas se mantienen constantes en intervalos de tiempo sucesivos.

Se procede entonces a realizar dicha prueba de estacionariedad de los rendimientos, utilizando el método de `adf.test()` de R, el cuál recibe una variable numerica, en este caso, con los rendimientos de una entidad, y al objeto obtenido con este método se le puede extraer el valor p, como se menciona en Mushtaq (2011), la estacionariedad depende del valor p resultante de esta prueba. Como se observa en la tabla 20, el valor p de la prueba dickey fuller es menor a  $\alpha = 0,05$  y por tanto se puede rechazar la hipotesis nula al 95% de confianza. Como la hipotesis alternativa es que los datos se comportan de manera estacional concluimos que este es el caso.

Tabla 20: Resultados de la prueba ADF para las columnas 2 a 7 de rentabilidadAnu.

	Valor p	Estadístico ADF
BAC	0.01	-4.76
BCR	0.01	-5.10
BN	0.01	-4.42
CCSS	0.01	-5.01
POP	0.01	-4.04
VP	0.01	-4.88

Fuente: Elaboracion Propia con datos SUPEN

### 5.3. Fichas de resultados

Resultado	Comportamiento de los rendimientos por fondo
Resumen de una oración	Se visualizo la distribucion para ver regiones con mayor masa.
Principal característica	Noción empirica de los datos
Problemas con el argumento o el tema:	Al ser una inferencia obtenida de los graficos generados puede presentar un grado de sesgo relacionado al error humano.
Resumen en un párrafo:	Se manipularon los datos para poder generar un histograma de cada variable a estudiar de la base cruda. Con base a estos histogramas se hizo un primer intento de inferencia sobre el comportamiento de los rendimientos de cada fondo.

Resultado	Identificacion de distorciones
Resumen de una oracion	Inferencia del efecto causado por valores extremos
Principal característica	Nocion empirica de efectos
Problemas con el argumento o el tema:	Al ser una inferencia obtenida de los graficos generados puede presentar un grado de sesgo relacionado al error humano.
Resumen en un párrafo:	Se generó un grafico de cajas y bigotes por entidad, en base a su comportamiento medio y la diferencia en valores extremos se infirió el efecto causado por estos.

Resultado	Regresión Lineal
Resumen de una oracion	Se visualiza la linea de mejor ajuste y se corrobora si los datos se adaptan a ella
Principal característica	Noción empírica de los datos
Problemas con el argumento o el tema:	Al ser un resultado obtenido de los graficos generados, se puede presentar un grado de sesgo relacionado al error humano.
Resumen en un párrafo:	Se generaron varios graficos en los cuáles se puede observar el comportamiento de los rendimientos por entidad, y la linea de mejor ajuste que se generó mediante regresión lineal para dichos comportamientos, y se puede observar como existen máximos y mínimos en los rendimientos que se alejan demasiado de la linea de mejor ajuste, por lo que los rendimientos no tienen un comportamiento que se ajuste a este modelo, lo que ayuda a concluir que se debe utilizar otro tipo de método para la proyección de los datos..

Resultado	Identificación de estacionariedad
Resumen de una oracion	Se realiza una prueba de estacionariedad sobre los fondos
Principal característica	Descripción del comportamiento de los datos
Problemas con el argumento o el tema:	La interpretación de estos resultados esta sujeta a entender como funcionan las pruebas de hipotesis.
Resumen en un párrafo:	Se aplicó la prueba de Dickey-Fuller Aumentada cuya hipotesis alternativa es que los datos tienen un comportamiento estacional, al tener los valores p de esta prueba todos menores a 0.05 se rechaza la hipotesis nula y se concluye que los datos tienen un comportamiento estacional

Resultado	Identificación de distribuciones
Resumen de una oración	Se realiza una prueba de hipótesis para la distribución de los fondos
Principal característica	Descripción del comportamiento de los datos
Problemas con el argumento o el tema:	La interpretación de estos resultados está sujeta a entender como funcionan las pruebas de hipótesis.
Resumen en un párrafo:	Se aplicó la prueba de Kolmogorov-Smirnov cuya hipótesis alternativa es que los datos provienen de distribuciones diferentes, al tener los valores p menores a 0.05 se rechaza la hipótesis nula y se concluye que los datos provienen de distribuciones diferentes

