| **Coderhouse - Data Analytics**  **Tercer Entrega**   | Caso de estudio: | **Siniestros viales**  **en la provincia de Buenos Aires** | | --- | --- | |  |  | |  |  | | Alumno: | **Juan Pablo Ludueña** | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |

ÍNDICE

[Tabla de versionado 2](#_nkf6bd1fl7zt)

[Temática: 3](#_m8scqiy0j9zi)

[Alcance: 3](#_kgi6h0oi9mto)

[Objetivo: 3](#_2un4sy822qlm)

[Hipótesis: 3](#_dir95vutpz4w)

[Origen del Dataset: 4](#_5fk33wqb0h81)

[Diagrama Entidad - Relación (DER) 5](#_vy35uy536it0)

[Listado de Tablas 6](#_7vsbnfwrj4xu)

[Detalle de columnas 7](#_vgeqifmawaxz)

[Transformación de datos 9](#_ab34ns5a7825)

[Tablas auxiliares: 17](#_yj8t4wgz785v)

[Modelo de datos en Power Bi **19**](#_g0zjq02uh5zr)

# 

# **Tabla de versionado**

| **Version** | **fecha** | **observaciónes** |
| --- | --- | --- |
| 1 | 15/05/2024 | Primer entregable |
| 2 | 05/06/2024 | Segundo entregable |
| 3 | 08/07/2024 | Tercer entregable |
|  |  |  |

# 

# **Temática:**

Abordaremos el desarrollo en relación a Información sobre lesiones en siniestros viales ocurridos en la Ciudad de Buenos Aires. Los datos incluyen fecha y ubicación del hecho y tipo de transporte involucrado. Además se especifica el género y edad de las víctimas y el tipo de lesión sufrida.

# **Alcance:**

El desarrollo del dashboard apuntará a mostrar información a un nivel táctico, para que los líderes y mandos medios del Observatorio de Movilidad y Seguridad Vial de la Ciudad de Buenos Aires (OMSV) revisen, estudien y tomen las decisiones pertinentes al aumento de siniestros que mostraron un crecimiento sostenido entre los años 2019 y 2021

# **Objetivo:**

Desarrollar un dashboard que permita a los interesados obtener información de calidad para la futura toma de decisiones en relación a los siniestros registrados en la ciudad de Buenos Aires, entre los años 2019 y 2021

# **Hipótesis:**

Es necesario entender si se han aplicado correctas políticas y controles viales dentro del período mencionado que hayan hecho bajar la cantidad de siniestros viales y la gravedad de los mismos, en caso de no ser así es importante tener en cuenta ésta información para futuras tomas de decisiones.

# 

# 

# **Origen del Dataset:**

Éste dataset toma como su principal fuente de información datos policiales, tal y como recomiendan los estándares internacionales. Las estadísticas elaboradas se realizan en base a los sumarios que instruye la Policía de la Ciudad ante la clase de delito: lesiones culposas (Arts. 94 y 94 bis del Código Penal de la Nación)

En el año 2018, comenzó a operar la Policía de la Ciudad, dependiente del Ministerio de Justicia y Seguridad del GCBA (MJyS).

La Policía de la Ciudad cubrió la totalidad del territorio e integró la información que hasta entonces recopilaban las fuerzas de manera independiente. En virtud de ello, a partir del segundo cuatrimestre de 2018, la principal fuente de información sobre siniestros viales con la que trabaja el Observatorio es la del MJyS, fuente que permite conocer una mayor cantidad de casos por cubrir la totalidad del territorio de la Ciudad.

Las estadísticas de delitos por Lesiones Culposas informadas por el MJyS son procesadas por el equipo del Observatorio para verificar que los incidentes se ajusten a las definiciones de siniestro vial vertidas en este documento.

