

# **Proyecto Final Business Intelligence: Power BI**

## **Accidentes según causa conductor gestionados por la Guàrdia Urbana a la ciutat de Barcelona en el año 2022**

**Francisco J. Regalado López**

### **Introducción**

La recopilación y análisis de datos sobre la ubicación y las causas de los accidentes en una ciudad desempeña un papel fundamental en la mejora de la planificación urbana y la seguridad vial. Conocer dónde han ocurrido los accidentes y comprender sus causas subyacentes proporciona información valiosa para implementar medidas preventivas y estrategias eficaces de gestión del tráfico.

Al examinar la distribución espacial de los accidentes, se pueden identificar áreas de alta concentración de incidentes, lo que permite a los responsables de la planificación enfocar sus recursos y acciones en aquellos lugares que presentan mayor riesgo. Además, comprender las causas principales de los accidentes, como el exceso de velocidad o el comportamiento imprudente de los conductores, brinda la oportunidad de abordar esos problemas específicos mediante la implementación de políticas y medidas adecuadas.

La recopilación y análisis de datos sobre accidentes también permiten evaluar la eficacia de las intervenciones realizadas, monitorear las tendencias a lo largo del tiempo y realizar comparaciones entre diferentes áreas geográficas. En última instancia, el objetivo es lograr una mejor planificación urbana que promueva la seguridad vial, reduzca los accidentes y mejore la calidad de vida de los ciudadanos.

### Descripción de los datos

- Nombre del conjunto de datos: Accidentes según causa conductor gestionados por la Guàrdia Urbana a la ciutat de Barcelona
- Año recurso: 2022
- Fuente: Ajuntament de Barcelona
- Fecha de publicación: 14/02/2023
- URL datos: [https://opendata-ajuntament.barcelona.cat/data/es/dataset/accidents\\_causa\\_conductor\\_gu\\_bc\\_n](https://opendata-ajuntament.barcelona.cat/data/es/dataset/accidents_causa_conductor_gu_bc_n)

El conjunto de datos titulado proporciona información detallada sobre los accidentes en la ciudad de Barcelona y las causas atribuidas a los conductores involucrados. La tabla contiene El conjunto de datos consta de 20 columnas y 9320 filas.

#	Columna	Tipo de dato	Descripción
1	Numero_expedient	object	Número d'expedient
2	Codi_districte	int64	Codi del districte
3	Nom_districte	object	Nom del districte
4	Codi_barri	int64	Nom del barri
5	Nom_barri	object	Nom del barri
6	Codi_carrer	int64	Codi del carrer
7	Nom_carrer	object	Nom_carrer
8	Num_postal	object	Número postal
9	Descripcio_dia_setmana	object	Nom del dia de la setmana
10	NK_Any	int64	Any
11	Mes_any	int64	Mes de l'any
12	Nom_mes	object	Nom del mes
13	Dia_mes	int64	Dia del mes
14	Hora_dia	int64	Hora del dia
15	Descripcio_causa_mediata	object	Descripció causa mediata
16	Descripcio_torn	object	Descripció del torn
17	Coordenada_UTM_X_ED50	float64	Coordenada X en format UTM (ED50)
18	Coordenada_UTM_Y_ED50	float64	Coordenada Y en format UTM (ED50)
19	Longitud	float64	Longitud
20	Latitud	Float64	Latitud

### Elección de la herramienta: Power BI

Power BI es una elección popular para la visualización de datos debido a varias razones clave. En primer lugar, ofrece una interfaz intuitiva y fácil de usar, lo que permite a los usuarios crear visualizaciones atractivas y dinámicas sin necesidad de conocimientos técnicos avanzados. Además, Power BI permite la conexión y combinación de datos de múltiples fuentes, lo que facilita el análisis completo y enriquecido.

## **Proyecto Final – Business Intelligence**

Otra ventaja es la amplia gama de visualizaciones y gráficos personalizables que ofrece Power BI, lo que permite presentar los datos de manera efectiva y comprensible. Además, cuenta con funciones de interactividad, filtrado y exploración, lo que brinda a los usuarios la capacidad de profundizar en los datos y descubrir insights relevantes.

