

Trabajo Final – Data Analytics

Análisis de los suicidios ocurridos durante el período entre 1985 y 2015.



Data Analytics - Comisión 18905

Alumnos:

- Barbona, Neri Axel
- Alegria, Néstor

Profesora:

- Rattia, Ana Daniela

Tutor:

- Rubin, Fernando

Contenido |

Descripción de la temática.....	2
Objetivo, alcance y usuario final del proyecto.....	2
Manual de marca	3
Diagrama entidad relación.....	4
Listados de tablas, con definición de clave primaria.....	5
Países.....	5
Generación.....	6
Edades	6
Modelo relacional.....	6
Especificaciones	7
Columnas, medidas y tablas calculadas	8
Análisis funcional del tablero	9
Medidas calculadas	15
Herramientas empleadas.....	17
Consideraciones finales.....	18

Data Analytics

Descripción de la temática.

El presente informe tiene por objetivo realizar un análisis descriptivo de los suicidios ocurridos entre los años 1985 y 2015.

Para llevar a cabo este análisis se utilizó una base de datos obtenida en Kaggle. A partir de los datos se podrá obtener información acerca de la cantidad de hombres y mujeres fallecidos por año, el porcentaje de población que representa el total de fallecidos y cuál fue el rango etario con más registros. Por otra parte, estos datos nos permitirán hacer una comparación con los datos obtenidos de Brasil, Uruguay y Chile.

En primer lugar, se aborda el análisis desde una perspectiva global. Luego, nos centraremos en el impacto a nivel latinoamericano, tomando como referencia a Argentina, Brasil y Chile.

Por último, se realizó una descomposición de los datos de Argentina durante el periodo entre 1999 y 2008 con el fin de ver el impacto de la crisis económica sobre la tasa de suicidios, tomando como dato el PBI per cápita.

Para más información sobre esta problemática, pueden visitar los siguientes enlaces:

- [Suicidio - Prevención | Argentina.gob.ar](https://Argentina.gob.ar)
- [Prevención del suicidio - OPS/OMS | Organización Panamericana de la Salud \(paho.org\)](https://paho.org)

Objetivo, alcance y usuario final del proyecto

El objetivo general es crear conciencia sobre la prevención del suicidio en todo el mundo. Los objetivos incluyen promover la colaboración de las partes interesadas y el auto-empoderamiento para abordar las autolesiones y el suicidio a través de acciones preventivas.

Se busca brindar a los potenciales grupos de interés (gobiernos, organizaciones, etc) una manera de poder analizar la evolución de la tasa de suicidios a nivel

global y local para llevar políticas o actividades destinadas a disminuir la problemática.

Además, permite detectar cuales son los grupos más vulnerables, diferenciándolos por edad, generación, sexo.

Manual de marca

El logo seleccionado para implementar durante el proyecto fue el diseñado y utilizado por la Organización Mundial de la Salud (OMS) para realizar las campañas de prevención del suicidio.



En referencia a la fuente empleada en el panel de Power BI es de tipo Segoe UI y en el presente documento es Helvetica Neue. El tamaño puede diferir según donde se emplee, aún así se aseveró que fueran lo más uniforme posible entre secciones, títulos, subtítulos, etc.

La paleta de colores utilizada para el fondo de Power BI es colores oscuros:



Mientras tanto, los colores utilizados para los gráficos son colores con tonalidad pastel:

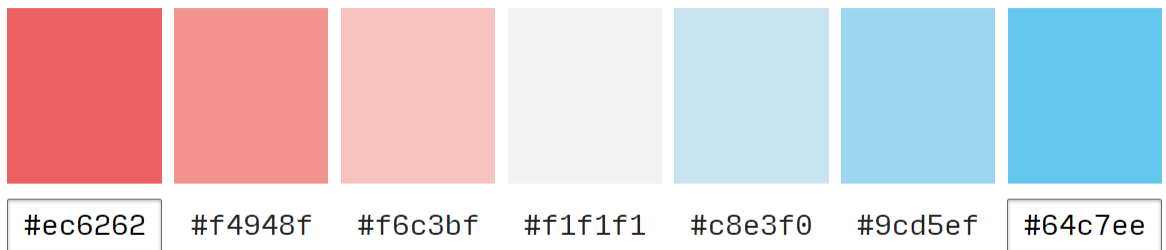
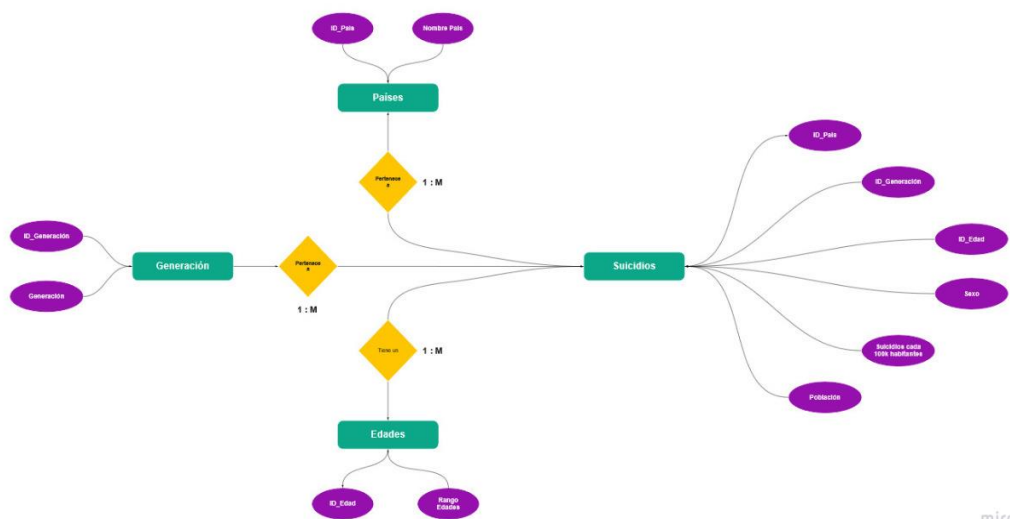


Diagrama entidad relación.

Se adjunta el diagrama entidad-relación:





Listados de tablas, con definición de clave primaria.

En el siguiente apartado se exponen las tablas utilizadas, las columnas que las componen y el tipo de dato que agrupan. A su vez, se definen las claves primarias y claves foráneas.

Tipo de clave	Campo	Tipo de campo
PK-index	id_suicidio	int
FK	id_pais	int
FK	id_generacion	int
FK	id_edad	int
-	Sexo	varchar (6)
-	Suicidios cada 1000 habitantes	int
-	población	int

Países

Tipo de clave	Campo	Tipo de campo
PK-index	id_pais	int

-	Nombre país	varchar(100)
---	-------------	--------------

Generación

Tipo de clave	Campo	Tipo de campo
PK-index	id_generacion	int
-	Generación	varchar(100)

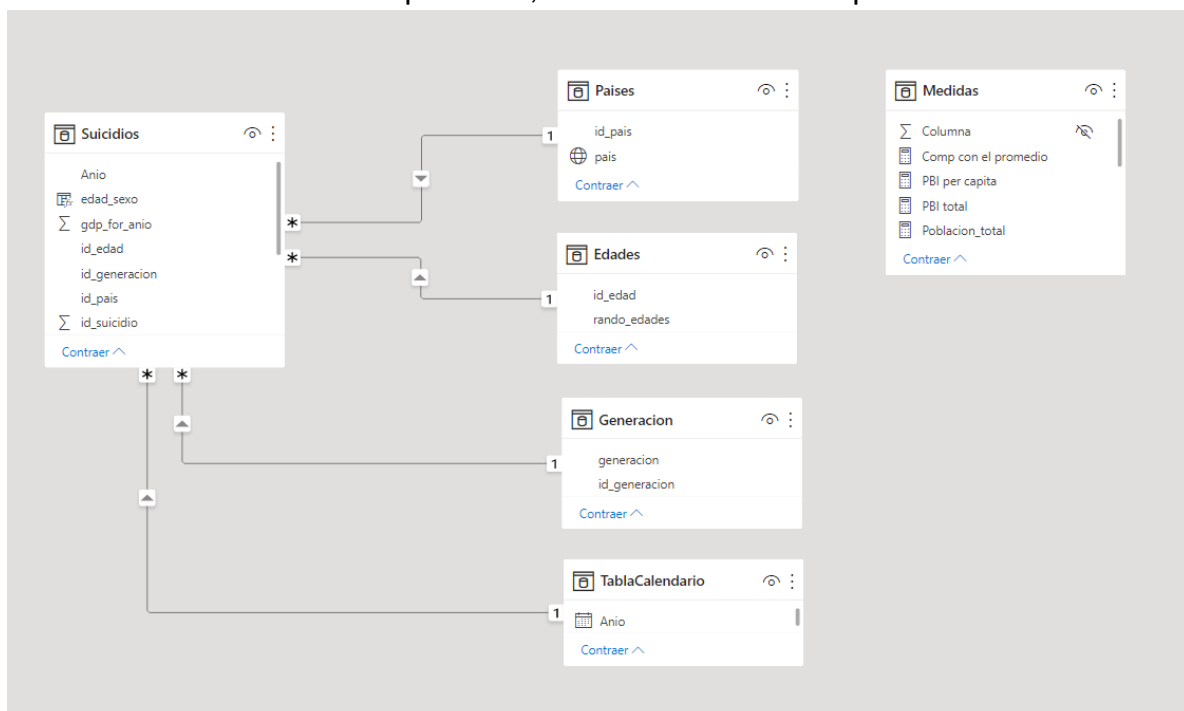
Edades

Tipo de clave	Campo	Tipo de campo
PK-index	id_edad	int
-	Rango edades	varchar(100)

Modelo relacional

El siguiente diagrama cuenta con:

- Listado de tablas, con definición de clave primario y foránea, según corresponda.
- Listado de columnas por tabla, con definiciones del tipo de dato.



Especificaciones

Antes de comenzar el análisis de la información proveniente de nuestra base de datos, debemos realizar una serie de transformaciones para un análisis eficiente.

En primer lugar, debemos vincular la base de datos a nuestro Power BI y distinguir entre las diferentes tablas presentes en el modelo. Luego, debemos identificar los distintos formatos de los datos que se encuentran en cada columna de las distintas tablas. En este caso, debemos transformar la columna “gdp_for_years” a una columna de tipo numérica para poder hacer el análisis posterior.

Por otra parte, debemos crear una nueva tabla llamada “tabla calendario” que nos permita la incorporación de un parámetro temporal al modelo.

Una vez realizadas las transformaciones necesarias en las tablas, se cargan a Power BI y se verifica que las tablas estén relacionadas entre sí. En este caso distinguimos a una principal llamada “Suicidios” y las tablas secundarias “Países”, “Edades”, “Generación” y la tabla calendario.

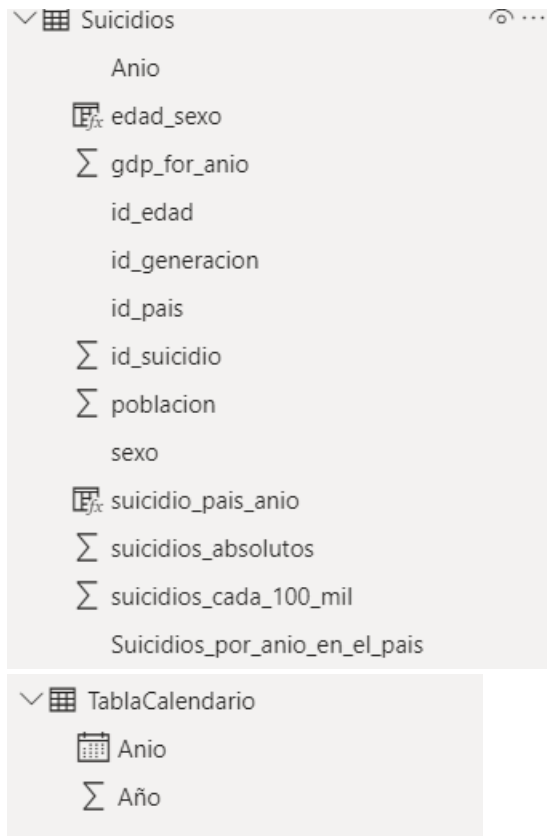
Por otra parte, se crearon dos columnas calculadas llamadas “edad_sexo” que relaciona la edad y el sexo del caso, y otra llamada “Suicidio_pais_año”, la cual concatena el id del país y el año del evento.

Además se crearon medidas calculadas, entre las cuales se encuentra “PBI Total” que nos permite relacionar el PBI de los países de acuerdo al año analizado, “Poblacion total” que nos permite analizar la relación de acuerdo al crecimiento demográfico de los países bajo análisis, “Suicidios_por_100m_hab” que nos arroja información acerca de los suicidios cada 100 mil habitantes en determinado año y país, y por último “Suicidios_por_generación” que nos permite analizar los suicidios de acuerdo a la variable generacional.

Columnas, medidas y tablas calculadas

The image shows a data model view with a hierarchy of tables and calculated measures. The tables are: Medidas, Edades, Generacion, Países, Suicidios, and TablaCalendario. The 'Medidas' table is expanded, showing a list of calculated measures: Columna, Comp con el promedio, PBI per capita, PBI total, Poblacion_total, Promedio tasa suicidios, Suicidios promedio Europa, Suicidios_abs, Suicidios_por_100m_hab, Suicidios_por_generacion, total_por_población, Ultima_tasa_suicidios, and Ultimo_año. The 'Edades' table is expanded, showing 'id_edad' and 'rando_edades'. The 'Generacion' table is expanded, showing 'generacion' and 'id_generacion'. The 'Países' table is expanded, showing 'id_pais' and 'pais'.

- > Medidas
 - > Edades
 - > Generacion
 - > Países
 - > Suicidios
 - > TablaCalendario
- Medidas
 - Columna
 - Comp con el promedio
 - PBI per capita
 - PBI total
 - Poblacion_total
 - Promedio tasa suicidios
 - Suicidios promedio Europa
 - Suicidios_abs
 - Suicidios_por_100m_hab
 - Suicidios_por_generacion
 - total_por_población
 - Ultima_tasa_suicidios
 - Ultimo_año
- Edades
 - id_edad
 - rando_edades
- Generacion
 - generacion
 - id_generacion
- Países
 - id_pais
 - pais



Dentro de la tabla suicidios, se creo una columna calculada llamada “Suicidios_absolutos” de la siguiente manera:

```
= Table.TransformColumns(#"Tipo cambiado3",{ "suicidios_absolutos", each Number.Round(_, 2), type number })
```

Análisis funcional del tablero

El tablero inicia con la solapa de “Portada”, la cual cuenta con marcadores que nos permiten dirigirnos hacia las distintas solapas, las cuales son:

- Introducción
- Análisis global
- Análisis Latinoamericano
- Argentina
- Brasil
- Chile
- Crisis económica Argentina



En primer lugar podremos dirigirnos a la solapa de introducción que nos brinda la descripción del informe y la forma en la cual se abordó.

Introducción

El siguiente informe tiene por objetivo realizar un análisis descriptivo de los suicidios ocurridos entre los años 1985 y 2015.

En primer lugar, se aborda el análisis desde una perspectiva global. Luego, nos centraremos en el impacto a nivel latinoamericano, tomando como referencia a Argentina, Brasil y Chile.

Por último, se realizó una descomposición de los datos de Argentina durante el periodo entre 1999 y 2008 con el fin de ver el impacto de la crisis económica sobre la tasa de suicidios, tomando como dato el pbi per cápita.

Para más información sobre esta problemática, pueden visitar los siguientes enlaces:

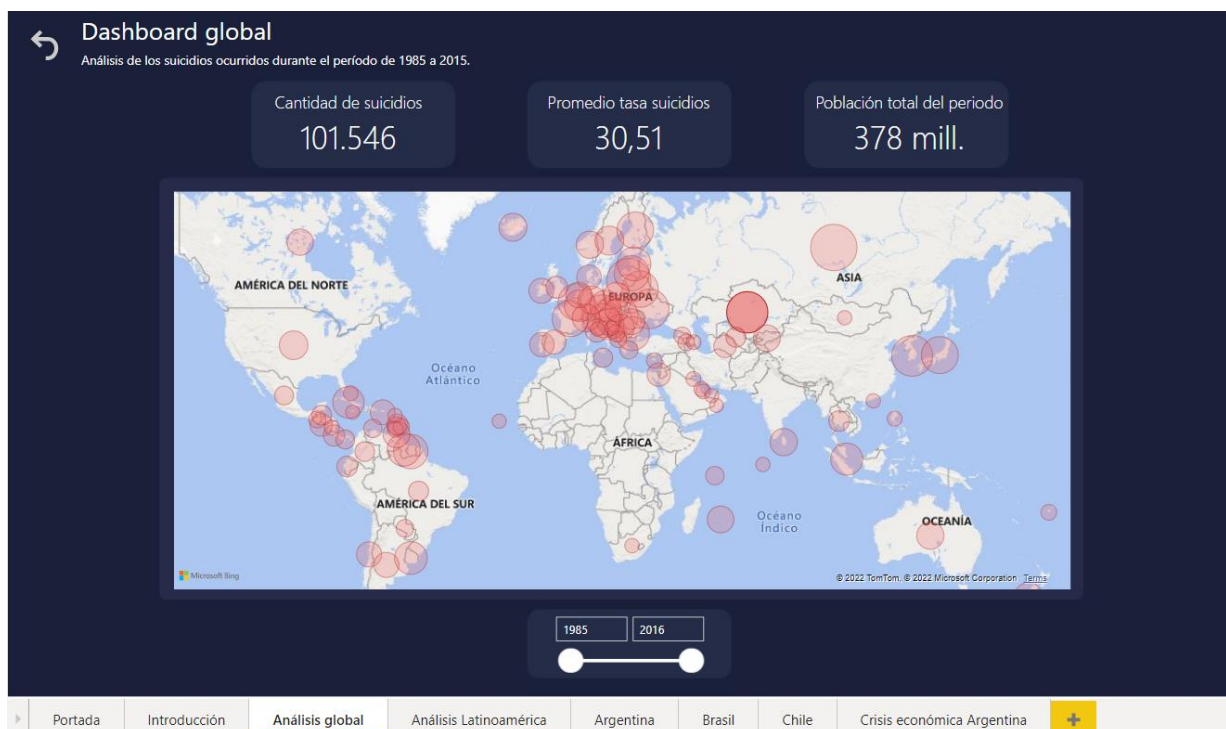
- Prevención de suicidios - Ministerio de Salud
- Organización Panamericana de la Salud

Luego comenzamos con el informe propiamente dicho, empezando por el “Análisis global”, el cuál muestra información sobre la tasa de suicidios de los distintos países en el período entre 1985 y 2015. En esta solapa podemos encontrar las siguientes tarjetas:

- Cantidad de suicidios (en términos absolutos)
- Promedio de tasa de suicidios a nivel mundial
- Población total del periodo analizado

Debajo vamos a encontrar un mapa que mediante burbujas refleja la tasa de suicidios por país. A mayor circunferencia de las burbujas, mayor será la tasa de suicidios del país.

Por último, encontramos un filtro por años que nos permita seleccionar el periodo a considerar.



El apartado “Análisis Latinoamérica” cuenta con tres tarjetas, a saber:

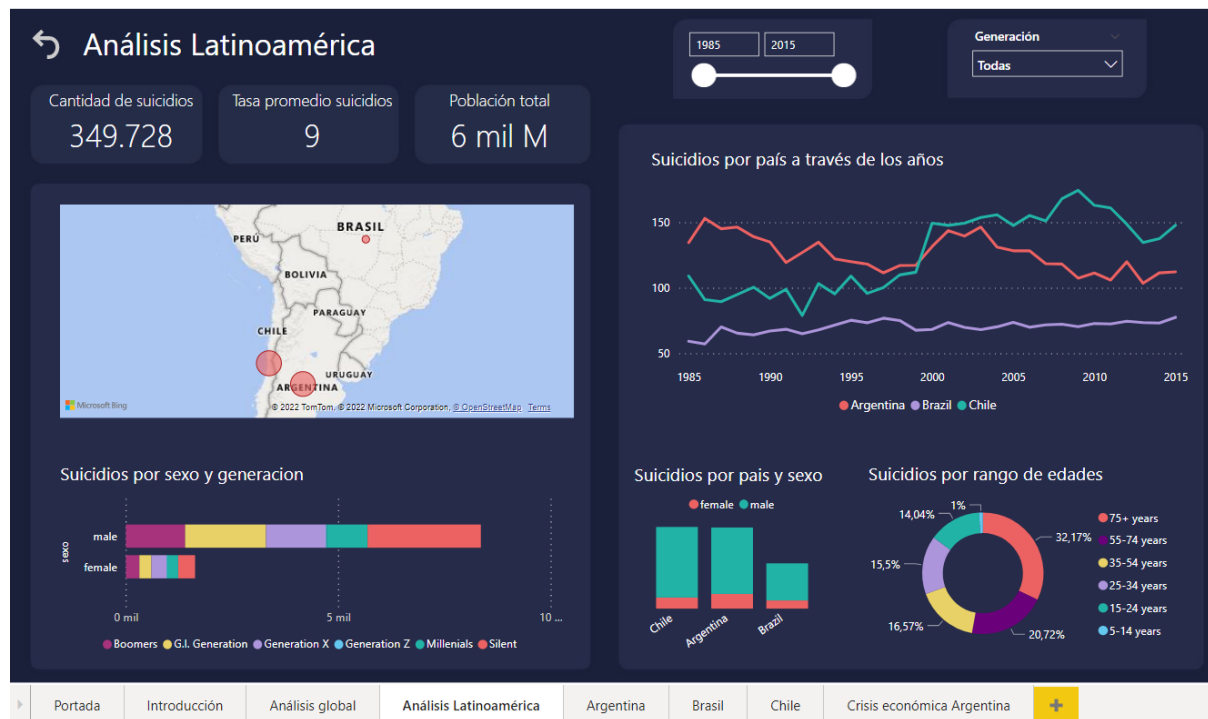
- Cantidad de suicidios (en términos absolutos)
- Tasa promedio de suicidios
- Población total

Si bien la página cuenta con un filtro por página de los países estudiados (Argentina, Brasil y Chile), en la parte superior también se pueden visualizar dos tipos de filtros adicionales que permiten especificar los datos a considerar, los cuales son la generación y el período de años.

Al análisis se suma un mapa que indica la tasa de suicidios sobre los países considerados.

Además, se suman otros gráficos que comparan los datos de Argentina, Brasil y

Chile en términos relativos, es decir, cada 100 mil habitantes. En formato de gráfico de líneas se analiza la evolución de la tasa de suicidios a través de los años, mediante un gráfico de barras apiladas se compara los suicidios de acuerdo al sexo y, a su vez, de acuerdo a la generación a la cual pertenecen. Además, se analiza los suicidios por país y sexo mediante un gráfico de columnas apiladas y, por último, un mediante un gráfico de anillos se presenta un análisis de los suicidios de acuerdo al rango de edades.



Por otro lado, se encuentra la solapa “Argentina” el cuál analiza la información de ese solo país aplicándolo como filtro de la página, además de contar con la posibilidad de utilizar los filtros por año y generación.

Las tarjetas que se encuentran en la parte superior de la página indican:

- Cantidad de suicidios (en términos absolutos)
- La población total bajo análisis.
- Un medidor que indica la tasa promedio de suicidios del periodo como meta, comparándola con la última tasa de suicidios registrada(en términos relativos).

Mediante un gráfico de columnas y líneas se busca representar la información de los suicidios a través de los años a la vez que se indica la evolución del PBI per cápita, intentando buscar una relación entre estos datos. Por otra parte, se

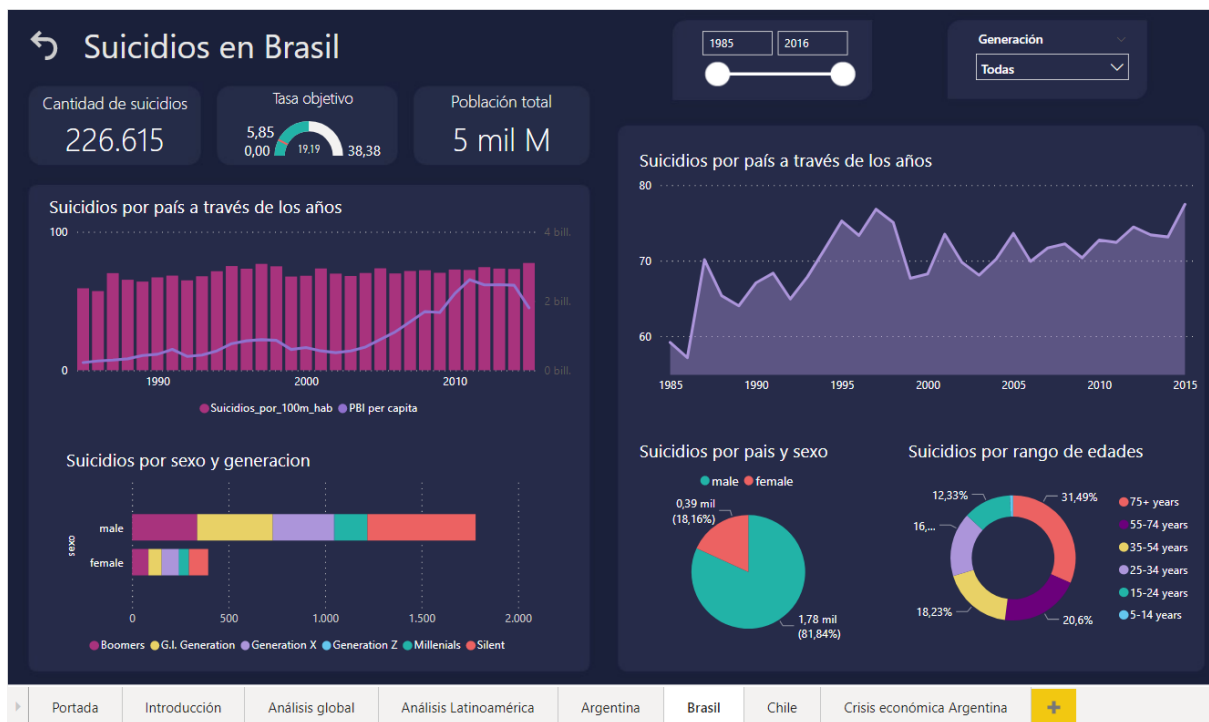
presentan los gráficos presentados en el “Análisis Latinoamérica”, pero especificando, en este caso, el país. Este último cambio llevo a tener que implementar un grafico circular en el análisis de suicidios por sexo, por ser este más eficiente que un grafico de barras.

De la misma manera que se implementó el análisis sobre los datos de Argentina, se realizo sobre Brasil y luego Chile.

- Argentina:



- Brasil:



○ Chile:



Por último, a partir de los datos de Argentina durante el período entre 1999 y 2008 se intentó evaluar un posible impacto de la crisis económica sobre la tasa de suicidios, sumando como elemento al estudio la variación del PBI per cápita.

En la solapa podemos observar mediante tarjetas:

- PBI per cápita de 1999 (precrisis)

- PBI per cápita de 2002 (durante la crisis)
- PBI per cápita de 2008 (postcrisis)

Por otro lado, podemos ver la variación de la tasa de suicidios, de manera relativa, durante el periodo analizado.

Además, se sumaron los gráficos utilizados anteriormente para la observación de los suicidios por sexo, por rango de edades y por sexo y generación.



Medidas calculadas

Dentro de la tabla "Medidas" podemos encontrar las distintas medidas calculadas aplicadas en el tablero. Entre ellas se encuentran:

- Comparación con el promedio: Considera el número de suicidios cada 100 mil personas y la tasa promedio de suicidios, arrojando si se encuentra en crecimiento o no.

```
1 Comp con el promedio =
2   VAR verdadero = "En crecimiento"
3   VAR falso = "En descenso"
4   RETURN
5   IF(SUM('Suicidios'[suicidios_cada_100_mil])>[Promedio tasa suicidios], verdadero,falso)
```

- PBI total: Suma el pbi por año, considerando el filtro de año y país.


```

1 PBI total =
2 CALCULATE (
3     SUM ( Suicidios[gdp_for_anio] ),
4     FILTER(
5         Suicidios,
6         'Suicidios'[Anio] = 'Suicidios'[Anio] &&
7         'Suicidios'[id_pais] = Suicidios[id_pais]
8     )
9 )

```

- PBI per capita: Nos permite obtener el pbi por habitante. Se divide por 12 para no sumar el pbi por cada categoría.

```

1 PBI per capita = calculate(DIVIDE([PBI total],12))

```

- Población total: Considera la población de cada país, considerando el filtro del año y país.

```

1 Poblacion_total =
2 CALCULATE (
3     SUM ( Suicidios[poblacion] ),
4     FILTER(
5         Suicidios,
6         'Suicidios'[Anio] = 'Suicidios'[Anio] &&
7         'Suicidios'[id_pais] = Suicidios[id_pais]
8     )
9 )

```

- Promedio tasa de suicidios: Calcula el promedio de suicidios.

```

1 Promedio tasa suicidios = AVERAGE(Suicidios[suicidios_cada_100_mil])

```

- Suicidios abs: Calcula la cantidad de suicidios en términos absolutos.

```

1 Suicidios_abs =
2 CALCULATE (
3     SUM ( Suicidios[suicidios_absolutos] ),
4     FILTER(
5         Suicidios,
6         'Suicidios'[Anio] = 'Suicidios'[Anio] &&
7         'Suicidios'[id_pais] = Suicidios[id_pais]
8     )
9 )

```

- Suicidios_por_100m_hab

```

1 Suicidios_por_100m_hab =
2 CALCULATE (
3     SUM ( Suicidios[suicidios_cada_100_mil] ),
4     FILTER(
5         Suicidios,
6         'Suicidios'[Anio] = 'Suicidios'[Anio] &&
7         'Suicidios'[id_pais] = Suicidios[id_pais]
8     )
9 )

```

- Suicidios_por_generacion: Calcula la sumatoria de los suicidios, filtrando de acuerdo a la generación, año y país.

```

1 Suicidios_por_generacion =
2 CALCULATE (
3     SUM (Suicidios[suicidios_cada_100_mil] ),
4     FILTER(
5         Suicidios,
6         'Suicidios'[id_generacion] = 'Suicidios'[id_generacion] &&
7         'Suicidios'[Anio] = 'Suicidios'[Anio] &&
8         'Suicidios'[id_pais] = Suicidios[id_pais]
9     )
10 )

```

- Ultima_tasa_suicidios: Esta medida nos arroja la última tasa de suicidio.

```

1 Ultima_tasa_suicidios =
2 MAXX(
3     FILTER(
4         'Suicidios',
5         'Suicidios'[Anio] = 'Medidas'[Ultimo_año]
6     ),
7     'Suicidios'[suicidios_cada_100_mil]
8 )

```

- Ultimo_año: Esta medida nos arroja el último año a considerar y de esta manera poder calcular la última tasa de suicidios.

```

1 Ultimo_año = CALCULATE(
2     MAX('Suicidios'[Anio]),
3     ALLEXCEPT('Suicidios', 'Suicidios'[id_pais])
4 )

```

Herramientas empleadas

- Power BI Desktop

Para ensamblar el tablero interactivo y el análisis del proyecto.

- Power BI on Cloud

Para presentar en el SaaS de Power BI el tablero construido.

- Excel

Para manipular la data inicial en formato CSV y presentarla de manera de confeccionar el modelo de datos necesario posteriormente en Power BI.

- Image Color Picker

Para seleccionar la paleta de colores.

- Google Docs

Para contar con un medio colaborativo al momento de confeccionar la documentación.

- Google Drive

Para gestionar entregables en las distintas iteraciones del proyecto en un medio colaborativo de almacenamiento.

Consideraciones finales

Si bien nuestra investigación se limita al análisis del período entre 1985 y 2015, creemos que sería útil para la implementación de una base de datos actualizada con las mismas variables y adicionales para poder profundizar el proyecto.

Algunas variables que podrían sumarse de manera que aporten una mayor profundidad a la investigación podrían ser momentos del año en los que los hechos pasan más frecuentemente, tasa de desempleo, índice de pobreza, índices que se relacionan con la violencia de género y familiar, acceso a vivienda digna, entre otros.