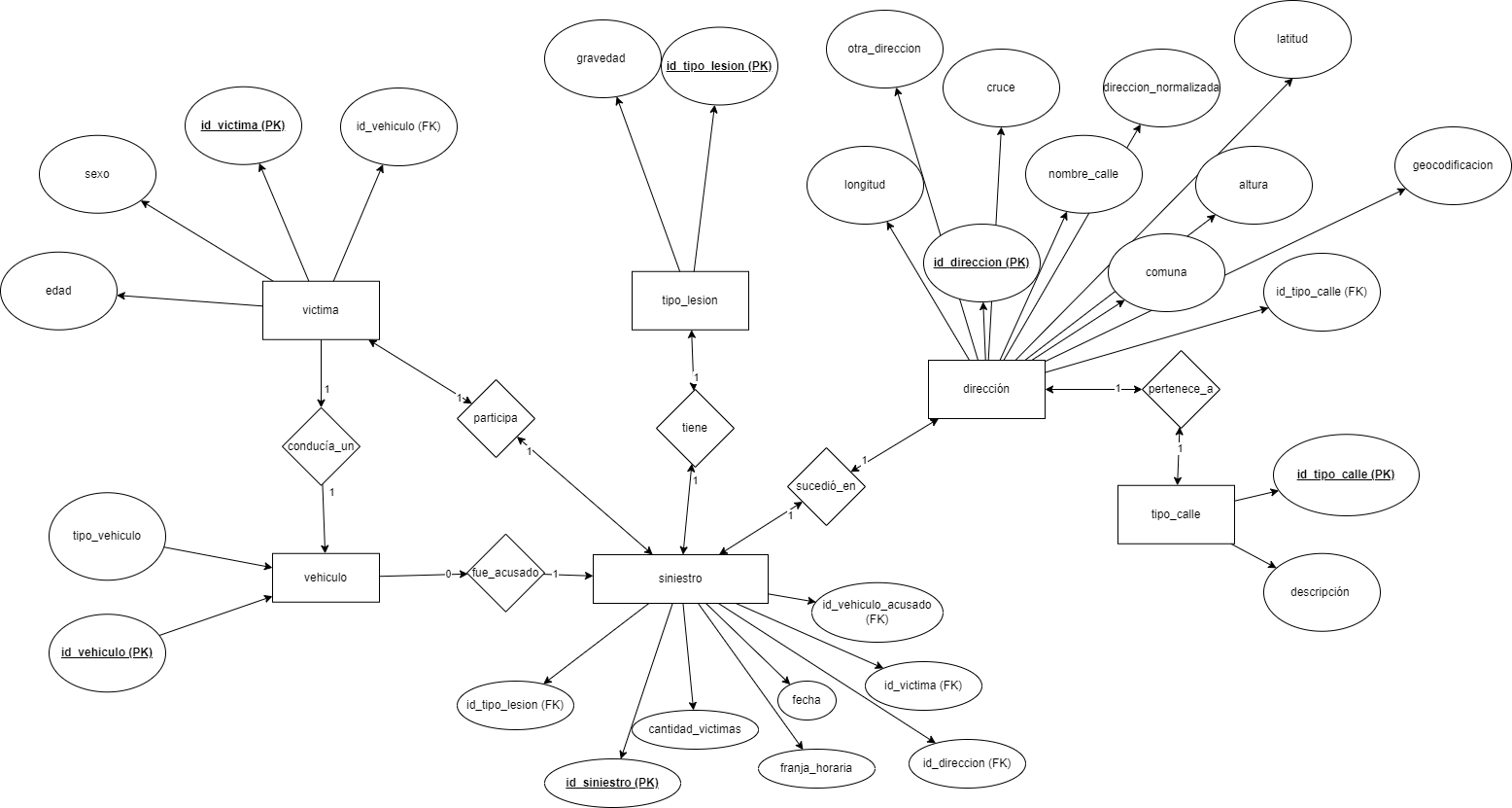
Asimismo, con el relevamiento propio a los trece hospitales de agudos de la Ciudad de Buenos Aires se logró asignar gravedad de la lesión, considerando siniestros graves a aquellos que tuvieron víctimas que debieron ser hospitalizadas por al menos 24hs.