En general, elegir Power BI para la visualización de datos brinda una solución potente, flexible y accesible para convertir datos en información significativa y accionable.

### **Preparación de los datos para el análisis**

Al cargar un conjunto de datos en Power BI, pueden surgir varios problemas que afecten la calidad y precisión de los resultados. Algunos de estos problemas incluyen la falta de limpieza y transformación de los datos antes de cargarlos en Power BI, lo que puede llevar a visualizaciones incorrectas o engañosas. Además, es posible que los datos no estén correctamente modelados o relacionados, lo que dificulta la creación de análisis y visualizaciones coherentes.

Es fundamental abordar estos problemas mediante una adecuada preparación de los datos, asegurando su calidad, coherencia y estructura adecuada antes de la carga en Power BI. Esto garantizará que las visualizaciones y análisis sean precisos, confiables y útiles para la toma de decisiones basadas en datos.

Al cargar el conjunto de datos hubo algunos tipos de datos erróneos que han de cambiarse y también existen algunas transformaciones que hay que realizar para tener los datos estructurados de la mejor manera posible.

DAX (Data Analysis Expressions) es un lenguaje de fórmulas utilizado en Power BI y otras herramientas de análisis de datos de Microsoft. Permite realizar transformaciones y cálculos avanzados en los datos para obtener información relevante.

El propósito de utilizar fórmulas tipo DAX en Power BI es realizar cálculos avanzados y transformaciones en los datos, permitiendo obtener métricas y análisis personalizados. DAX proporciona funciones para realizar operaciones condicionales, agregaciones, cálculos de tiempo y porcentaje, entre otros. Estas fórmulas ayudan a crear medidas y columnas calculadas que permiten obtener información relevante y realizar análisis más profundos en los informes y paneles de Power BI.

## Proyecto Final – Business Intelligence

#	Columna	Descripción carga del dato
1	Numero_expedient	Ok
2	Codi_districte	Ok
3	Nom_districte	Texto a Ubicación
4	Codi_barri	Ok
5	Nom_barri	Texto a Ubicación
6	Codi_carrer	Ok
7	Nom_carrer	Ok
8	Num_postal	Ok
9	Descripcio_dia_setmana	Ok
10	NK_Any	Ok
11	Mes_any	Ok
12	Nom_mes	Ok
13	Dia_mes	Ok
14	Hora_dia	Numero entero a Hora
15	Descripcio_causa_mediata	Ok
16	Descripcio_torn	Ok
17	Coordenada_UTM_X_ED50	Ok
18	Coordenada_UTM_Y_ED50	Ok
19	Longitud	Número decimal a Longitud
20	Latitud	Número decimal a Latitud

### Cambio de tipo de dato

El cambio de datos para las columnas de ubicación es simplemente clasificar los valores al tipo de ubicación, desde la Herramienta de columnas > categoría de datos. Los valores de texto pueden clasificarse a ubicaciones como ciudad, dirección y lugar, los valores de número decimal pueden clasificarse como longitudes o latitudes.

### Transformación de datos

#### Hora

La columna Hora\_dia es un número entero, pero no tiene el formato correcto (HH), en este caso se utilizará una formula DAX para la transformación del dato a datetime, la formula Time recibe tres valores (H, M, S). El valor Hora tiene un rango de 1 a 24, cualquier valor mayor a 23 se dividirá entre 24 y el resto se tratará como el valor de hora, representado como una fracción de un día. En el conjunto de datos los valores tienen un rango no pasan de 24. Se crea una columna y se aplica la función.

El formato nos devuelve una fecha aleatoria para todos los valores, pero con la hora correcta.

## Proyecto Final – Business Intelligence

	Descripción_causa_mediate	Descripción_torn	Coordenada_UTM_X_ED50	Coordenada_UTM_Y_ED50	Longitud	Latitud	Hora
7	No respetar distancias	Mati	430080.36	4582095.48	2.16261527843894	41.3855046566338	30/12/1899 7:00:00
11	Altres	Mati	430062.55	4582005.38	2.16241269697104	41.3846916052752	30/12/1899 11:00:00
11	Canvi de canil sense precaució	Mati	429629.38	4581777.7	2.15965062520641	41.3826264020394	30/12/1899 11:00:00
12	No respetar distancias	Mati	430035.23	4582321.24	2.16623538089505	41.387564496894	30/12/1899 12:00:00
9	No respetar distancias	Mati	430307.64	4582245.3	2.1653160888093	41.386873776318	30/12/1899 9:00:00
11	Desconegut	Mati	430299.85	4582237.63	2.16522380978352	41.3868040208313	30/12/1899 11:00:00
15	Altres	Tarda	431314.18	4582320.39	2.17724185676572	41.3958328428663	30/12/1899 15:00:00
16	No respetar distancias	Tarda	430934.78	4582858.79	2.17274626758835	41.39245343302	30/12/1899 16:00:00
18	No determinada	Tarda	430418.29	4582353.68	2.16662693156219	41.3878595095211	30/12/1899 18:00:00
19	Manca d'atenció a la conducció	Tarda	431024.51	4582946.76	2.17380944523285	41.3932534932386	30/12/1899 19:00:00
21	No respetar distancias	Tarda	431112.06	4583023.89	2.17484780884279	41.3939556563002	30/12/1899 20:00:00
21	No respetar distancias	Tarda	431690.71	4582793.42	2.1840742573486	41.4009529703384	30/12/1899 21:00:00
11	Canvi de canil sense precaució	Mati	429731.19	4581680.88	2.1584876352699	41.3817400392224	30/12/1899 11:00:00
20	No respetar distancias	Nit	429175.13	4581209.22	2.15189325464331	41.3774431561658	30/12/1899 22:00:00
14	Manca d'atenció a la conducció	Tarda	430069.43	4582012.05	2.16249420388788	41.3847522784463	30/12/1899 14:00:00
13	No respetar distancias	Mati	430111.63	4582053.06	2.16299413656188	41.3851253128875	30/12/1899 13:00:00
8	No respetar distancias	Mati	430103.84	4582045.42	2.1629018502635	41.3850536244507	30/12/1899 8:00:00
10	No respetar distancias	Mati	430080.36	4582095.48	2.16261527843894	41.3855046566338	30/12/1899 10:00:00
20	Manca precaució incorporació circulació	Tarda	431112.06	4583023.89	2.17484780884279	41.3939556563002	30/12/1899 20:00:00
16	Canvi de canil sense precaució	Tarda	430934.78	4582858.79	2.17274626758835	41.39245343302	30/12/1899 16:00:00
12	Manca d'atenció a la conducció	Mati	430773.34	4582701.12	2.171083342124943	41.3910194562991	30/12/1899 12:00:00
8	Manca d'atenció a la conducció	Mati	431513.75	4583425.21	2.17960686322714	41.3976045465574	30/12/1899 8:00:00

Lo que nos interesa es el formato corto de hora, así que vemos que los valores de datos ya tienen el tipo de dato de Hora y cambiamos el formato a Short Time para obtener solamente la hora.

	Descripción_causa_mediate	Descripción_torn	Coordenada_UTM_X_ED50	Coordenada_UTM_Y_ED50	Longitud	Latitud	Hora
7	No respetar distancias	Mati	430080.36	4582095.48	2.16261527843894	41.3855046566338	7:00
11	Altres	Mati	430062.55	4582005.38	2.16241269697104	41.3846916052752	11:00
11	Canvi de canil sense precaució	Mati	429629.38	4581777.7	2.15965062520641	41.3826264020394	11:00
12	No respetar distancias	Mati	430035.23	4582321.24	2.16623538089505	41.387564496894	12:00
9	No respetar distancias	Mati	430307.64	4582245.3	2.1653160888093	41.386873776318	9:00
11	Desconegut	Mati	430299.85	4582237.63	2.16522380978352	41.3868040208313	11:00
15	Altres	Tarda	431314.18	4582320.39	2.17724185676572	41.3958328428663	15:00
16	No respetar distancias	Tarda	430934.78	4582858.79	2.17274626758835	41.39245343302	16:00
18	No determinada	Tarda	430418.29	4582353.68	2.16662693156219	41.3878595095211	18:00
19	Manca d'atenció a la conducció	Tarda	431024.51	4582946.76	2.17380944523285	41.3932534932386	19:00
21	No respetar distancias	Tarda	431112.06	4583023.89	2.17484780884279	41.3939556563002	20:00
21	No respetar distancias	Tarda	431690.71	4582793.42	2.1840742573486	41.4009529703384	21:00
11	Canvi de canil sense precaució	Mati	429731.19	4581680.88	2.1584876352699	41.3817400392224	11:00
20	No respetar distancias	Nit	429175.13	4581209.22	2.15189325464331	41.3774431561658	22:00
14	Manca d'atenció a la conducció	Tarda	430069.43	4582012.05	2.16249420388788	41.3847522784463	14:00
13	No respetar distancias	Mati	430111.63	4582053.06	2.16299413656188	41.3851253128875	13:00
8	No respetar distancias	Mati	430103.84	4582045.42	2.1629018502635	41.3850536244507	8:00
10	No respetar distancias	Mati	430080.36	4582095.48	2.16261527843894	41.3855046566338	10:00
20	Manca precaució incorporació circulació	Tarda	431112.06	4583023.89	2.17484780884279	41.3939556563002	20:00
16	Canvi de canil sense precaució	Tarda	430934.78	4582858.79	2.17274626758835	41.39245343302	16:00
12	Manca d'atenció a la conducció	Mati	430773.34	4582701.12	2.171083342124943	41.3910194562991	12:00
8	Manca d'atenció a la conducció	Mati	431513.75	4583425.21	2.17960686322714	41.3976045465574	8:00

## Fecha

Para la fecha la tenemos separada en tres columnas (Año, Mes, Dia), desde el editor de Power Bi, seleccionamos las tres columnas y la pestaña agregar columna combinamos las columnas y le asignamos un separador personalizado en este caso una barra diagonal (/).

Combinar columnas

Elige cómo combinar las columnas seleccionadas.

Separador: /

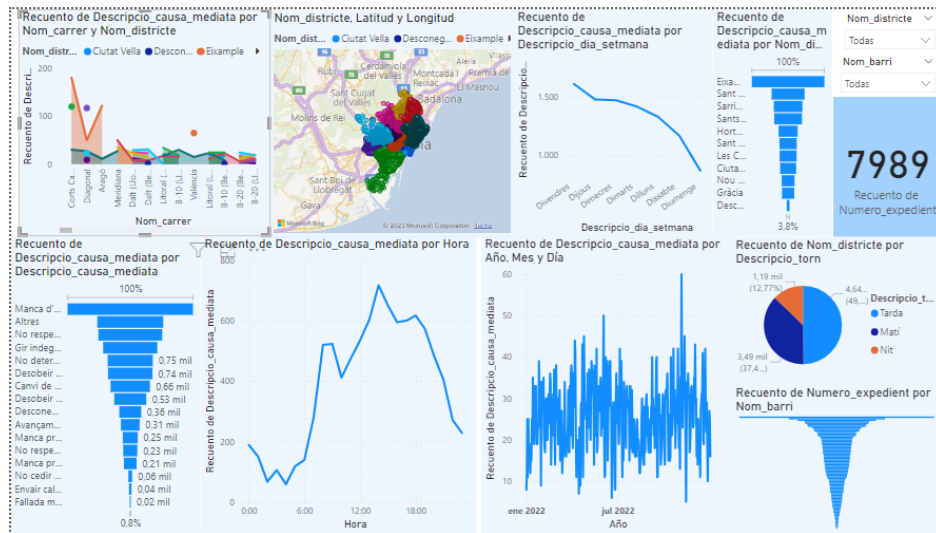
Nuevo nombre de columna (opcional): Fecha

Aceptar Cancelar

## Proyecto Final – Business Intelligence

La columna Fecha fue creada y los valores tienen el tipo de dato de Texto por lo que dentro del mismo editor seleccionamos el tipo de dato Fecha.

### Estructura de cuadro de mando



El cuadro de mando se compone de 10 visualizaciones y dos tipos de segmentación de datos. Se tiene un gráfico de áreas, mapa, tarjeta, circular, 3 gráficos de líneas y 3 de embudo.

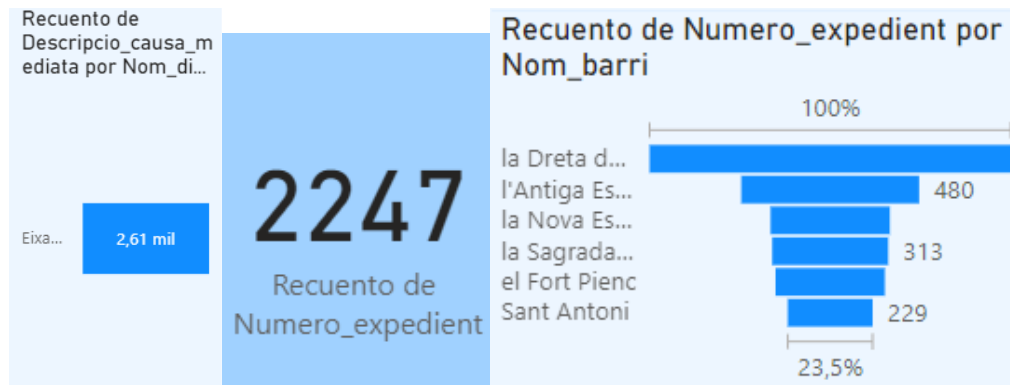
### Análisis de datos

El cuadro de mando en su forma inicial muestra una imagen general de la accidentabilidad de la ciudad de Barcelona con 7989 expedientes realizados, así podemos decir que en el año 2022 los días viernes a las 14:00 hrs fue el momento donde hubo un mayor número de accidentes debido a la falta de atención de los conductores y el 5 de noviembre es el día con mayor número de accidentes. La calle con mayor número de accidentes es Corts Catalanes perteneciente al distrito de Eixample en el barrio de la Dreta de l'Eixample por la tarde.

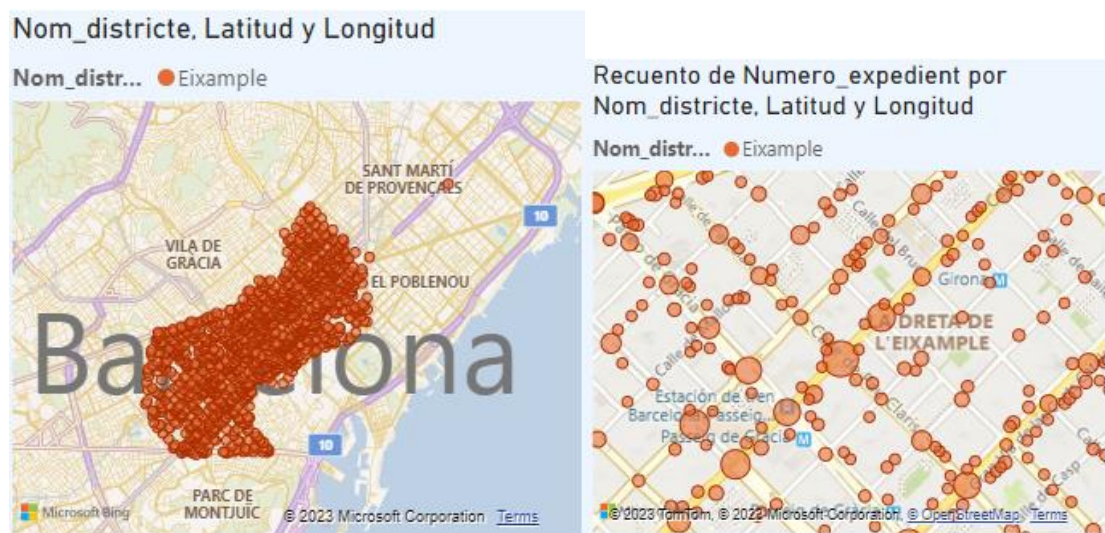
La segmentación de datos puede realizarse mediante el nombre del distrito y barrio, sabemos que el distrito con mayor numero de accidentes es Eixample y el menor Gracia, existe un segmento Desconocido que tiene 79 expedientes que equivalen al 3.8%.

Seleccionando el distrito con mayor recuento de accidentes que es Eixample se tienen 2247 números de expedientes con 2614 causas de accidentes, siendo el barrio de la Dreta de l'Eixample el que tiene un mayor número de incidencias.

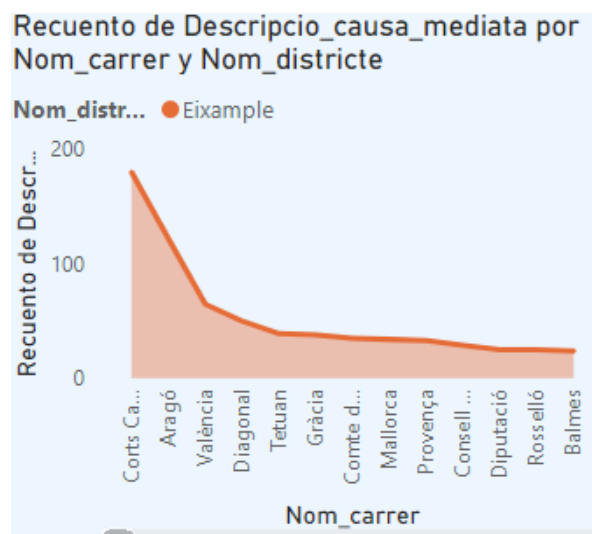
## Proyecto Final – Business Intelligence



Eixample se puede apreciar la gran densidad en sus calles dejando zonas en blanco que corresponden a parques y que la intersección de Corts Catalanes y Plaça Claris es la peligrosa de Eixample.

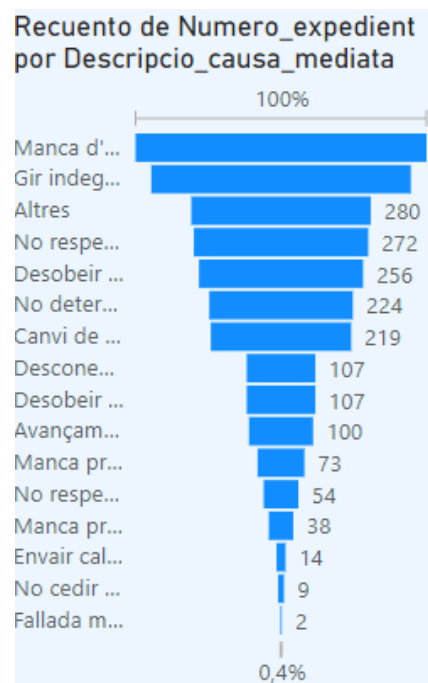


Siendo las tres primeras calles con un mayor número de accidentes Cort Catalanes, Aragón y Valencia.

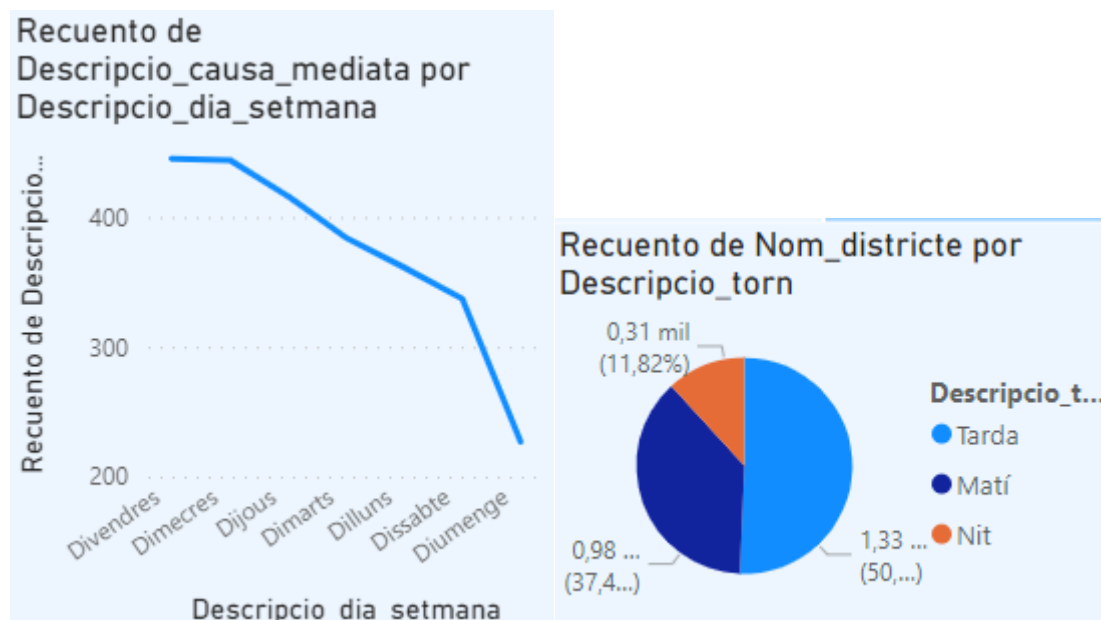


## Proyecto Final – Business Intelligence

Gran parte de dichos accidentes en estas calles fueron debido a la falta de atención del conductor y el giro indebido o sin precaución.



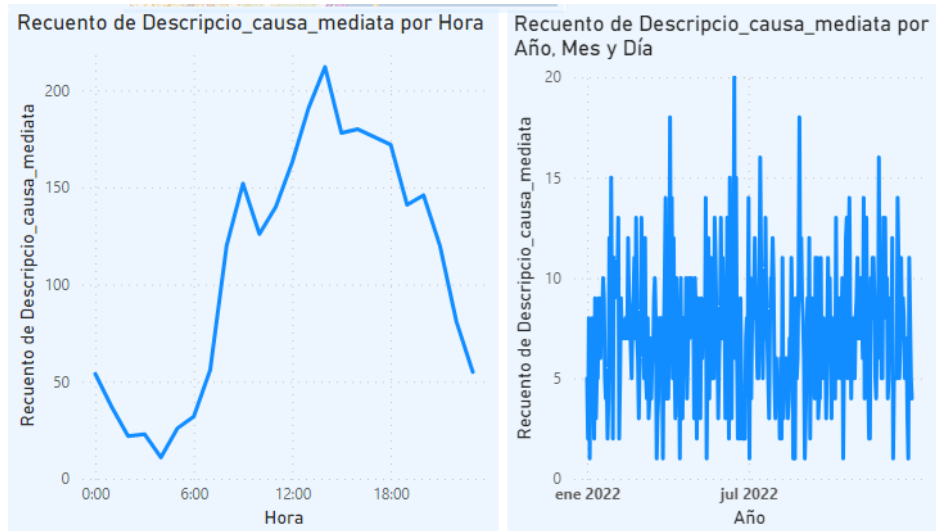
La distribución durante la semana de los accidentes es mayor los viernes, seguido de los miércoles, y más del 50% suceden por la tarde.



La mayor distribución temporal de los accidentes es de las 12:00 a 18:00 hrs, parece que los datos presentan cierto ciclo o estacionalidad, gran parte de la disminución de accidentes se debe a los fines de semana, días festivos y vacaciones.



## Proyecto Final – Business Intelligence



## Conclusiones

Barcelona es una ciudad dinámica y productiva por lo que tiene sentido que los datos muestren que gran parte de los accidentes suceden los jueves y viernes por la tarde en vías de entrada/salida de la ciudad. Sugiriendo que son desplazamientos de personas que se van a su segunda residencia los fines de semana.

Esto también se puede comprobar en que en el mes de agosto se tiene un valle en los accidentes debido a que gran parte de la población se encuentra fuera de la ciudad reduciendo los desplazamientos.

Con esta información se puede realizar una mejor planificación de los servicios de seguridad para poder gestionar mejor los puntos más críticos de la ciudad, por ejemplo, en el Eixample en el barrio de la Dreta de l'Eixample en el tramo de Corts Catalanes entre Pau Claris y Paseo de Sant Joan.