# 19/03/2023

Proyecto Final

Data Analytics - Coderhouse



Profesor: Darío Ceballos Tutor: Martín Felipe Segovia Álvarez

Comisión: 32645 Alumna: Marianela Miguel

# Introducción al proyecto

# La empresa Financial Company ha dispuesto la presentación de un reporte de su situación financiera en los años 2019, 2020 y 2021. Pretende que el enfoque este dado en su núcleo fuerte de inversores formado por 300 personas de 10 países diferentes de Latinoamérica. Teniendo este horizonte se realizó un Dashboard en Power Bi para el área de Análisis Financiero del Departamento de Relaciones Comerciales de la empresa. En el mismo se muestran las inversiones de los inversionistas asociados tanto en los distintos sectores como en los diferentes años de análisis. Además, se evalúa cuan diferente fue la expectativa de inversión en 2021 respecto a la inversión real y cuantos proyectos nuevos fueron presentados para llevarse a cabo en 2022.

# Dataset escogido

# La base de datos elegida para este proyecto presenta información referente a diversos proyectos de inversión de Financial Company, empresa financiera especializada en relaciones comerciales, durante los años 2019, 2020 y 2021. Los principales sectores de inversión son: Tecnología, Educación, Construcción, Agroindustrial e Inmobiliaria y los Inversionistas pertenecen a 10 países de Latinoamérica: Colombia, Venezuela, Argentina, Perú, Ecuador, Bolivia, Chile, Uruguay, Paraguay y Brasil. El Dataset original tenía un total de 300 filas y 13 columnas que luego fueron redistribuidas en 5 tablas.

# Dataset original

Objetivo del proyecto:

El objetivo del proyecto es analizar los principales sectores en los que invierte una empresa financiera mediante sus asociados localizados en 8 países de Latinoamérica entre 2019 y 2021. Además, interesa analizar cuáles son los detalles de estas inversiones y detectar correlaciones de estos datos con el desarrollo de la economía de cada país. Por último, interesa realizar proyecciones para 2022 con el objetivo de conocer que áreas serán las de mayor inversión, teniendo en cuenta quienes han presentado proyectos.

Alcance:

El alcance de este proyecto es el Área de Análisis Financiero del Departamento de Relaciones Comerciales de la empresa.

Usuario final y nivel de aplicación del análisis:

El usuario final es el Director de Relaciones Comerciales de la empresa.

El nivel de aplicación del análisis es operativo ya que es un procedimiento que la empresa realiza todos los años tendientes a preparar los planes estratégicos financieros, para sí misma y para sus inversores. Es también estratégico ya que el análisis de estos datos permitirá tomar importantes decisiones de índole financiero y comercial.

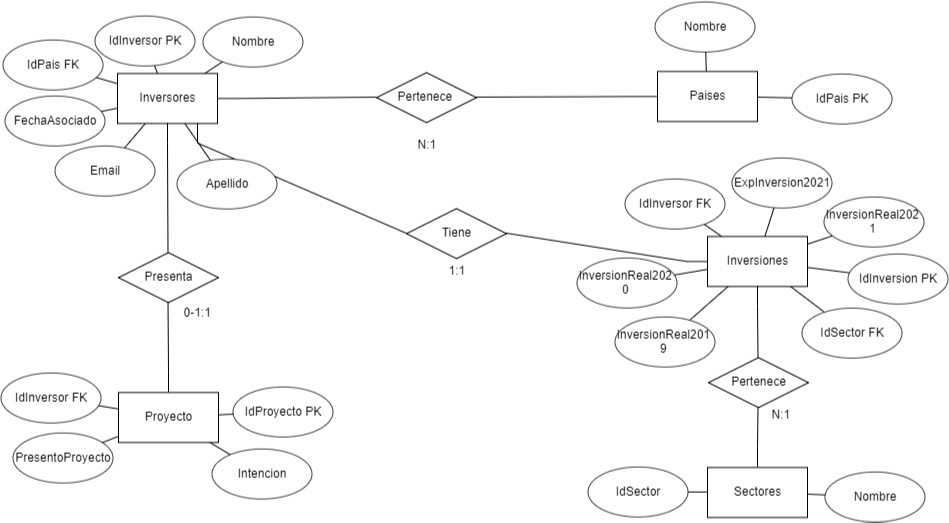
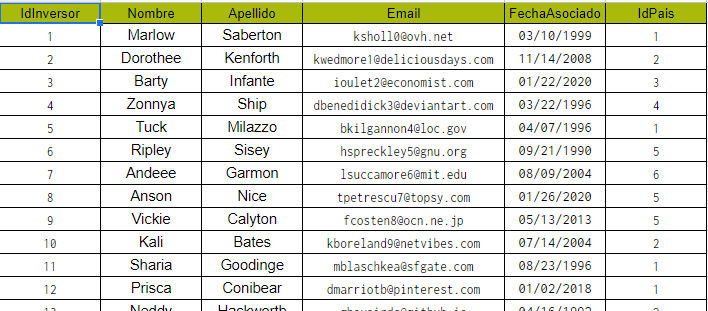