[Link al dataset (xls)](https://docs.google.com/spreadsheets/d/1cJ6EZyFmS4iDRSlUQy6jZ30bEpuPeRZ0/edit?usp=drive_link&ouid=104511010569512327633&rtpof=true&sd=true)

[Link al dataset de comunas (csv)](https://drive.google.com/file/d/1mtejtdyEZk0YhGsyO_pM-lnROL301_-h/view?usp=drive_link)

# **Diagrama Entidad - Relación (DER)**

Para una mejor visualización, ver el siguiente [link](https://drive.google.com/file/d/1-_TwIbOkDpL_TT-NXGdwwhFuRoEFjW2t/view?usp=sharing)

****

# **Listado de Tablas**

En este apartado, se hará mención de cada una de las tablas junto a una breve

descripción de las mismas y la definición de la clave primaria y foránea:

* **siniestro:** Incidente ocurrido en una vía de acceso público en el que participa al menos un vehículo y que resulta en al menos una persona lesionada o fallecida.
  + PK: id\_siniestro
  + FK: id\_tipo\_lesion; id\_calle; id\_victima; vehiculo\_acusado
* **victima:** Datos de quien haya fallecido o se haya lastimado a raíz del hecho, o bien peatón/a
  + PK: id\_victima
  + FK: id\_vehiculo
* **vehículo:** Datos del vehículo que ocupaba la víctima o que ocupaba quien resultó acusado/a del hecho, sin implicar culpabilidad legal
  + PK: id\_vehiculo
  + FK: N/A
* **tipo\_de\_lesión:** Nivel máximo conocido de gravedad de la lesión de la(s) víctima(s) del siniestro en función del tiempo de hospitalización
  + PK: id\_tipo\_de\_lesion
  + FK: N/A
* **direccion:** Dirección donde se produjo el siniestro
  + PK: id\_dirección
  + FK: id\_tipo\_de\_calle
* **tipo\_de\_calle:** Tipo de arteria. En el caso de intersecciones a nivel se clasifica según la de mayor jerarquía
  + PK: id\_tipo\_de\_calle
  + FK: N/A

# **Detalle de columnas**

En la siguientes páginas presentamos las tablas el DER del proyecto final, con sus

Columnas, Descripción, Tipo de Dato, Primary key y Foreign key.

| **Tabla: siniestro** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_siniestro | identificador unico del siniestro | numerico | SI | - |
| cantidad\_victimas | cantidad de víctimas | numerico | - | - |
| fecha | fecha en formato dd/mm/aaaa | fecha | - | - |
| id\_tipo\_lesion | id del tipo de lesion | numerico | - | SI |
| id\_vehiculo\_acusado | id del vehiculo acusado | numerico | - | SI |
| id\_victima | id de la victima | numerico | - | SI |
| id\_direccion | id de la direccion | numerico | - | SI |
| franja\_horaria | franja horaria entera | numerico | - | - |

| **Tabla: direccion** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_direccion | identificador de direccion | numerico | SI | - |
| longitud | longitud con separador punto. WGS84 | numerico | - | - |
| otra\_direccion | direccion en formato alternativo | texto | - | - |
| cruce | cruce en caso de que sea una encrucijada | texto | - | - |
| direccion\_normalizada | direccion en formato normalizado USIG | texto | - | - |
| nombre\_calle | nombre de la arteria donde se produjo el hecho | texto | - | - |
| comuna | comuna de CABA con jurisdicción según la ubicación | texto | - | - |
| latitud | latitud con separador punto. WGS84 | numerico | - | - |
| altura | altura de la arteria donde se produjo el hecho | numerico | - | - |
| id\_tipo\_calle | identificador del tipo de calle | numerico | - | SI |
| geocodificacion | geocodificación plana | numerico | - | - |

| **Tabla: tipo\_calle** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_tipo\_calle | identificador unico del tipo de calle | numerico | SI | - |
| descripción | Tipo de arteria. En el caso de intersecciones a nivel se clasifica según la de mayor jerarquía | texto | - | - |

| **Tabla: tipo\_lesion** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_tipo\_lesion | identificador unico del tipo de lesion | numerico | SI | - |
| gravedad | Nivel máximo conocido de gravedad de la lesión de la(s) víctima(s) del siniestro en función del tiempo de hospitalización | texto | - | - |

| **Tabla: victima** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_victima | identificador unico de la victima | numerico | SI | - |
| id\_vehiculo | identificador del vehiculo | numerico | - | SI |
| sexo | Sexo de la víctima según registro policial | texto | - | - |
| edad | Edad en años enterios de la víctima en registro policial | numerico | - | - |

| **Tabla: vehiculo** | | | | |
| --- | --- | --- | --- | --- |
| **columna** | **descripcion** | **tipo de dato** | **PK** | **FK** |
| id\_vehiculo | identificador unico del vehiculo | numerico | SI | - |
| tipo\_vehiculo | tipo de vehiculo que participó del siniestro | texto | - | - |

