

# **Estimación de retiros netos para la administración de recursos de terceros**

**Juan Esteban Calderón Gómez**

**Reporte**

Análisis inicial de datos



Universidad Nacional De Colombia

## Contenido

I.	Alistamiento de los datos (Meta data set up) .....	3
II.	Diagrama de flujo (Flow chart of the study).....	3
III.	Procesamiento y limpieza de Datos (data processing and cleaning) .....	4
IV.	Data screening.....	4
V.	Afectaciones en la interpretación de resultados .....	6
VI.	Afectaciones en el análisis estadístico .....	6

El presente reporte tiene como finalidad presentar los resultados exploratorios del análisis inicial de los datos obtenidos, así como la descripción del proceso de obtención de estos y su respectiva limpieza y transformación.

## I. Alistamiento de los datos (Meta data set up)

El presente proyecto en su concepción está enmarcado en el propósito de establecer una técnica de analítica predictiva para series de tiempo, que permita calcular de forma diaria el nivel de retiro neto probable a 1 día de los diferentes portafolios que componen el Fondo Voluntario de Pensiones de Fiduciaria Davivienda S.A.

En línea con lo anterior el estudio aborda tres frentes que apalancan la definición de alcanzables específicos, el primero de ellos se relaciona con la búsqueda de una implementación computacional que sea práctica, y que requiera la menor interacción posible por parte de un usuario. El segundo frente se enmarca en el uso de información disponible públicamente, aunque un desarrollo a futuro puede estar enmarcado en el uso de datos internos de la organización. El último pilar se enfoca en la forma de explicar el modelo como tal, ya que a pesar de que en general la audiencia es especializada en temas de inversiones, no necesariamente lo es en temas de analítica.

Dado lo anterior la necesidad de establecer una predicción radica en la importancia de mitigar el riesgo de liquidez al administrar portafolios; un factor que ayuda a materializar dicho riesgo se da en el momento en el que un cliente quiere

sacar su dinero de los portafolios y no se encuentran de forma líquida (cuentas de ahorro) para entregárselos.

Buscando responder a los planteamientos presentados previamente, este proyecto recurre en principio a información que se encuentra disponible de forma pública en el sitio web <https://www.datos.gov.co/dataset/Fondos-de-Pensiones-Voluntarias/gpzw-wmxz> en el cual la Superintendencia Financiera de Colombia publica los datos que las entidades que administran fondos voluntarios de pensiones tienen la obligación de reportarle a dicho ente, entre estos datos como el valor de portafolio, valor de unidad para realizar operaciones de ingreso y salida de clientes y valores de retiros de recursos de los diferentes fondos.

## II. Diagrama de flujo (Flow chart of the study)

A continuación, se presenta el diagrama 1, el cual permite evidenciar la población objetivo del estudio, así mismo se muestra una cascada donde se resalta el proceso a grandes rasgos de la eliminación de registros de la base, bajo diferentes condiciones establecidas:

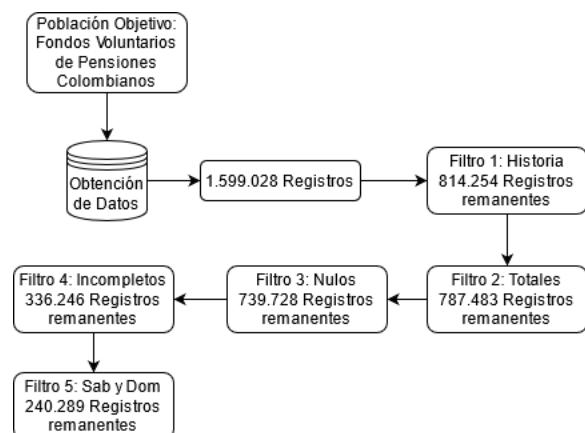


Diagrama 1: Población objetivo y registros remanentes.

### III. Procesamiento y limpieza de Datos (data processing and cleaning)

Como se mencionó previamente el proceso de obtención de datos se realiza a través de un API (Application programming interface) que se encuentra disponible públicamente, en términos generales este repositorio cuenta con información estructurada y que contiene una organización que es consistente en el tiempo, en principio la base contiene información desde el año 2014, sin embargo, como se enuncia a continuación, se aplicaron diversos criterios de selección:

1. **Historia:** Dada una limitante de procesamiento computacional, el periodo máximo de información que fue posible cargar fueron 5 años, los cuales siempre se calculan desde el mismo día en que se realiza la ejecución del script de Python.
2. **Totales:** la base contiene para cada día y cada compañía que administra los portafolios, un agrupador que refleja todos los demás, por ejemplo, si una empresa en particular tiene los fondos  $f_1$  y  $f_2$  existe un  $F$  tal que  $f_1 \subset F$  y  $f_2 \subset F$ . Por ello todos los  $F$  son excluidos ya que el interés del estudio recae sobre los fondos en particular y no sobre el total.
3. **Nulos:** aquellas entidades cuyo número de faltantes sea mayor al 1% en un periodo de tiempo diario de la venta de tiempo cargada (5 años) son desestimadas del análisis.

4. **Incompletos:** de la masa de fondos resultantes, todos aquellos que no se encuentren dentro de la totalidad de días de la ventana de tiempo, con un margen de tolerancia sobre el numero máximo de ocurrencias de la compañía, son igualmente excluidos.

Vale la pena aclarar que en este punto se incluyen **todas** las entidades que ofrecen fondos voluntarios de pensiones dados los filtros presentados anteriormente, sin perjuicio de que dentro del desarrollo del modelo se termine decidiendo trabajar únicamente con los datos de Fiduciaria Davivienda.

### IV. Data screening

En Colombia existen alrededor de 14 entidades que ofrecen fondos voluntarios de pensiones, que para el mes de Junio de 2020, administran alrededor de 29 Billones de pesos, lo cual equivale al cerca del 10% del saldo<sup>1</sup> que se ahorra en el país de forma obligatoria de acuerdo a lo establecido por la regulación.

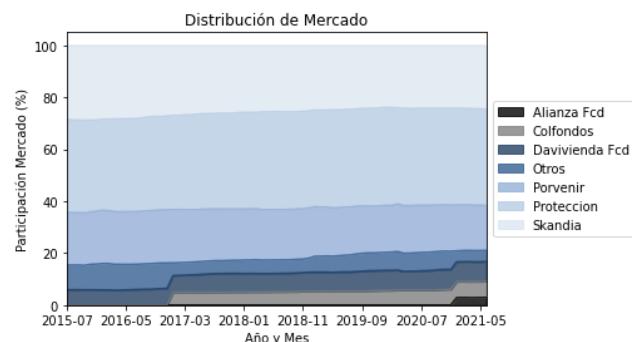


Gráfico 1: distribución de los activos Administrados

En ese sentido Fiduciaria Davivienda administra alrededor del 7.5% de los activos de la industria y tal como se

<sup>1</sup> Obtenido de [www.superfinanciera.gov.co/jsp/9125](http://www.superfinanciera.gov.co/jsp/9125)

## Estimación de retiros netos para la administración de recursos de terceros

aprecia en el gráfico 1 en general ha venido manteniendo su participación a lo largo del tiempo, resulta llamativo como a pesar de que se presentan crisis en los mercados financieros (ejemplo Marzo 2020), la distribución entre las diferentes compañías continua siendo homogénea, donde Protección es el líder con una participación de mercado del 37% aproximadamente.

En el gráfico 2 se presenta la evolución de los activos administrados de la industria frente a Fiduciaria Davivienda, esto permite confirmar en cierta medida lo comentado previamente, y es que conforme pasa el tiempo, se observan movimientos significativamente similares lo que puede aportar a la idea de que se está frente a un tipo de negocio que pareciera comportarse de forma homogénea.

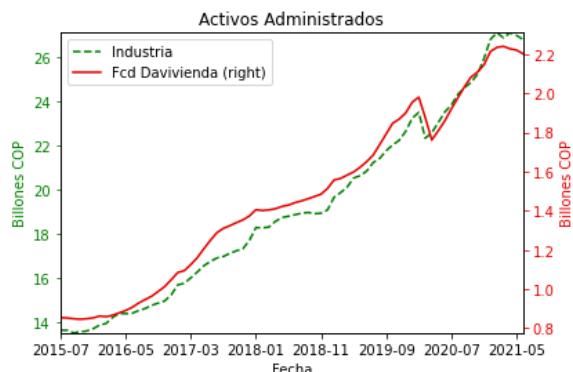


Gráfico 2: evolución de los activos Administrados

Los activos que administra una entidad financiera (como las mencionadas anteriormente) son propiedad de los clientes no de la compañía, por tanto un cliente está en facultad de decidir retirar sus recursos en cualquier momento, salvo ciertas condiciones especiales.

En ese sentido una primera aproximación al entendimiento de la dinámica de los retiros netos de los clientes (aportes menos retiros) se puede ver en el gráfico 3 donde para valores negativos se

observan movimientos especialmente altos en los meses de Marzo y Diciembre, en el caso del primero se da por un efecto especial de la pandemia del Covid-19 la cual generó altos volúmenes de retiros por parte de los clientes dado el pánico generado en los mercados financieros, en el caso del segundo resulta curioso al no ser un mes que represente normalmente choques fuertes en la rentabilidad de los portafolios.

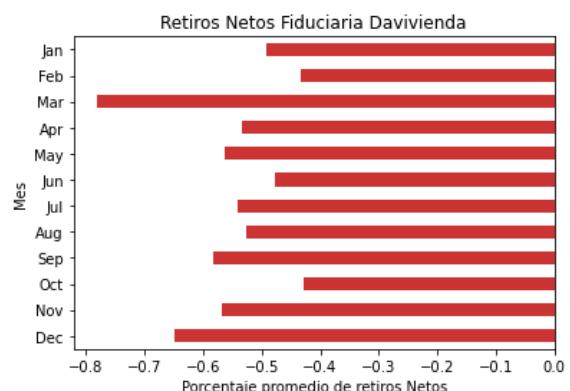


Gráfico 3: promedios de retiros netos por mes Total Fiduciaria

Cuando se analiza la dinámica de retiros netos cuyo valor es negativo, se puede apreciar como hay una concentración importante alrededor de 0, con un grupo de eventos que podrían considerarse de alto impacto y baja frecuencia que se encuentran en la cola de la distribución de los datos, esto es posible apreciarlo en el gráfico 4

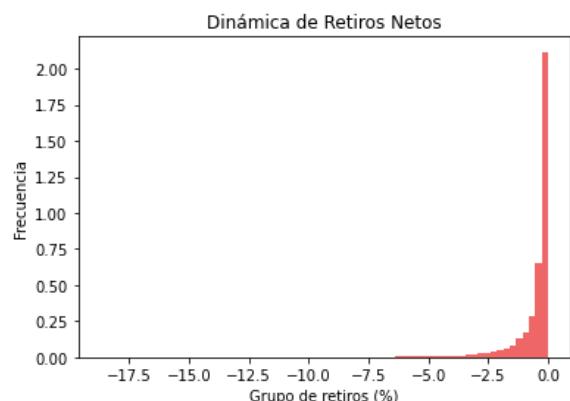


Gráfico 4: distribución de retiros netos negativos

Resulta interesante ver que esa distribución concertada en cero es muy similar para todos los años, en el gráfico 5 es posible evidenciar, que si bien es cierto, para años como el 2020 hay eventos con severidades más altas, en general entre los diferentes períodos de tiempo hay un comportamiento similar entre si.

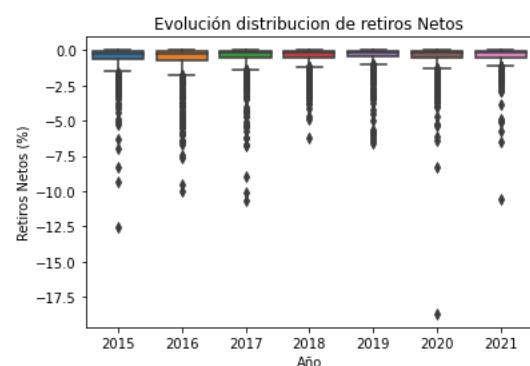


Gráfico 5: evolución de retiros netos negativos

En línea con lo anterior es pertinente pensar si cuando un cliente realiza un retiro de alguno de los fondos administrados, lo hace porque ha observado rendimientos negativos en su inversión, sin embargo en la tabla 1 es posible ver que, para la mayoría de fondos administrados por Fiduciaria Davivienda que cuentan con todas las observaciones en el periodo de estudio, no pareciera existir una relación lineal entre el porcentaje de retiro neto y el porcentaje de retorno del fondo diario.

Portafolio	Coef Correlación <i>p</i>
Renta Fija Pesos	42.14%
Estable	21.40%
PCP	4.93%
Dinámico	2.70%
Moderado	4.05%
Acciones Global	2.14%
Acciones Colombia	-2.60%
Renta Fija Global	-3.37%

Tabla 1 : Coeficiente de correlación de Pearson entre retorno diario y valor de retiro

## V. Afectaciones en la interpretación de resultados

En General los datos provienen de una fuente estructurada que en el tiempo es bastante estable, sin embargo hay algunas particularidades que podrían eventualmente afectar la interpretación de los resultados observados, tal como los que se enuncian a continuación:

- 1. Datos Faltantes:** Es posible evidenciar que para ciertos días no se presentan datos en diferentes compañías, lo cual haría que la serie de tiempo se encuentre incompleta
- 2. Periodo de experiencia:** dependiendo del número de años de experiencia elegido, la muestra se reduce considerablemente, ya que al no tener en cuenta fondos que estén durante todo el periodo, se pierde información de diferentes compañías y portafolios.

## VI. Afectaciones en el análisis estadístico

Existen algunos hallazgos preliminares que permiten esperar que a futuro se presenten algunas dificultades en el análisis estadístico, algunas de ellas son:

- 1. Sesgo en los retiros:** la distribución de los retiros netos muestra un claro sesgo hacia valores cercanos a cero, esto a futuro podría afectar el desarrollo de modelos, ya que una alta concentración en valores pequeños podría afectar la estimación de las severidades altas los cuales afectan el

- portafolio de manera más importante.
- 2. No relación clara frente a los retornos:** resultando un poco contra intuitivo desde una perspectiva de negocio, que los datos muestren que no existe una relación entre los niveles de retiros y el retorno de los portafolios, podría ser de esperarse que cuando un cliente ve una rentabilidad que no es favorable, realice el retiro de sus recursos
- 3. Homogeneidad en la industria:** aparentemente el análisis preliminar indica que la industria en general se comporta de una forma muy similar, esto podría eventualmente afectar la interpretación de los resultados que se obtengan de incluir estos datos dentro del análisis.