Diagrama entidad-relación

Listado de tablas:

-**Inversores**: en esta tabla se encuentran los datos personales y de contacto de los asociados inversores. Las columnas con su tipo de dato son: IdInversor (int PK), nombre (text), apellido (text), email (text), fechaAsociado (datetime), idPais (int FK).



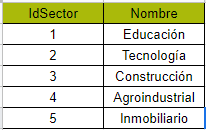
-

**Países**: en esta tabla se encuentra el listado de países representados por los asociados. Las columnas con su tipo de dato son: idPais (int PK) y nombre (text).

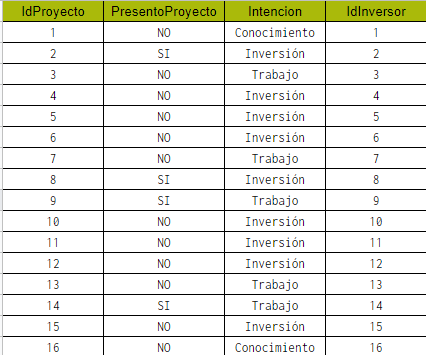


-**Sectores**: en esta tabla se encuentra el listado de los sectores comerciales representados por los diferentes inversores. Las columnas con su tipo de dato son: idSector (int PK), nombre (text).

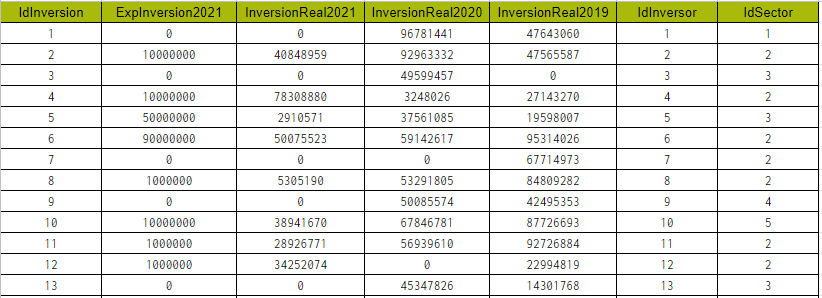
.



-**Proyecto**: en esta tabla se encuentra el listado de inversores que presentaron proyecto para el año 2022 (año correlativo según el dataset escogido). Las columnas con su tipo de dato son: idProyecto (int PK), PresentoProyecto (boolean), Intencion (text), IdInversor (int FK).



-**Inversiones**: en esta tabla se encuentra en detalle las inversiones realizadas por los inversionistas en los años 2019, 2020 y 2021 (en este caso la inversión real y la estimada). Las columnas con su tipo de dato son: IdInversion (int PK), ExpInversion2021 (float), InversionReal2021 (float), InversionReal2020 (float), InversionReal2019 (float), IdInversor (int FK), IdSector(int FK).



Creando el Dashboard

Como primera acción se realizó una reunión con el equipo del Área financiera de la empresa para conocer las especificaciones que requerían para su reporte y a la vez para concretar la entrega del dataset original.

Una vez que se tuvo el dataset se realizó una limpieza de datos y la disposición en 5 tablas que mantuvieran la integridad de los datos del dataset original y que permitan la claridad de la información al momento de buscar representarlas visualmente en un gráfico.

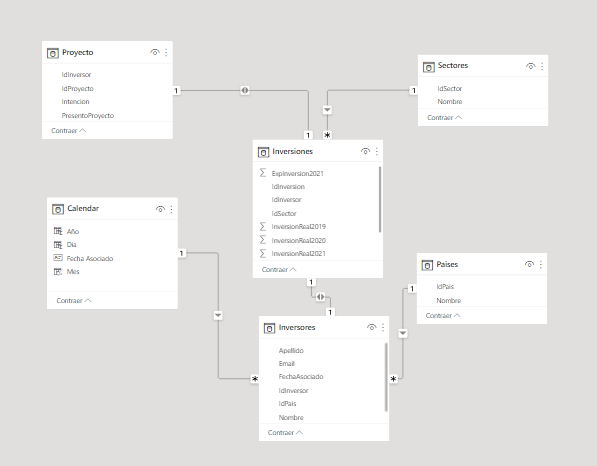
Luego se realizó un Mockup para identificar la manera más eficiente de disponer la información y se presentó nuevamente al área de finanzas para su aprobación. En esta reunión se concretó que el Dashboard quedaría dividido en 3 partes principales pudiendo contener información adicional. Estas áreas fueron: Inversionistas, Inversiones y Proyectos.

1. Inversionistas

Como la empresa había sido muy clara en la preferencia en enfocar el reporte en sus inversionistas se decidió que ésta sería la primera página. La misma contiene información detallada acerca de la cantidad de inversores, su país de origen, su fecha de asociación y sus diferentes inversiones en la firma.

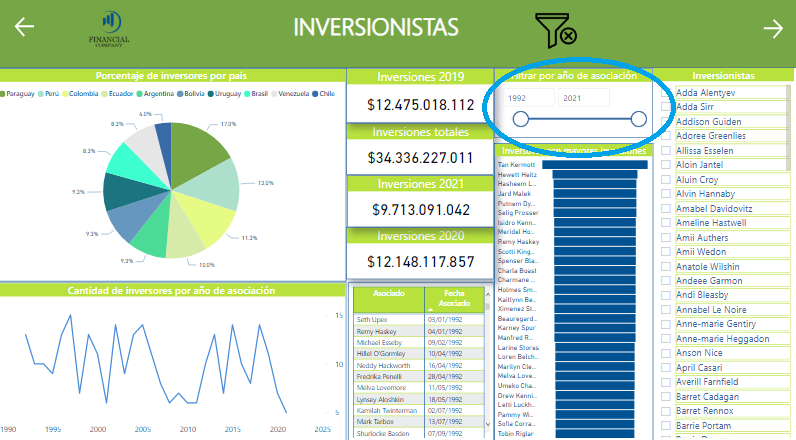
El modelo relacional original fue modificado debido a la creación de una tabla calendario creada con el motivo de mantener la persistencia de los datos considerando que Power Bi frente a la ausencia de registros los reemplaza automáticamente con alguna fecha que aun este libre considerando un calendario anual completo. La tabla Calendar tiene una relación de uno a muchos con la tabla Inversores, es decir que, cada inversor tiene solo una fecha de asociación, pero una misma fecha de asociación puede ser compartida por más de un inversor. Entonces el modelo relacional quedo conformado de la siguiente manera:

Modelo relacional en Power Bi



Al iniciar la página nos encontramos con un gráfico de torta que muestra porcentualmente la cantidad de inversores por país, siendo Paraguay el país que tiene más cantidad de inversores. También encontramos un gráfico de embudo que nos muestra los inversores que mayor capital en inversiones tienen y una tabla donde podemos apreciar los inversores con mayor antigüedad. El listado total de inversores actúa como filtro, al seleccionar uno de ellos podremos conocer su país de origen, el año en el que se asoció a la empresa, las inversiones totales que registra y estas mismas discriminadas por año.

Se agregó un filtro para poder realizar una búsqueda personalizada de usuarios por año de asociación a la empresa. El mismo se ilustra en la siguiente imagen:



La página cuenta con un icono (X) que permite reestablecer a los valores originales antes de los filtros y también símbolos de flechas a ambos lados que permiten navegar a la página anterior y posterior.

Especificaciones:

* Como primera acción se realizó una tabla calendario a partir de la columna fechaAsociado para tener una mejor persistencia de los datos. La tabla se llama Calendar y en primer lugar se creó una columna que engloba todas las posibles fechas sin repetir ninguna de ellas. La misma se creó con la fórmula *CALENDARAUTO ()* y fue establecida como tabla de fechas*.* Luego a esta columna se le modificó el nombre por fechaAsociado y se dio el formato short date. Después se crearon 3 columnas adicionales que filtran sólo los años, los meses y los días de la columna fechaAsociado. Para crearlas se utilizaron las siguientes fórmulas:

*Año = YEAR ('Calendar'[Fecha Asociado])*

*Mes = FORMAT (Calendar [Fecha Asociado],"mmmm")*

*Dia = DAY ('Calendar'[Fecha Asociado])*

* Se creó una nueva columna llamada nombre completo cuya función es concatenar el nombre y el apellido. Se utilizó la siguiente fórmula:

*NombreCompleto =CONCATENATE(Inversores[Nombre], Inversores[Apellido]).*

* En el gráfico de embudo se realizó un filtro para visualizar a aquellos inversores que poseen un capital mayor a 200.000.000 millones de pesos para limitar la cantidad de registros en la visualización y poder apreciar la finalidad del mismo, que es justamente, representar los inversores que poseen mayor capital invertido en la firma.



1. Inversiones

En esta segunda página nos encontramos con toda la información referente a las inversiones, donde podemos visualizar el porcentaje que representa cada sector sobre las inversiones totales, siendo tecnología el sector con mayor capital. También podemos visualizar mediante un gráfico de cintas la diferencia que hubo entre la expectativa de inversión y la inversión real en el año 2021.

Por último, filtrando por país podemos conocer la cantidad de inversiones que hubo en total y también por cada año, conocer los sectores en los que se realizaron dichas inversiones, así como sus capitales totales.

La página cuenta con un icono (X) que permite reestablecer a los valores originales antes de los filtros y también un símbolo de flecha en el borde superior izquierdo que permite navegar a la página anterior.

Especificaciones:

Se realizaron varias medidas para esta página, las mismas se encuentran agrupadas en la tabla Medidas, cuyas fórmulas se detallan a continuación:

* *Cantidad inversiones 2019 = DISTINCTCOUNTNOBLANK(Inversiones[InversionReal2019])*

Esta medida realiza un conteo del total de inversiones realizadas en 2019 sin considerar aquellas que están en blanco.

* *Cantidad inversiones 2020 = DISTINCTCOUNTNOBLANK(Inversiones[InversionReal2020])*

Esta medida realiza un conteo del total de inversiones realizadas en 2020 sin considerar aquellas que están en blanco.

* *Cantidad inversiones 2021 = DISTINCTCOUNTNOBLANK(Inversiones[InversionReal2021])*

Esta medida realiza un conteo del total de inversiones realizadas en 2021 sin considerar aquellas que están en blanco.

* *Cantidad total de inversiones = [Cantidad inversiones 2019]+[Cantidad inversiones 2020]+[Cantidad inversiones 2021]*

Esta medida hace una sumatoria de la cantidad total de inversiones.

* *ExpInversionTotal2021 = sum(Inversiones[ExpInversion2021])*

Esta medida realiza una sumatoria del total de expectativa de inversión para el año 2021.

* *Total inversiones 2019 = SUM(Inversiones[InversionReal2019])*

Esta medida realiza una sumatoria del valor de todas las inversiones del año 2019.

* *Total inversiones 2020 = SUM(Inversiones[InversionReal2020])*

Esta medida realiza una sumatoria del valor de todas las inversiones del año 2020.

* *Total inversiones 2021 = SUM(Inversiones[InversionReal2021])*

Esta medida realiza una sumatoria del valor de todas las inversiones del año 2021.

* Se creó una nueva columna llamada *Suma de inversiones totales por Inversor.* La misma contiene la sumatoria de las inversiones de los 2019, 2020 y 2021 para cada inversor. La fórmula utilizada fue la siguiente:

*Suma de inversiones totales por Inversor =Inversiones[InversionReal2019]+Inversiones[InversionReal2020]+Inversiones[InversionReal2021].*

Luego se agregó el formato ‘$’ para todos los registros de esta columna.

1. Proyectos

La última página está destinada a los proyectos que fueron presentados para llevarse a cabo en 2022, teniendo en cuenta que el dataset es de hace un año. En este apartado podemos visualizar el porcentaje de inversores que presentaron proyectos, la cantidad por sector y por intención. Esta última hace referencia al destino prioritario de ese capital, pudiendo ser trabajo, conocimiento o inversión.

Filtrando por inversionista podemos conocer si presento o no proyecto para 2021 y, si presento proyecto, podemos conocer a que sector pertenece, su país de origen y su intención.

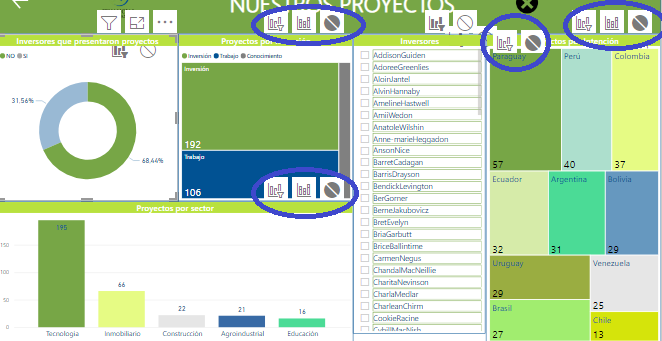
La página cuenta con un icono (X) que permite reestablecer a los valores originales antes de los filtros y también símbolos de flechas a ambos lados que permiten navegar a la página anterior y posterior.

Especificaciones:

* En el listado de inversores se realizó un filtro para que éste sólo muestre aquellos inversores que han presentado proyecto para el año 2021 para evitar la inconsistencia en los datos arrojados.



* Para el grafico de torta que muestra el porcentaje de inversores que presentaron proyecto para el 2021 se deshabilitaron los filtros para el resto de los gráficos para evitar la inconsistencia en los datos arrojados, ya que no tendría sentido mostrar la intención o el sector en un proyecto que no existe porque no fue presentado.

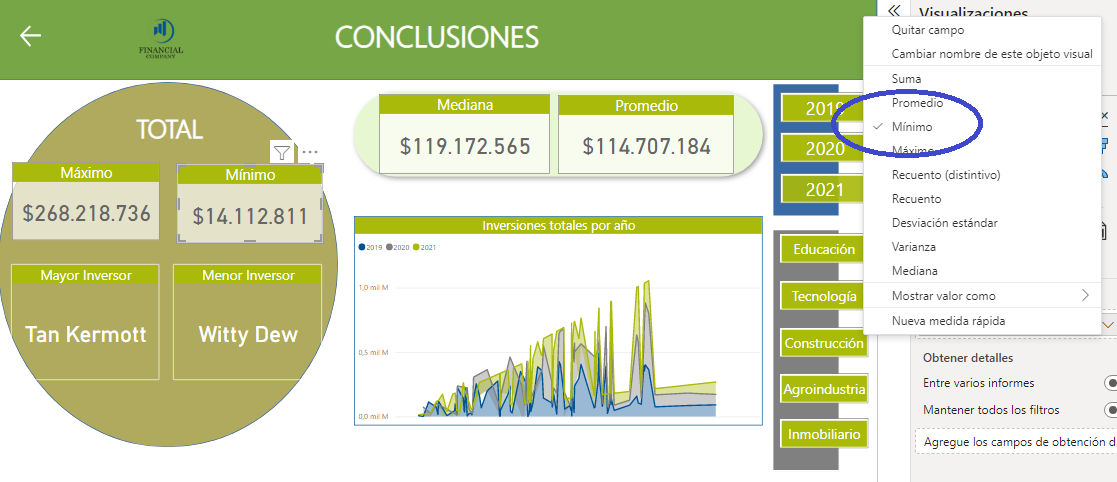


Conclusiones

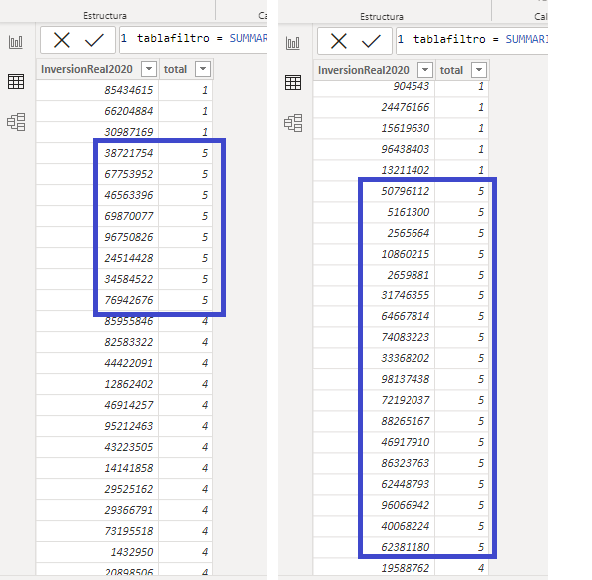
Para realizar las conclusiones de este trabajo se realizaron diferentes cálculos con medidas de tendencia central teniendo en cuenta los tres años analizados (2019,2020,2021) y los 5 sectores involucrados (Educación, Tecnología, Construcción, Agroindustrial e Inmobiliario).

En la página central se presentan los datos de promedio, mediana, máximo y mínimo para los valores totales, es decir, teniendo en cuenta la sumatoria del total de años y el total de inversiones. También se identificó al inversor que tuvo el mejor desempeño respecto al total invertido, así como el que menos invirtió.

Como estos datos no se iban a utilizar para posteriores análisis se realizó el cálculo simplemente filtrando desde el mismo dato que arrojaba la suma total.



Se decidió no considerar la medida central MODA ya que eran muchos los valores diferentes que se repetían la misma cantidad de veces. En esta imagen se muestra a modo de ejemplo las inversiones para el año 2020, pero pasaba lo mismo con el 2019 y 2021.



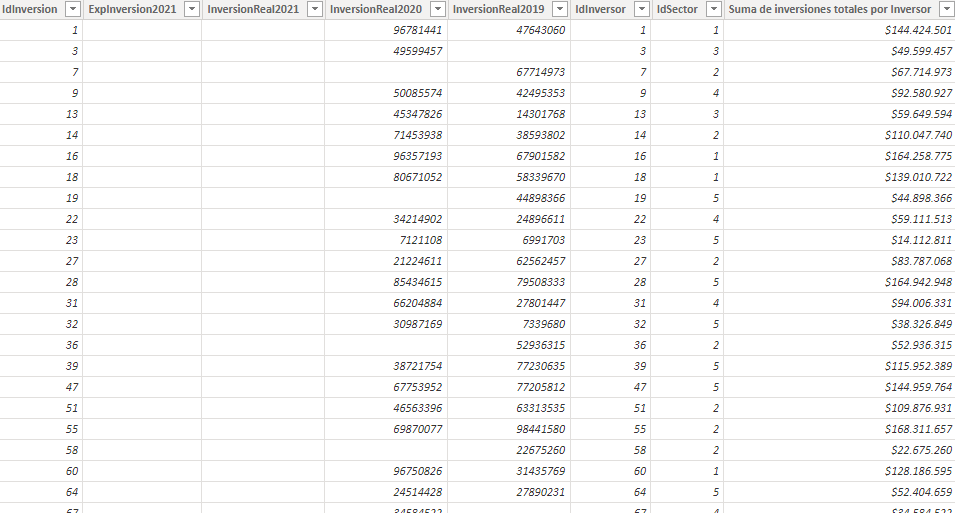
En la misma página se encuentra un menú de botones que redirigen a diferentes páginas según el dato que se quiera visualizar.



Para cada año se muestra la mayor y menor inversión, la medida mediana, el promedio y la suma total. También se puede visualizar rápidamente quien fue el mayor y el menor inversor para ese periodo. Al presionar el icono que se muestra debajo se regresa a la página principal de las conclusiones.



Debido a inconsistencias en los datos arrojados para el valor promedio y moda se modificaron los registros de la tabla inversiones que tenían valor 0, siendo reemplazados por valor null. De esta forma no son tenidos en cuenta en las diferentes medidas calculadas y se evita distorsionar los datos.



Para los años 2019, 2020 y 2021 se realizaron las siguientes medidas calculadas tomando como ejemplo el año 2020:

*maximo2020 = MAX(Inversiones[InversionReal2020])*

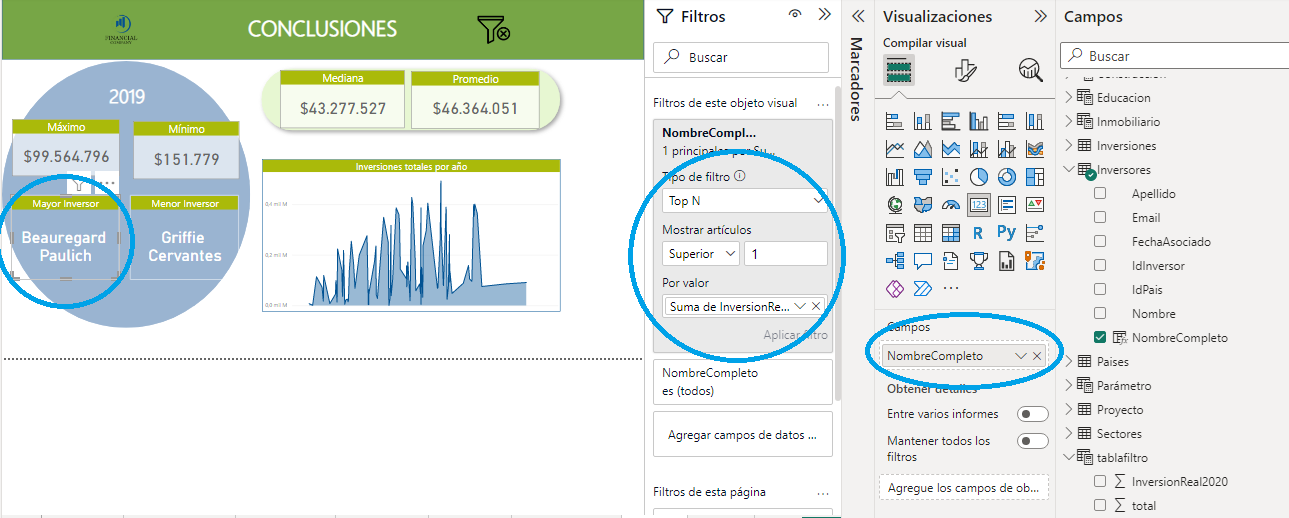
*minimo2020 = MIN(Inversiones[InversionReal2020])*

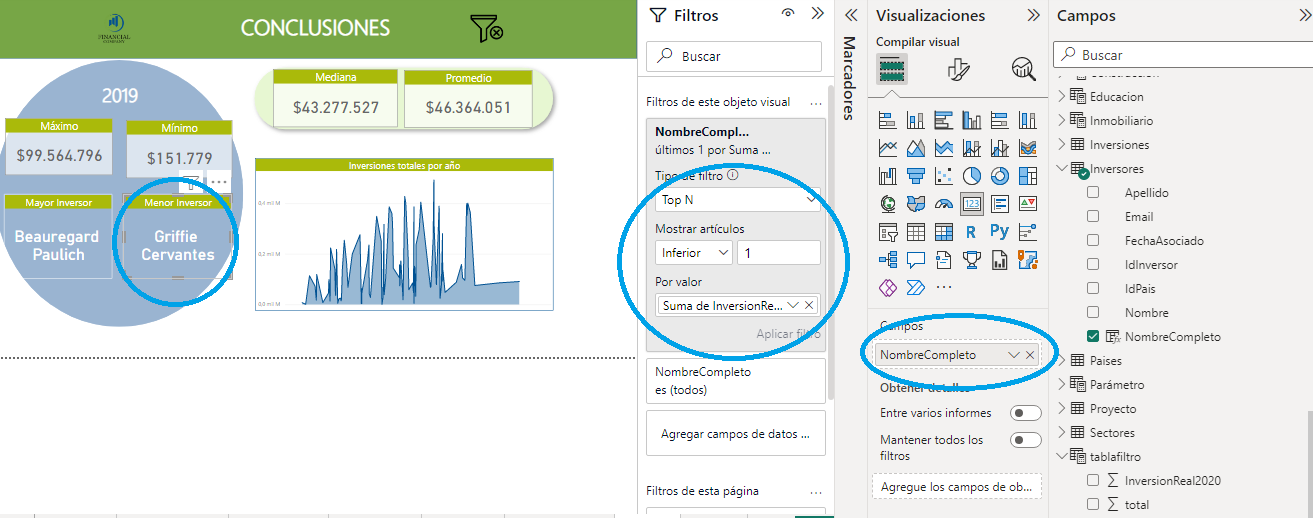
*promedio2020 = AVERAGE(Inversiones[InversionReal2020])*

*mediana2020 = MEDIAN(Inversiones[InversionReal2020])*

Para mostrar el mayor y el menor inversor se creó una tarjeta que muestra los nombres de los asociados colocándole un filtro que toma sólo el primer valor superior e inferior respectivamente teniendo en cuenta la sumatoria de inversiones (medida explicada en la sección inversores).

La primera imagen corresponde al filtro utilizado para obtener el mayor inversor y la segunda imagen corresponde al filtro utilizado para obtener el menor inversor.





Para segmentar los sectores se realizó una nueva tabla por cada sector filtrando de la tabla Inversiones. A modo de ejemplo se muestra debajo la fórmula utilizada para el sector Educación representado por el número 1 en idSector de la tabla Inversiones. De esta misma manera se completó para el resto de los sectores:

*Educacion = FILTER(Inversiones,Inversiones[idSector]=1)*

Una vez que se ingresa por medio de los botones correspondientes a cada sector se puede visualizar un valor total que representa la suma de inversiones que hubo en ese sector en los 3 años consecutivos (2019,2020,2021).

Para realizar el cálculo del total de inversiones por sector se realizó la siguiente medida calculada, la misma ejemplifica para el sector Educación.

*EducacionTotal = CALCULATE(Medidas[Inversionessectoreducacion2019]+Medidas[Inversionessectoreducacion2020]+Medidas[Inversionessectoreducacion2021])*

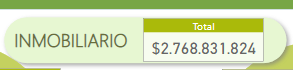
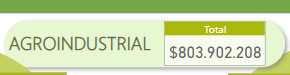
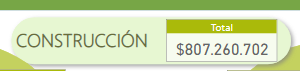
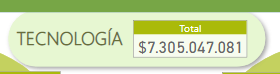
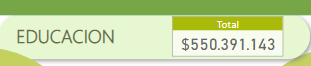
Para realizar la sumatoria de inversiones por año y sector se realizó la siguiente medida tomando como ejemplo para el sector Educación y el año 2020:

*Inversionessectoreducacion2020 = CALCULATE(SUM(Inversiones[InversionReal2020]),Inversiones[idSector]=1)*

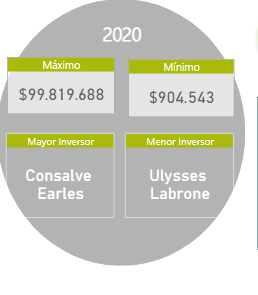
Luego apretando los botones 2019, 2020 y 2021 se puede ver el total para cada año respectivamente. También se puede visualizar la mediana, el promedio, el valor máximo y mínimo por cada año. Para volver a la página principal del sector se aprieta el icono del filtro.

Conclusiones finales

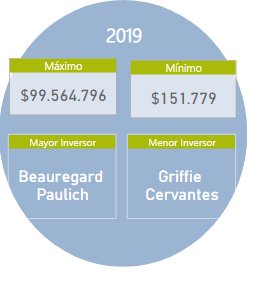
* Por una diferencia de casi 5 millones el sector Tecnología es el que ha tenido mayores inversiones y es, por lo tanto, el más influyente. Por el contrario, el sector que menos inversiones ha recibido es el de Educación.

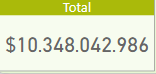


* El total de inversiones muestra un considerable descenso, siendo de -4% de 2019 a 2020 y de -19% de 2020 a 2021. Los valores máximos se muestran muy parejos, así como el valor promedio. No tanto así el valor mínimo de las inversiones.





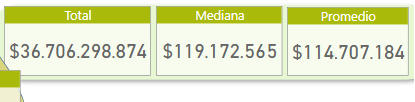




* Para el sector Educación no hubo inversiones en el año 2021.
* En el sector Educación la mayor inversión se realizó en el 2020 con el 74% y en 2019 fue de un 26%. Podemos inferir que hubo un gran crecimiento de 2019 a 2020 pero luego decreció a 0 en 2021.
* En el sector Tecnología las inversiones se mantuvieron muy parejas siendo de 39% en 2019, 30% en 2020 y 31% en 2021.
* En el sector Construcción se muestra un decrecimiento siendo de un 44% las inversiones en 2019, 32% en 2020 y de 24% en 2021.
* En el sector Agroindustrial las inversiones representaron un 25% para 2019, luego subieron a 45% en 2020 para terminar desciendo a 30% en 2021.
* En el sector Inmobiliario se muestra una situación similar al del Agroindustrial. Comienza con una representación de 35% en 2019, luego aumenta a un 40% en 2020, para terminar en 25% en 2021.
* El mayor inversor considerando los 3 años consecutivos fue Tan Kermott y el menor inversor fue Witty Dew, ambos pertenecientes a Perú.
* Se presentaron un total de 320 proyectos para el año 2022 siendo el sector Tecnología el más representado con 195 proyectos. La Intención predominante es la Inversión con 192 proyectos. Sólo el 31.56% de los asociados presentaron proyecto para el 2022.
* La incorporación de asociados es muy variante pasando de años en donde hay un aumento considerable a otros años de baja con la misma intensidad tal como se muestra en la imagen siguiente:



* El país que tiene más cantidad de inversores es Paraguay y también es el más influyente en total de inversiones con un total de $6.192.123.383.
* La expectativa de inversión para 2021 muestra ser inferior significativamente respecto a la inversión real para los sectores Tecnología, Inmobiliario y Agroindustrial. Por el contrario, es superior a la inversión real para el sector de la Construcción.
* El dataset analizado presenta una ligera simetría en sus medidas de tendencia central, con una leve asimetría hacia la izquierda ya que el promedio es levemente inferior a la mediana, tal como se muestra en la siguiente imagen:



Recomendaciones

* Para mejorar el análisis de datos se recomienda en adelante especificar las fechas de cada inversión y no solo el año, ya que de esa manera se podría conocer el comportamiento de las inversiones a lo largo del mismo.
* Para tener un conocimiento más actualizado de la distribución geográfica de los socios interesaría conocer el país de residencia de los mismos, y no solo su país de origen. De esta manera podríamos conocer en qué países o ciudades se concentran las mayores inversiones.