# **Transformación de datos**

**hechos-ppal:** contiene los datos del excel principal de “hechos”

* se cambian los tipos de datos
* se reemplazan valores SD por null
* se eliminan columnas "calle", "altura", "cruce", "hora", "tipo\_calle", "acusado" ya que no contienen datos o son datos irrelevantes

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla1\_Table = Source{[Item="Tabla1",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla1\_Table,{{"id", type text}, {"n\_victimas", Int64.Type}, {"aaaa", Int64.Type}, {"mm", Int64.Type}, {"dd", Int64.Type}, {"fecha", type date}, {"hora", type datetime}, {"franja\_hora", Int64.Type}, {"direccion\_normalizada", type text}, {"comuna", Int64.Type}, {"tipo\_calle", type text}, {"otra\_direccion", type text}, {"calle", type any}, {"altura", type any}, {"cruce", type any}, {"geocodificacion\_CABA", type text}, {"longitud", type text}, {"latutid", type text}, {"victima", type text}, {"acusado", type text}, {"participantes", type text}, {"moto", type text}, {"auto", type text}, {"transporte\_publico", type text}, {"camion", type text}, {"ciclista", type text}, {"gravedad", type text}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type",null,"SD",Replacer.ReplaceValue,{"direccion\_normalizada"}),

#"Removed Columns" = Table.RemoveColumns(#"Replaced Value",{"calle", "altura", "cruce"}),

#"Replaced Value1" = Table.ReplaceValue(#"Removed Columns",null,"SD",Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1",null,"SD",Replacer.ReplaceValue,{"longitud"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2",null,"SD",Replacer.ReplaceValue,{"latutid"}),

#"Removed Columns1" = Table.RemoveColumns(#"Replaced Value3",{"hora", "tipo\_calle", "acusado"}),

#"Uppercased Text" = Table.TransformColumns(#"Removed Columns1",{{"gravedad", Text.Upper, type text}}),

#"Replaced Value4" = Table.ReplaceValue(#"Uppercased Text","GRAVE","1",Replacer.ReplaceText,{"gravedad"}),

#"Replaced Value5" = Table.ReplaceValue(#"Replaced Value4","FATAL","2",Replacer.ReplaceText,{"gravedad"})

in

#"Replaced Value5"

**victimas-ppal:** carga de la base de datos de victimas

* corrección de tipos de datos
* join con el dataset “tipo\_vehiculo”

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla2\_Table = Source{[Item="Tabla2",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla2\_Table,{{"ID hecho", type text}, {"AAA", Int64.Type}, {"MM", Int64.Type}, {"DD", Int64.Type}, {"FECHA ", type date}, {"VEHICULO\_VICTIMA", type text}, {"SEXO", type text}, {"EDAD\_VICTIMA", type any}, {"GRAVEDAD", type text}}),

#"Merged Queries" = Table.NestedJoin(#"Changed Type", {"VEHICULO\_VICTIMA"}, tipo\_vehiculo, {"tipo\_vehiculo"}, "tipo\_vehiculo", JoinKind.LeftOuter),

#"Expanded tipo\_vehiculo" = Table.ExpandTableColumn(#"Merged Queries", "tipo\_vehiculo", {"id\_tipo\_vehiculo"}, {"tipo\_vehiculo.id\_tipo\_vehiculo"}),

#"Renamed Columns" = Table.RenameColumns(#"Expanded tipo\_vehiculo",{{"tipo\_vehiculo.id\_tipo\_vehiculo", "id\_tipo\_vehiculo"}})

in

#"Renamed Columns"

**comunas:** carga de base de datos de comunas, para tener mas detalle de la dirección del siniestro

* corrección de tipo de datos

let

Source = Csv.Document(File.Contents("D:\coderhouse\data analytics flex\comunas.csv"),[Delimiter=";", Columns=5, