

Siniestros Viales en Ciudad Autónoma de Buenos Aires

Institución: Coderhouse

Integrantes:

María Candela Curto

María García Samartino

Diego Martin Rigoli

Tutora: Yamileth Cabrera

Mayo 2022



Tabla de contenido

Tabla de contenido	2
1. Introducción	3
2. Descripción de la temática de los datos	3
3. Alcance y usuario	3
4. Objetivo	4
5. Herramientas tecnológicas implementadas	4
6. Dataset	5
7. Diagrama Entidad-Relación (DER)	5
8. Listado de tablas	5
9. Listado de columnas por tabla	6
10. Modelo relacional en Power BI	8
10.1. Eliminación de PK Null	8
10.2. Cambio en el tipo de dato	8
10.3. Renombre de campos	9
10.4. Reemplazo de valores	9
10.5. Creación de columnas	10
10.6. Eliminación de columnas	10
10.7. Tabla Calendario	10
10.8. Tabla para medidas	11
11. Medidas calculadas	11
12. DER Power BI	14
13. Segmentaciones elegidas	15
14. Paleta de colores elegida	15
15. Visualización de los datos	16
15.1. Portada	16
15.2. Glosario	17
15.3. General	17
15.4. Ubicación	19
15.5. Geolocalización	20
15.6. Víctimas	22
16. Conclusión	23
17. Futuras líneas	24
18. Bibliografía	25

1. Introducción

Cada vez es mayor la cantidad de vehículos que circulan a diario por las calles y rutas, en especial en los grandes conglomerados urbanos. Para conducir se requiere conocer y respetar una determinada cantidad de normas que permiten transitar sin sobresaltos ni contratiempos. Muchas veces, aunque sean de público conocimiento, estas pautas no se cumplen y acciones consideradas arriesgadas derivan en finales poco felices.

La Organización Mundial de la Salud afirma que esta es la principal causa de muerte entre los jóvenes de 15 a 29 años, y que las lesiones derivadas de estos hechos representan la octava razón de muerte a nivel global.

Adicionalmente, se conoce que el 90% de las defunciones relacionadas con accidentes viales se producen en los países de ingresos bajos y medianos, grupo en el que se encuentra Argentina.

Si bien en muchas ocasiones estos no pueden evitarse, existen múltiples acciones que los gobernantes y los propios usuarios de la vía pública, pueden llevar adelante para intentar prevenirlos y/o reducirlos.

En el presente trabajo se abordó esta temática, por considerarse de relevancia a toda la población y donde, especialmente, contar con datos y tableros de control limpios y exactos, ayudará a conocer los fenómenos relaciones para poder tomar las decisiones correctas. La Ciudad Autónoma de Buenos Aires sufre diariamente las consecuencias de estos sucesos, por lo que el análisis estará enfocado precisamente allí.

2. Descripción de la temática de los datos

Para el presente trabajo, se extrajo información de los siniestros viales en la Ciudad Autónoma de Buenos Aires (CABA). Los datos se extienden en el periodo 2015 - 2018.

3. Alcance y usuario

El dashboard diseñado se encuentra orientado a ser utilizado por un nivel táctico y/o estratégico, es decir, por los mandos medios o líderes del ministerio de transporte y vialidad de la CABA.

Su utilización, les permitirá comprender las características generales de estos acontecimientos y a quienes afectan. En el tablero de control, podrían identificar las comunas, las calles y los horarios más sensibles a los accidentes.

La información relevada puede ser de utilidad para definir políticas públicas de prevención de estos acontecimientos, ya sea en infraestructura, señalización y control, como en educación. Contar con información detallada y limpia ayuda a mejorar la eficacia de las acciones que se puedan implementar.

4. Objetivo

Nuestro objetivo es comprender las principales dimensiones que caracterizan a los siniestros viales y sus víctimas en la CABA. Para esto, se verán la evolución de los casos a lo largo del periodo bajo análisis, geolocalizaremos los siniestros para buscar puntos y horarios específicos en los que se den con mayor frecuencia. Además, intentaremos definir las características demográficas de las víctimas y los medios de transporte más frecuentemente utilizados.

Lograr una correcta descripción de estos aspectos, podría ayudar a nuestros usuarios a realizar planes de acción concretos, para prevenir este tipo de siniestros que son una de las causas de muertes más importantes en la ciudad.

5. Herramientas tecnológicas implementadas

En el desarrollo del presente proyecto, se utilizaron los siguientes programas:

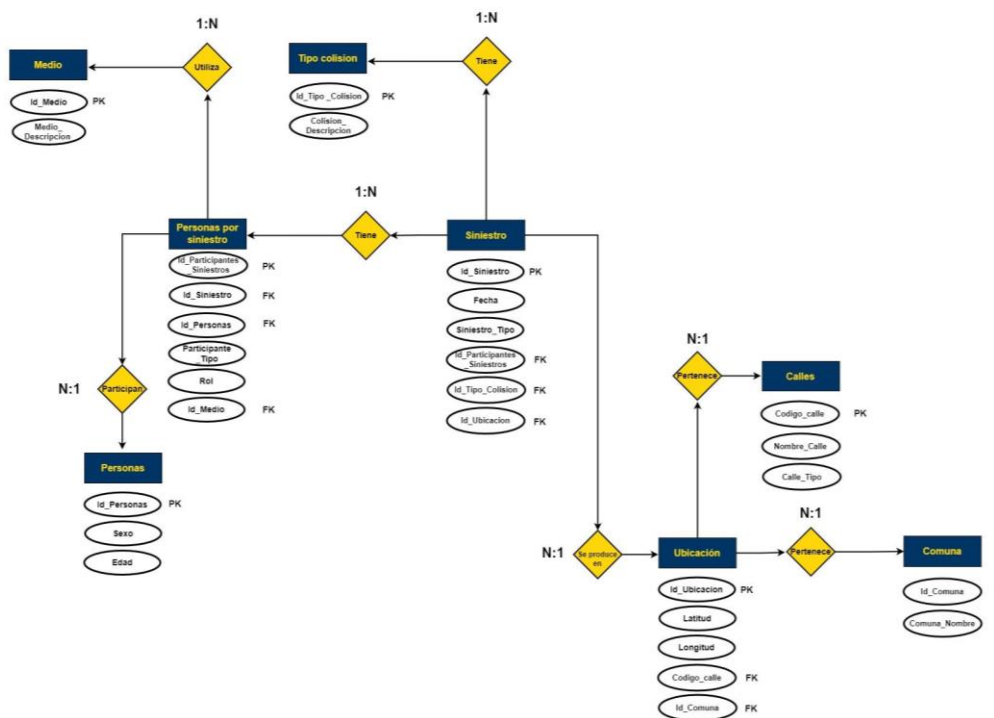
- Google Sheets para la lectura y limpieza del dataset.
- Google Slides para la creación del diseño del mockup.
- Google Doc para documentación y organización del equipo.
- Canvas para estructuración del informe.
- Diagrams Net para la creación del diagrama entidad-relación (<https://app.diagrams.net/>).
- Power BI Desktop para la creación del tablero de control.

6. Dataset

A continuación, se adjunta el link al archivo xlsx que contiene la base de datos que se utilizó para el análisis: [SiniestrosVialesEnCaba.xlsx](#)

7. Diagrama Entidad-Relación (DER)

A continuación, se presenta el gráfico del diagrama de entidad - relación a inicios del proyecto:



8. Listado de tablas

A continuación, se menciona cada tabla junto a una breve descripción de estas:

1. Medio: contiene la información de los distintos medios con los que puede circular una persona en la vía pública. Esta es una tabla maestro, ya que solo se limita a listar los medios posibles.

2. Tipo colisión: también es una tabla maestro y contiene todas las combinaciones de siniestros entre distintos medios de transporte, es decir el tipo de choque o colisión.
3. Comuna: contiene el listado de comunas de la Ciudad Autónoma de Buenos Aires, también es una tabla maestro.
4. Calle: informa el listado de calles de la Ciudad Autónoma de Buenos Aires y una clasificación sobre el tipo de calle considerada.
5. Ubicación: informa puntos de geolocalización en la Ciudad Autónoma de Buenos Aires donde ocurrió algún siniestro vial.
6. Personas: contiene información sobre las personas, su la edad y el sexo, entendemos que este registro en un maestro y no se limita solo a quienes participaron de siniestros como víctimas sino a todas las personas que participaron ya de alguna manera.
7. Personas por siniestro: es una tabla intermedia que nos permitió relacionar "Personas" y "Siniestros". Debimos implementarla, ya que una misma persona puede participar en más de un siniestro y teniendo distintos roles (víctima / victimario).
8. Siniestro: contiene el registro de cada accidente vial ocurrido en la Ciudad Autónoma de Buenos Aires y los datos de la víctima. Todos los siniestros tienen dato de una sola víctima, por lo que suponemos que es la principal.

9. Listado de columnas por tabla

A continuación, se mencionará las columnas que posee cada tabla junto con su tipo de campo y clave:

Tabla 1: MEDIO		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Medio	nvarchar(100)
	Medio_Descripcion	nvarchar(100)

Tabla 2: PERSONAS		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Personas	nvarchar(100)
	Sexo	nvarchar(100)
	Edad	int

Tabla 3: PERSONAS POR SINIESTRO		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Participantes_Siniestro	nvarchar(100)
FK	Id_Siniestro	nvarchar(100)
FK	Id_Personas	nvarchar(100)
	Participante_Tipo	nvarchar(100)
	Rol	nvarchar(100)
FK	Id_Medio	nvarchar(100)

Tabla 4: SINIESTRO		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Siniestro	nvarchar(100)
	Fecha	datetime
	Victimas_Cantidad	nvarchar(100)
	Siniestro_Tipo	nvarchar(100)
FK	ID_Participantes_Siniestro	nvarchar(100)
FK	Id_Tipo_Colision	nvarchar(100)
FK	Id_Ubicacion	nvarchar(100)

Tabla 5: UBICACION		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Ubicacion	nvarchar(100)
	Latitud (X)	float(18,3)
	Longitud (Y)	float(18,3)
FK	Codigo_Calle	nvarchar(100)
FK	Id_Comuna	nvarchar(100)

Tabla 6: COMUNA		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Comuna	nvarchar(100)
	Comuna_Nombre	nvarchar(100)

Tabla 7: CALLE		
Tipo de clave	Campo	Tipo de campo
PK - index	Codigo_Calle	nvarchar(100)
	Nombre_Calle	nvarchar(100)
	Calle_Tipo	nvarchar(100)

Tabla 8: TIPO COLISION		
Tipo de clave	Campo	Tipo de campo
PK - index	Id_Tipo_Colision	nvarchar(100)
	Colision_Descripcion	nvarchar(100)

10. Modelo relacional en Power BI

Una vez que los archivos planos fueron subidos a Power BI, se hicieron los siguientes cambios en los datos:

10.1. Eliminación de PK Null

En un primer momento se habían contemplado en las columnas de ID de todas las tablas un valor NULL, para aquellos registros de esa tabla que venían NULL. Pero esta interpretación no es correcta ya que una PK no puede ser NULL. Por lo tanto, se eliminaron estos registros en cada tabla.

10.2. Cambio en el tipo de dato

- La columna Id_Comuna de la tabla Comuna se modifica de tipo de dato número entero a texto.
- La columna Id_Medio de la tabla Medio se modifica de tipo de dato número entero a texto.

- La columna Id_Tipo_Colision de la tabla Tipo Colisión se modifica de tipo de dato número entero a texto.
- La columna Codigo_Calle de la tabla Calle se modifica de tipo de dato número entero a texto.
- La columna Id_Ubicación de la tabla Ubicación se modifica de tipo de dato número entero a texto.
- La columna Id_Personas de la tabla Personas se modifica de tipo de dato número entero a texto.
- La columna Id_Participantes_Siniestro de la tabla Personas por Siniestro se modifica de tipo de dato número entero a texto.
- La columna Id_Siniestro de la tabla Siniestros se modifica de tipo de dato número entero a texto.
- La columna Participante_Tipo de la tabla Personas por Siniestro se eliminó, ya que todos los valores contienen el valor "1", es decir que todos son víctimas.
- La columna Codigo_Calle_1 de la tabla Ubicación se modifica de tipo de dato número entero a texto.
- La columna Victimas_Cantidad de la tabla Siniestros se modifica el tipo de dato de texto a número entero.

10.3. Renombre de campos

- Se renombró la columna Nombre_Calle por Calle de la tabla Calle.
- Se renombró la columna Tipo_Calle por Tipo de calle de la tabla Calle.
- Se renombró la columna Comuna_Nombre por Comuna de la tabla Comuna.
- Se renombró la columna Medio_Descripcion por Medio de transporte de la tabla Medios.
- Se renombró la columna Siniestro_Tipo por Tipo de siniestro de la tabla Siniestro.
- Se renombró la columna Colisión_Descripcion por Tipo de colisión de la tabla Tipo Colisión.

10.4. Reemplazo de valores

Se reemplazó el valor "desconocida" en la columna Edad de la tabla Personas por un vacío. Ya que esta es una variable en la que podemos calcular promedio de edad u otras métricas y para ello necesitamos que el tipo de dato sea numérico entero. Por lo tanto, también se transforma el tipo de dato en número entero.

10.5. Creación de columnas

Se creó la columna Rango Etario en la tabla Personas en base a una subclasificación de las edades. Esta columna nos va a permitir resumir en qué rango etario se encuentra la mayor cantidad de víctimas de siniestros viales.

```
Rango etario = if(and(Personas[Edad] >= 1, Personas[Edad] < 10) , "1-9 años",  
if(and(Personas[Edad] >= 10, Personas[Edad] < 20) , "10-19 años",  
if(and(Personas[Edad] >= 20, Personas[Edad] < 30) , "20-29 años",  
if(and(Personas[Edad] >= 30, Personas[Edad] < 40) , "30-39 años",  
if(and(Personas[Edad] >= 40, Personas[Edad] < 50) , "40-49 años",  
if(and(Personas[Edad] >= 50, Personas[Edad] < 60) , "50-59 años",  
if(and(Personas[Edad] >= 60, Personas[Edad] < 70) , "60-69 años",  
if(and(Personas[Edad] >= 70, Personas[Edad] < 80) , "70-79 años",  
if(Personas[Edad] >= 80, "80 o más años"))))))))
```

Desde la columna Fecha de la tabla Siniestro, agregamos cuatro columnas extrayendo los distintos datos temporales:

```
Año = YEAR(Siniestro[Fecha])  
Mes = FORMAT(Siniestro[Fecha], "MMMM")  
Dia = FORMAT(Siniestro[Fecha], "DDDD")  
Hora del siniestro = HOUR(Siniestro[Fecha])
```

10.6. Eliminación de columnas

Se eliminó la columna Participante_Tipo de la tabla Personas por Siniestro dado que la misma contiene solo el valor 1 indicando que la persona es una víctima y no un acusado del accidente, caso en el que tendría 0. El motivo de esta acción se debe a que es una variable que no utilizamos en el análisis.

10.7. Tabla Calendario

Se creó una nueva columna Fecha sin hora en la tabla Siniestros, en base a la cual se va a construir la tabla calendario. Se realizó de esta manera, porque al querer generar una medida calculada con funciones de inteligencia de tiempo con la tabla calendario, toma como correctos solo aquellos registros que tienen hora en punto.

La columna se hizo duplicando la de Fecha y haciendo un split de dd/mm/yyyy con hh:mm:ss, quedándonos con la primera parte y desechando la segunda.

Tabla calendario = `CALENDAR(MIN(Siniestro[Fecha sin hora]), MAX(Siniestro[Fecha sin hora]))`

En esta tabla se crearon tres columnas:

Año = `YEAR('Tabla calendario'[Date])`

Mes = `FORMAT('Tabla calendario'[Date], "MMMM")`

Día de la semana = `FORMAT('Tabla calendario'[Date], "DDDD")`

Mes en numero = `MONTH('Tabla calendario'[Date])`

10.8. Tabla para medidas

Se creó una tabla denominada Medidas calculadas, en la que agrupamos todas las medidas que fuimos creando en el proceso. Esto nos permitirá tener el trabajo más organizado.

11. Medidas calculadas

1. Cantidad de siniestros en base a la tabla Siniestros: esta medida nos permitirá saber la cantidad de siniestros que ocurrieron en total y teniendo en cuenta distintos filtros que agregaremos al tablero. Por ejemplo: por comuna, año, hora, entre otros.

Cantidad de siniestros = `COUNT(Siniestro[Id_Siniestro])`

2. Edad promedio de las víctimas en base a la tabla Personas: esta medida permitirá identificar qué sector de la población se concentra la mayor cantidad de víctimas.

Edad Promedio = `AVERAGE(Personas[Edad])`

3. Variación de siniestros año a año en base a la tabla Siniestros:

Variación de siniestros respecto el año anterior =

```
var A = Calculate(DISTINCTCOUNT(Siniestro[Id_Siniestro]),DATEADD('Tabla calendario'[Date], -1, YEAR))
var B = Calculate(DISTINCTCOUNT(Siniestro[Id_Siniestro]),DATEADD('Tabla calendario'[Date], 0, YEAR))
return B- A
```

4. Variación de siniestros respecto del año anterior en base a la tabla Siniestros en porcentajes:

Siniestros año a año =

```
VAR __PREV_YEAR =
CALCULATE( COUNTA('Siniestro'[Id_Siniestro]),
DATEADD('Tabla calendario'[Date], -1, YEAR) )
RETURN
DIVIDE(COUNTA('Siniestro'[Id_Siniestro]) - __PREV_YEAR, __PREV_YEAR)
```

5. Cantidad de Víctimas discriminadas por género: con estas medidas se busca poder identificar como se dividen las victimas según el sexo, es decir si los siniestros viales afectan a uno más que al otro y de qué manera.

Cantidad de víctimas masculinas = `var D = CALCULATE(COUNT(Personas[Sexo]), Personas[Sexo] = "masculino")`
`return D`

Cantidad de víctimas femeninas = `CALCULATE(COUNT(Personas[Sexo]), Personas[Sexo] = "femenino")`

6. Porcentajes por tipo de rol de la víctima: se crearon estas medidas para determinar cuáles son los roles que tuvieron las víctimas en los accidentes. Si bien tenemos el medio de transporte, puede que en un automóvil la víctima haya sido acompañante o conductor.

Porcentaje de ciclistas = `var A = CALCULATE(COUNT(Siniestro[Id_Siniestro]), 'Personas por siniestro'[Rol] = "ciclista")`
`var B = COUNT(Siniestro[Id_Siniestro])`
`return A/B`

Porcentaje de conductores = `var A = CALCULATE(COUNT(Siniestro[Id_Siniestro]), 'Personas por siniestro'[Rol] = "conductor")`
`var B = COUNT(Siniestro[Id_Siniestro])`
`return A/B`

Porcentaje de pasajeros = `var A = CALCULATE(COUNT(Siniestro[Id_Siniestro]), 'Personas por siniestro'[Rol] = "pasajero")`
`var B = COUNT(Siniestro[Id_Siniestro])`
`return A/B`

Porcentaje de peatones = `var A = CALCULATE(COUNT(Siniestro[Id_Siniestro]), 'Personas por siniestro'[Rol] = "peaton")`
`var B = COUNT(Siniestro[Id_Siniestro])`
`return A/B`

A lo largo del proyecto, realizamos distintas modificaciones con el fin de cumplir con las distintas entregas, pero que no tienen relevancia para nuestro análisis y que no utilizamos. A continuación, detallamos:

Creamos la columna Medios de transporte nueva clasificación en la tabla Medios, la misma se realizó a partir de los 22 tipos de medios de transporte de la columna Medio de transporte:

Medios de transporte nueva clasificación = `if(or(Medio[Medio de transporte] = "autobomba",
or(Medio[Medio de transporte] = "camion",
or(Medio[Medio de transporte] = "traccion a sangre",
or(Medio[Medio de transporte] = "transporte de cargas / grua",
or(Medio[Medio de transporte] = "tranvia",
or(Medio[Medio de transporte] = "utilitario",
Medio[Medio de transporte] = "vehiculo rural"))))))),
"Transporte de carga",
if(or(Medio[Medio de transporte] = "ambulancia",
or(Medio[Medio de transporte] = "automovil",
or(Medio[Medio de transporte] = "bicicleta",
or(Medio[Medio de transporte] = "camioneta",
or(Medio[Medio de transporte] = "cuatriciclo",
or(Medio[Medio de transporte] = "fuerza seguridad",
or(Medio[Medio de transporte] = "moto",
or(Medio[Medio de transporte] = "omnibus / minibus",`

```

or(Medio[Medio de transporte] = "patineta",
or(Medio[Medio de transporte] = "peaton",
or(Medio[Medio de transporte] = "taxi/remis",
or(Medio[Medio de transporte] = "transporte escolar",
Medio[Medio de transporte] = "tren / subte")))))))))))",
"Transporte de pasajeros", "Otros"))

```

Parámetro

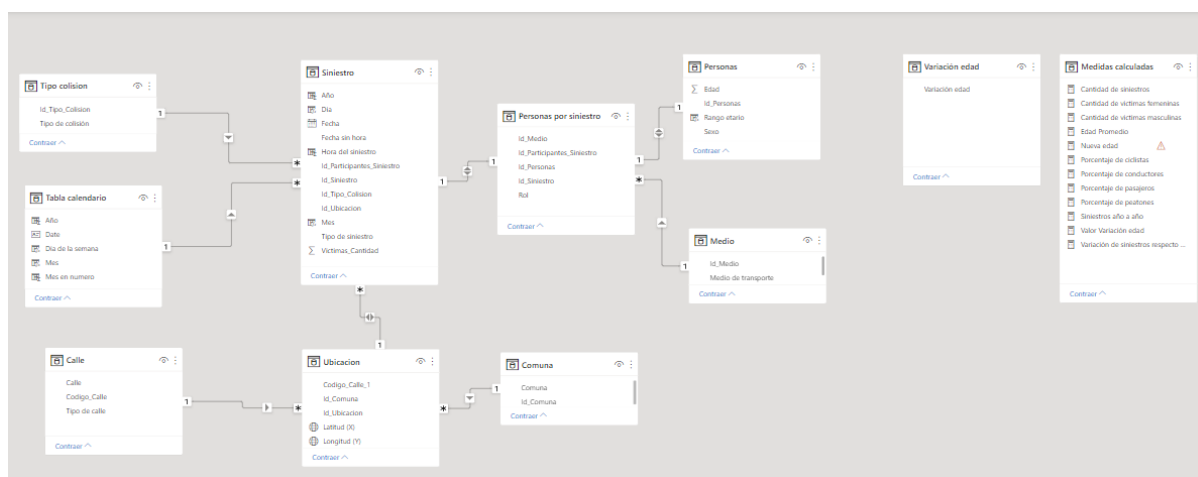
Variación edad = `GENERATESERIES(1, 20, 1)`

Medida calculada en base al parámetro:

Nueva edad = `AVERAGE(Personas[Edad]) + 'Variación edad'[Valor Variación edad]`

12. DER Power BI

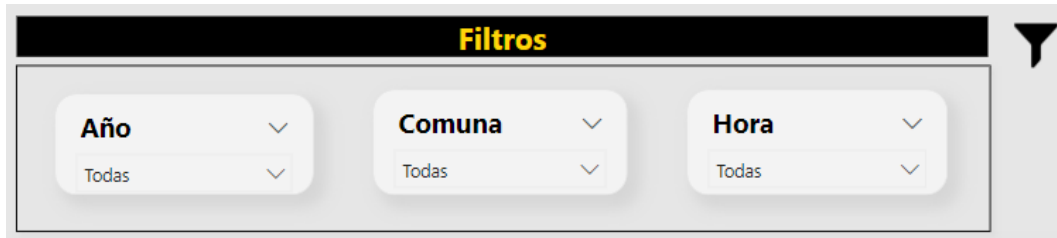
Por último, adjuntamos imagen del diagrama entidad-relación graficado en Power BI luego de las transformaciones y las medidas que incorporamos, como se puede observar a continuación:



13. Segmentaciones elegidas

En el tablero realizado, se utilizaron 2 tipos de segmentaciones:

1. Desplegables: utilizados para filtrar año, hora y comuna, que son los tres filtros principales. El botón incorporado en el margen superior derecho permite borrar todos los filtros seleccionados en la página.



2. Botones interactivos: utilizados para filtrar por sexo. Se diseñaron de la siguiente manera:



14. Paleta de colores elegida

Haciendo uso de una plataforma online¹, logramos una paleta de colores que se adaptara a nuestro proyecto, la misma puede visualizarse a continuación:

FFC300	FFCC38	F5D65B	013565	0A3C65	00284C	000000	BBBBBB	E7E7E7	FDFDFD
Mikado Yellow	Sunglow	Naples Yellow	Prussian Blue	Indigo Dye	Oxford Blue	Black	Gray X 11 Gray	Platinum	White

Luego, por medio de un archivo JSON importamos a Power BI de la siguiente manera:

```
{  "name": "Siniestros",  "dataColors": ["#FFD60A", "#FFCC38", "#F5D65B", "#013565", "#0A3C65", "#00284C", "#000000", "#BBBBBB", "#E7E7E7", "#FDFDFD"],  "background": "#FFFFFF",}
```

¹ <https://coolors.co/>

```
"foreground": "#3A6108",  
"tableAccent": "#568410" }
```

15. Visualización de los datos

La visualización del tablero partió de un mockup que luego fue llevado a Power BI. A continuación, se detalla un breve resumen de cada una de las solapas creadas:

15.1. Portada

En el mockup esta página fue incluida de manera muy simple como se muestra a continuación:

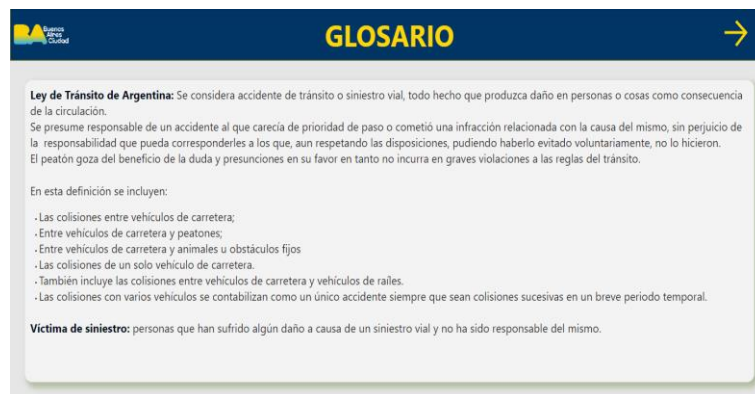


Ya en Power BI, consideramos importante incluir otros elementos como presentación e introducción a la temática, como el título, una imagen y un logo identificando al usuario final. Además, contiene botones interactivos que dirigen a cada una de las demás páginas. Adicionalmente, se hace referencia a la última actualización que tuvo el tablero y sus autores.



15.2. Glosario

Esta solapa contiene dos conceptos importantes a tener en cuenta para cualquier usuario del tablero, uno es que se entiende por siniestro vial o accidente de tránsito según la ley argentina y a quien se considera víctima.



15.3. General

Inicialmente en el mockup está solapa fue diseñada de la siguiente manera:



El producto final realizado en Power Bia mantiene la esencia, pero algunas grafías fueron modificadas de acuerdo lo que descubrimos con los datos reales y en el ir conociendo la herramienta.

En esta solapa se pretende mostrar información genérica de los siniestros analizados y con especial relevancia la ubicación temporal de los mismos, tanto en los años y meses como en los distintos momentos del día (horas) en los que suceden.



En cuanto a la información y como se visualiza, creamos un KPI que nos permite saber la cantidad de siniestros totales que se analizaran (Cantidad de siniestros). Luego, incorporamos cuatro gráficos:

- Siniestros por hora intenta determinar la existencia de una estacionalidad en el día en que ocurra una mayor cantidad de hechos al mismo, el filtro de hora no le es aplicable.
- Top tres de tipo de coalición gráfico de barras que muestra que medio de transporte colisionaron y el mismo puede ir variando en los distintos periodos.
- Evolución de los siniestros en el tiempo (años) permite ver la cantidad de siniestros a través de los años analizados.
- Evolución de los siniestros en el tiempo (meses) permite ver la cantidad de siniestros a través de los meses analizados.

Además, si incorporo una grilla en la que se puede observar la variación de siniestros año a año en cantidades, se puede observar una disminución de datos en el año 2018, quizás se debe a una falta de actualización de estos.

Por último, se incorporaron los tres filtros que mencionamos anteriormente y un botón en el margen superior derecho que dirige al usuario a la siguiente solapa. Estos elementos están presentes en todas las páginas que siguen.

15.4. Ubicación

En el diseño del tablero se planteó esta solapa con el objetivo de mostrar toda la información referida a la dimensión espacial de los incidentes, desde lo más genérico a lo más específico: comuna, tipo de calle, calle y el punto exacto de los incidentes:



Luego, para una mejor visualización de los gráficos se decidió incorporar dividir esta página en dos donde una solo este el mapa:



Se incorporaron los mismos filtros que en la portada general y dos gráficos:

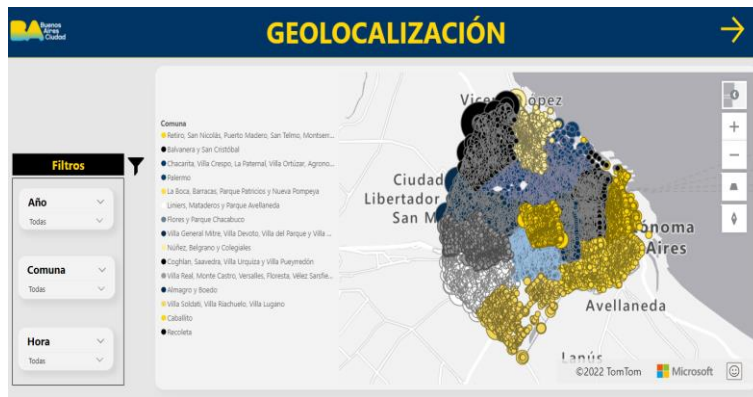
- Siniestros por comuna es un treemap, el mismo permite a simple vista identificar en qué comuna porteña son más frecuentes los accidentes viales y además, mediante la incorporación de un Tooltips, cual es el Top cinco de calles en las que ocurren. A este gráfico no le es aplicable el filtro de Comuna. Para el Tooltip creamos una sexta solapa que no será visualizada por el usuario.
- ¿Dónde ocurren? Gráfico de jerarquía en el que se muestra el tipo de calle en la que sucedieron y el nombre específico de la mismas.

Por último, como en las demás solapas los tres filtros y el botón en el margen superior derecho para pasar a la siguiente página.

15.5. Geolocalización

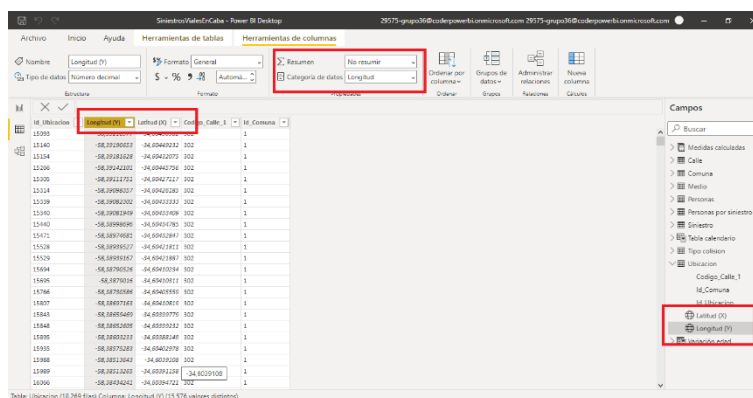
Como se mencionó en párrafos anteriores, esta solapa es un desprendimiento de Ubicación y solo contiene como visualización un mapa de la ciudad donde se pueden observar los puntos geolocalizados donde sucedieron los siniestros.

Cada comuna está identificada con un color distinto y el tamaño de los puntos indican la cantidad de accidentes que sucedieron en el lugar.



Añadimos a continuación capturas de pantalla que muestran la correcta configuración de Power BI para poder visualizar el mapa:

Es necesario que en las tablas que componen el DER los datos de Latitud y Longitud sean seteados como datos a No Resumir y con su correspondiente categoría tal como se observa en la imagen inferior.

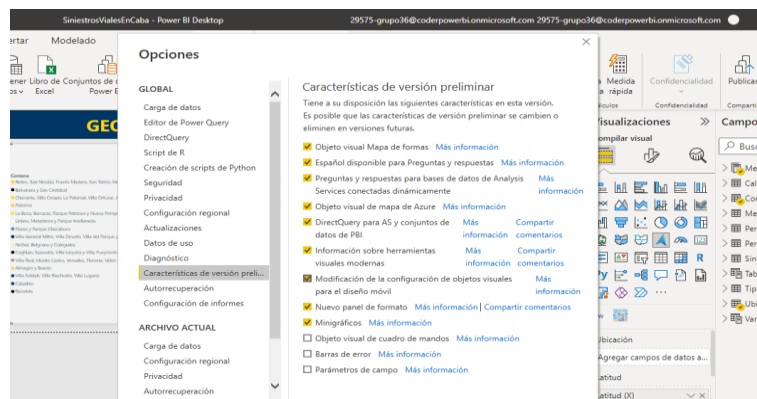


Además de esto, hay que saber que los valores de longitud/latitud pueden llegar a estar invertidos en el dataset, como paso en nuestro caso, esto se puede resolver simplemente asignando la categoría correcta a cada columna.

Las visualizaciones de mapas que vienen por defecto en PBI para poder ser activadas requieren de habilitar manualmente algunas *Características de Versión Preliminar* dentro de Opciones y configuración:



La más importante en este caso es la de Objeto Visual de Mapa de Azure:



15.6. Víctimas

La última solapa tiene por objetivo caracterizar a las víctimas de los siniestros. El primer diseño del mockup fue el siguiente:



El producto final en Power BI cuenta con dos gráficos, 7 KPI's y tres marcadores mencionados en la segmentación.



El primer gráfico Cantidad por rango etario muestra la pirámide poblacional de las víctimas de siniestros, permitiéndonos identificar segmentos específicos. Este gráfico es complementado con el KPI de edad promedio de la víctima.

El segundo gráfico corresponde al medio de transporte de estas. Cuatro KPI's complementan a este segundo gráfico permitiéndonos reconocer el rol que cumplía en el medio de transporte.

Por último, dos KPI que identifican la cantidad por sexo. Los marcadores que incorporamos intentan mostrar toda la solapa dividida en el sexo y una global, los filtros y el botón de transición que dirige al usuario a la portada.

16. Conclusión

Con los datos analizados en nuestro trabajo podemos concluir que los siniestros se dan con mayor frecuencia en la franja horaria comprendida entre las 12 del mediodía y las 18. Es decir, existe una estacionalidad en la hora y las causas de esta podrían ser analizadas en el futuro con información complementaria, por ejemplo, sobre los horarios de ingreso y egreso laboral, almuerzos, etc.

Por otro lado, se puede observar que la cantidad de accidentes interanual se mantuvo relativamente constante hasta 2017, en 2018 se observa una importante disminución. Esta disminución, creemos puede deberse a una fuente de datos incompleta para ese periodo.

Si se toman medidas de acción concretas para reducir los siniestros viales y el tablero se complementa con la información de los siguientes periodos, deberíamos observar una tendencia decreciente.

Si comparamos la evolución anual por cada comuna, observamos que no todas tienen un comportamiento constante. Por ejemplo, en la comuna Villa Real, Monte Castro, Versalles, Floresta, Vélez Sarsfield y Villa Luro, los eventos fueron en aumento de periodo a periodo, caso que amerita un análisis específico para identificar las causas.

En el top tres de tipo de colisión en todos los años, comunas y horas aparece en primer lugar Moto-Automóvil, rotando en segundo y tercer puesto Automóvil-Automóvil y Peatón-Automóvil dependiendo el filtro. Para la dimensión mensual, no encontramos un patrón de comportamiento uniforme, cada año y cada comuna varía de distintas formas.

La comuna con mayor cantidad de accidentes es Retiro, San Nicolás, Puerto Madero, San Telmo, Montserrat y Constitución y , si bien se mantiene primera en todos los periodos, los casos han ido disminuyendo periodo a periodo en la misma.

Por último, observamos que la víctima promedio es masculina de 36 años. Las víctimas masculinas duplican a las femeninas, se debería reforzar la educación vial en este sector de la población.

En resumen, existen múltiples variables que caracterizan a los siniestros y desde las cuales se puede atacar su prevención y eso es lo que intentamos mostrar en nuestro tablero.

17. Futuras líneas

En este apartado, se consideraron ciertos puntos que no fueron incluidos en el presente trabajo pero que se podrían añadir a futuro:

- Profundizar el análisis de la hipótesis planteada, con una actualización de la información de los siniestros viales a la fecha. Es importante mantener los datos actualizados para dar un seguimiento real.
- Complementar los tableros con información sobre el impacto de las acciones llevadas a cabo por el gobierno de la ciudad para intentar prevenir. Por ejemplo, zonas de mejoras en la señalización, controles, entre otros. De esta manera se podrá identificar las variaciones y si las medidas fueron correctas o si es necesario implementar otras.

- Incluir datos sobre los responsables de estos siniestros, si se encontraban bajo el efecto de alguna sustancia o no, permitiría identificar la negligencia y responsabilidad.
- También, encontrar patrones desde la geolocalización de puntos donde se den con mayor frecuencia y que no existan puntos de control de tránsito.

18. Bibliografía

[Buenos Aires Data - Datasets](#)