## Referencias

- Cuadros, G., Tarazona, E., Cárdenas Solís, C., Ramírez Infante, R., y cols. (s.f.). Estadística aplicada 2 (ma145), ciclo 2013-1.
- Dario Sanchez, C. D. (2022). *Efectos económicos de la pandemia covid 19, en los principales fondos de pensiones privados de Colombia* (Tesis Doctoral no publicada). Universidad Antonio Nariño.
- DeGroot, M., M. y Schervish. (2002). *Probability and statistics* (Vol. 2). Pearson Education, Inc.
- Fallas, J. (2012). Prueba de hipótesis. *Recuperado de: <http://www.ucipfg.com/Repositorio/MGAP/MGAP>, 5.*
- Feher, C., y de Bidegain, I. (2020). Los planes de pensiones en la crisis de la covid-19: Impacto y consideraciones en materia de políticas. *Fondo Monetario Internacional*.
- Granados, R. M. (2016). Modelos de regresión lineal múltiple. *Granada, España: Departamento de Economía Aplicada, Universidad de Granada*.
- INE. (2020). *Fondos de pensiones*. Descargado de <https://www.ine.es/DEFIne/es/concepto.htm?c=4245&tm=21&txt=&p=2&n=20> (28/08/2023)
- Mushtaq, R. (2011). Augmented dickey fuller test.
- Preece, M. R. M., y Badillo, C. Z. (2020). Impacto de la crisis económica por covid-19 en el sistema de pensiones mexicano y perspectivas ante el proyecto de su reforma. *Contaduría y administración*, 65(5), 16.
- RAE. (2023). *Definición de la palabra*. Descargado de <https://dle.rae.es/> (28/08/2023)
- Soto Escobar, G. P. (2021). El retiro de fondos de pensiones por el covid-19 y su impacto en la rentabilidad de las afp, Perú-2020.
- SUPEN. (2022). *Pilares de su pensión*. Descargado de <https://www.supen.fi.cr/como-se-compone> (28/09/2023)
- SUPEN. (2023). *Inicio*. Descargado de <https://www.supen.fi.cr/web/guest/inicio>
- Vinuesa, P. (2016). Regresión lineal simple y polinomial: teoría y práctica. *Recuperado de [https://www.ccg.unam.mx/~vinuesa/R4biosciences/docs/Tema9\\_regresion.pdf](https://www.ccg.unam.mx/~vinuesa/R4biosciences/docs/Tema9_regresion.pdf)*.
- WB. (2020). *Global economic prospects* (Inf. Téc.). World Bank.



## A. UVE heurística

### Conceptual

#### Conceptos:

- Rendimiento (de Capital).
- Fondo de Pensiones.
- Pensión.

#### Teorías y principios:

- Impacto económico del Covid-19.
- Pruebas de hipótesis.

### Metodología

#### Afirmaciones y resultados:

- Se concluye que los datos tienen un comportamiento estacional.
- Se identifica que 7 pares de fondos presentan diferencias significativas entre sus distribuciones.

#### Transformación de datos:

- Se eliminaron columnas con una cantidad significativa de faltantes y se extrajo los rendimientos anuales de los fondos para realizar el estudio.

#### Modelo:

- Prueba de Dickey-Fuller Aumentada.
- Prueba de hipótesis Kolmogorov-Smirnov.
- Modelo de regresión lineal.

#### Pregunta central:

¿Cuál es el impacto que tuvo la crisis sanitaria del Covid-19 sobre los rendimientos de los fondos de pensiones obligatorias complementarias Costarricenses?

#### Acontecimiento:

- Se presenta una bajada en los rendimientos de los fondos de pensiones obligatorias complementarias luego del 2021.

## A.1. Código

A continuación se incluye el código utilizado para realizar los cuadros y gráficos presentados en la bitácora 2:  
**Limpieza de los datos**

```
1 rentabilidadxENTIDADES <- read.csv("C:/2023/Ciclo II/Estadistica1/Proyecto
  Estadistica/RentabilidadxENTIDADES.csv", skip = 5, dec = ",") %>% remove_empty()
2
3 colnames(rentabilidadxENTIDADES) <- c("Mes", "histBAC", "anuBAC", "histBCR", "anuBCR",
  "histBN", "anuBN", "histCCSS", "anuCCSS", "histIBP", "anuIBP", "histINS", "
  anuINS", "histPOP", "anuPOP", "histVP", "anuVP")
4
5 rentRaw <- rentabilidadxENTIDADES
6
7 rentabilidadxENTIDADES <- rentabilidadxENTIDADES %>% clean_names() %>% select(-c(hist
  _ibp, anu_ibp, hist_ins, anu_ins))
8
9 nombres_meses <- c(
10   "ene", "feb", "mar", "abr", "may", "jun",
11   "jul", "ago", "sep", "oct", "nov", "dic"
12 )
13
14 rentabilidadxENTIDADES <- rentabilidadxENTIDADES %>% mutate(year = as.numeric(paste0(
  "20", substr(mes, 5, 6))), month = match(substr(mes, 1, 3), nombres_meses))
15
16 rentabilidadxENTIDADES <- rentabilidadxENTIDADES %>%
17   mutate(mes = my(paste0(month, "-", year))) %>%
18   select(-month, -year)
19
20 rentabilidadHist <- rentabilidadxENTIDADES %>%
21   select(matches("hist"))
22 rentabilidadHist <- cbind(rentabilidadxENTIDADES[, 1], rentabilidadHist)
23
24 rentabilidadAnu <- rentabilidadxENTIDADES %>%
25   select(matches("anu"))
26 rentabilidadAnu <- cbind(mes = rentabilidadxENTIDADES$mes, rentabilidadAnu)
```

**Cuadro 18**

```
1 tabla<-function(data, variable){
2   mean_temp <- mean(data) %>% round(3)
3   median_temp <- median(data) %>% round(3)
4   min_temp <- min(data) %>% round(3)
5   max_temp <- max(data) %>% round(3)
6   std_dev_temp <- sd(data) %>% round(3)
7   summary_df <- data.frame(
8     Statistic = c("Media", "Mediana", "Mínimo", "Máximo", "Desviación estándar"),
9     Value = c(mean_temp, median_temp, min_temp, max_temp, std_dev_temp)
10  )
11  colnames(summary_df) <- c("Estadístico", variable)
12  return(summary_df)
13 }
14
15 estBAC <- rentabilidadxENTIDADES$anu_bac %>% tabla("BAC")
16 estBCR <- rentabilidadxENTIDADES$anu_bcr %>% tabla("BCR")
17 estBN <- rentabilidadxENTIDADES$anu_bn %>% tabla("BN")
18 estCCSS <- rentabilidadxENTIDADES$anu_ccss %>% tabla("CCSS")
19 estPOP <- rentabilidadxENTIDADES$anu_pop %>% tabla("POP")
20 estVP <- rentabilidadxENTIDADES$anu_vp %>% tabla("VP")
21 estadisticos <- cbind(estBAC, BCR = estBCR[, 2], BN = estBN[, 2], CCSS = estCCSS[, 2],
  POP = estPOP[, 2], VP = estVP[, 2]) %>% t()
22 colnames(estadisticos) <- as.character(estadisticos[1, ])
23 estadisticos <- estadisticos[-1, ] %>% t()
24
25
26 estTex <- xtable(estadisticos, digits = rep(3, 7), align = rep("l", 7), caption = "
  Fuente: Elaboración propia con datos de SUPEN.")
```

```

27
28 print(estTex, include.rownames = T)

```

### Cuadro 17

```

1 # Seleccionar los fondos
2 subset_data <- rentabilidadAnu[, 2:7]
3 colnames(subset_data) <- toupper(substring(colnames(subset_data), first = 5))
4 # Calcular la matriz de correlación
5 correlacion <- cor(subset_data)
6
7 # Crear una tabla de correlación
8 tabCorrelacion <- as.data.frame(correlacion)
9
10 # Mostrar la tabla de correlación
11 print(tabCorrelacion)
12 tabCorrelacion %>% xtable(digits = rep(3,7), align = rep("l",7)) %>% print(include.
    rownames = T)

```

### Gráficos 2 y 3

```

1 # Convertir el dataframe a formato largo
2 rentAnuLong <- rentabilidadAnu %>% pivot_longer(cols = -mes, names_to = "Variable",
    values_to = "Rendimiento")
3 rentAnuLong <- rentAnuLong %>% mutate(Variable = toupper(substring(Variable, first =
    5)))
4
5 # GRAFICO DE LINEA
6 # Crear el gráfico de líneas
7 (plotLinRend <- rentAnuLong %>% ggplot(aes(x = mes, y = Rendimiento, color = Variable
    )) +
8   geom_line(size = 1) +
9   # labs(title = "Gráfico de Líneas de Rendimientos", x = "Fecha", y = "Rendimiento")
10  +
11  labs(x = "Fecha", y = "Rendimiento", caption = "Fuente: Elaboración propia con
    datos de SUPEN.") +
12  theme(plot.caption = element_text(hjust = 0)) +
13  theme(legend.position = "top") +
14  # scale_color_brewer(palette = "Dark2") +
15  scale_x_date(date_breaks = "2 year", date_labels = "%Y") +
16  scale_color_viridis(discrete = TRUE, option = "H") +
17  scale_fill_viridis(discrete = TRUE) +
18  theme(legend.position = "right")
19  )
20 pdf("linRend.pdf", width = 10)
21 plotLinRend
22 dev.off()
23
24 # GRAFICO DE DISPERSION
25 ( plotDispRend <- rentAnuLong %>% ggplot(aes(x = mes, y = Rendimiento, color =
    Variable)) +
26   geom_point(size = 1, alpha = 0.6) +
27   # labs(title = "Gráfico de Dispersión de Rendimientos", x = "Mes", y = "Rendimiento
    ") +
28   labs(x = "Mes", y = "Rendimiento", caption = "Fuente: Elaboración propia con datos
    de SUPEN.") +
29   theme(plot.caption = element_text(hjust = 0)) +
30   theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
31   scale_color_brewer(palette = "Dark2") +
32   scale_x_date(date_breaks = "2 year", date_labels = "%Y") +
33   scale_color_viridis(discrete = TRUE, option = "H") +
34   scale_fill_viridis(discrete = TRUE) +
35   theme(legend.position = "right")
36   )
37
38 pdf("dispRend.pdf", width = 10)

```

```

39 plotDispRend
40 dev.off()
41
42 # GRAFICO DE CAJAS
43 (plotBoxRend <- rentAnuLong %>% ggplot(aes(x = Variable, y = Rendimiento, fill =
  Variable)) +
44   geom_boxplot(color = "darkgrey", show.legend = F)+#(outlier.shape = NA) +
45   # labs(title = "Gráfico de Caja de Rendimientos", x = "Variable", y = "Rendimiento
    ") +
46   labs(x = NULL, y = "Rendimiento", caption = "Fuente: Elaboración propia con datos
    de SUPEN.")+
47   theme(plot.caption = element_text(hjust = 0)) +
48   theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
49   # scale_color_viridis(discrete = TRUE, option = "H") +
50   scale_fill_viridis(discrete = TRUE, option = "H") +
51   theme(legend.position = "right")
52 )
53
54 pdf("boxRend.pdf", width = 10)
55 plotBoxRend
56 dev.off()

```

**Gráfico 4**

```

1 data <- rentabilidadAnu[,-1]
2
3 hist_list_ggplot <- lapply(names(data), function(col) {
4   ggplot(data, aes_string(x = col, fill = col)) +
5     geom_histogram(binwidth = 0.5, fill = "lightblue", color = "black") +
6     labs(title = toupper(substring(col, first = 5)), x = "Rendimiento", y = "Cantidad
    ")
7 })
8
9 histArray <- grid.arrange(grobs = hist_list_ggplot, ncol = 2)
10 caption_text <- "Fuente: Elaboración propia con datos de SUPEN."
11 grid.text(
12   x = 0.5,
13   y = 0.01,
14   label = caption_text,
15   just = "right"
16 )
17 histArray

```

**Gráfico de autocorrelacion**

```

1 rentabilidadAnu %>% acf()

```

**Cuadro 16**

```

1 contadorND <- sapply(rentRaw, function(column) sum(column == "ND"))
2
3 # Filtrar las columnas con NA > 0
4 faltantes <- contadorND[contadorND > 0]
5
6 # Crear una tabla que muestra la cantidad de datos faltantes por columna (solo las
  que tienen NA > 0)
7 tablaND<- data.frame(
8   # Columna = names(filtered_missing),
9   Fondo = substr(names(faltantes), nchar(names(faltantes)) - 2, nchar(names(faltantes)
    )),
10  "Datos Faltantes" = faltantes
11 )
12
13 # Mostrar la tabla de datos faltantes por columna
14 # print(tablaND)
15 tablaND %>% xtable(caption = "Cantidad de datos faltantes por fondo") %>% print(
  include.rownames = F)

```

Cuadro 20

```

1 data <- rentabilidadAnu
2 resultados <- c()
3 # Realizar la prueba ADF para cada fondo
4 for (i in 2:7) {
5   # Realizar la prueba ADF
6   adf_result <- adf.test(data[, i])
7   # Extraer el valor p y el estadístico de la prueba
8   p_value <- adf_result$p.value
9   test_statistic <- adf_result$statistic
10  # Almacenar los resultados
11  resultados <- c(resultados, c(p_value, test_statistic))
12 }
13 # Crear una tabla LaTeX con los resultados
14 tabla_latex <- matrix(resultados, ncol = 2, byrow = TRUE)
15 colnames(tabla_latex) <- c("Valor p", "Estadístico ADF")
16 rownames(tabla_latex) <- toupper(substring(colnames(data)[2:7], first = 5))
17 # Imprimir la tabla LaTeX
18 print(xtable(tabla_latex, caption = "Resultados de la prueba ADF para las columnas 2
    a 7 de rentabilidadAnu."),
19       caption.placement = "top",
20       include.rownames = TRUE,
21       include.colnames = TRUE,
22       table.placement = "h")

```

Cuadro 19

```

1 data <- rentabilidadAnu[, 2:7]
2 # Crear una matriz para almacenar los resultados
3 num_var <- ncol(data)
4 resultados <- matrix(NA, ncol = 3, nrow = num_var * (num_var - 1) / 2)
5 colnames(resultados) <- c("Fondo 1", "Fondo 2", "Valor p KS")
6 # Realizar la prueba KS para todos los pares de variables
7 row_idx <- 1
8 for (i in 1:(num_var - 1)) {
9   for (j in (i + 1):num_var) {
10     var1 <- data[, i]
11     var2 <- data[, j]
12     # Realizar la prueba KS
13     ks_result <- ks.test(var1, var2)
14     # Almacenar los resultados en la matriz
15     resultados[row_idx, 1] <- toupper(substring(colnames(data)[i], first = 5))
16     resultados[row_idx, 2] <- toupper(substring(colnames(data)[j], first = 5))
17     resultados[row_idx, 3] <- round(ks_result$p.value, 3)
18     row_idx <- row_idx + 1
19   }
20 }
21 # Ordenar los resultados por valor p de menor a mayor
22 resultados <- resultados[order(resultados[, 3]), ]
23 # Imprimir la tabla LaTeX
24 print(xtable(resultados, caption = "Resultados de la prueba KS para todas las
    combinaciones de pares de variables (columnas 2 a 7) de rentabilidadAnu."),
25       digits = rep(3, 3),
26       caption.placement = "top",
27       include.rownames = FALSE,
28       include.colnames = TRUE,
29       table.placement = "h")

```

Gráfico 5

```

1 # Ajustar un modelo de regresión lineal
2 modelo <- lm(anu_bac ~ numero_mes, data = rentabilidadAnuNorm)
3 # Resumen del modelo
4 summary(modelo)
5
6 # Crear un gráfico de dispersión y de regresión con ggplot2
7 (lmBac <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_bac)) +

```

```

8   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
9   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
10  labs(title = NULL, x = "Fecha", y = "Rendimiento Bac")+#, caption = "Fuente:
    Elaboración propia con datos de SUPEN.")+
11  theme(plot.caption = element_text(hjust = 0))+
12  scale_x_date(date_breaks = "2 year", date_labels = "%Y")
13 )
14 pdf("lmBac.pdf", width = 10)
15 lmBac
16 dev.off()
17
18 # Ajustar un modelo de regresión lineal
19 modelo <- lm(anu_bcr ~ numero_mes, data = rentabilidadAnuNorm)
20 # Resumen del modelo
21 summary(modelo)
22
23 # Crear un gráfico de dispersión y de regresión con ggplot2
24 (lmBCR <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_bcr)) +
25   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
26   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
27   labs(title = NULL, x = "Fecha", y = "Rendimiento BCR")+#, caption = "Fuente:
    Elaboración propia con datos de SUPEN.")+
28   theme(plot.caption = element_text(hjust = 0))+
29   scale_x_date(date_breaks = "2 year", date_labels = "%Y")
30 )
31
32 pdf("lmBCR.pdf", width = 10)
33 lmBCR
34 dev.off()
35
36 # Ajustar un modelo de regresión lineal
37 modelo <- lm(anu_bn ~ numero_mes, data = rentabilidadAnuNorm)
38 # Resumen del modelo
39 summary(modelo)
40
41 # Crear un gráfico de dispersión y de regresión con ggplot2
42 (lmBN <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_bn)) +
43   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
44   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
45   labs(title = NULL, x = "Fecha", y = "Rendimiento BN")+#, caption = "Fuente:
    Elaboración propia con datos de SUPEN.")+
46   theme(plot.caption = element_text(hjust = 0))+
47   scale_x_date(date_breaks = "2 year", date_labels = "%Y")
48 )
49 pdf("lmBN.pdf", width = 10)
50 lmBN
51 dev.off()
52
53 # Ajustar un modelo de regresión lineal
54 modelo <- lm(anu_ccss ~ numero_mes, data = rentabilidadAnuNorm)
55 # Resumen del modelo
56 summary(modelo)
57
58 # Crear un gráfico de dispersión y de regresión con ggplot2
59 (lmCCSS <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_ccss)) +
60   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
61   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
62   labs(title = NULL, x = "Fecha", y = "Rendimiento CCSS")+#, caption = "Fuente:
    Elaboración propia con datos de SUPEN.")+
63   theme(plot.caption = element_text(hjust = 0))+
64   scale_x_date(date_breaks = "2 year", date_labels = "%Y")
65 )
66 pdf("lmCCSS.pdf", width = 10)
67 lmCCSS
68 dev.off()
69

```

```

70 # Ajustar un modelo de regresión lineal
71 modelo <- lm(anu_pop ~ numero_mes, data = rentabilidadAnuNorm)
72 # Resumen del modelo
73 summary(modelo)
74
75 # Crear un gráfico de dispersión y de regresión con ggplot2
76 (lmPOP <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_pop)) +
77   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
78   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
79   labs(title = NULL, x = "Fecha", y = "Rendimiento POP")+#, caption = "Fuente:
      Elaboración propia con datos de SUPEN.")+
80   theme(plot.caption = element_text(hjust = 0))+
81   scale_x_date(date_breaks = "2 year", date_labels = "%Y")
82 )
83 pdf("lmPOP.pdf", width = 10)
84 lmPOP
85 dev.off()
86
87 # Ajustar un modelo de regresión lineal
88 modelo <- lm(anu_vp ~ numero_mes, data = rentabilidadAnuNorm)
89 # Resumen del modelo
90 summary(modelo)
91
92 # Crear un gráfico de dispersión y de regresión con ggplot2
93 (lmVP <- ggplot(rentabilidadAnuNorm, aes(x = mes, y = anu_vp)) +
94   geom_point(color = "blue", size = 1) + # Puntos azules para los datos
95   geom_smooth(method = "lm", color = "red", se = FALSE) + # Línea de regresión
96   labs(title = NULL, x = "Fecha", y = "Rendimiento VP")+#, caption = "Fuente:
      Elaboración propia con datos de SUPEN.")+
97   theme(plot.caption = element_text(hjust = 0))+
98   scale_x_date(date_breaks = "2 year", date_labels = "%Y")
99 )
100 pdf("lmVP.pdf", width = 10)
101 lmVP
102 dev.off()
103
104 lmList <- list(lmBac, lmBCR, lmBN, lmCCSS, lmPOP, lmVP)
105
106 grid.arrange(grobs = lmList, ncol = 2)
107 caption_text <- "Fuente: Elaboración propia con datos de SUPEN."
108 grid.text(
109   x = 0.5, # Adjust the x-coordinate to center the caption
110   y = 0.01, # Adjust the y-coordinate to set the position of the caption
111   label = caption_text,
112   just = "right"
113 )
114
115
116 pdf("regresiones.pdf", width = 10)
117 grid.arrange(grobs = lmList, ncol = 2)
118 caption_text <- "Fuente: Elaboración propia con datos de SUPEN."
119 grid.text(
120   x = 0.5, # Adjust the x-coordinate to center the caption
121   y = 0.01, # Adjust the y-coordinate to set the position of the caption
122   label = caption_text,
123   just = "right"
124 )
125 dev.off

```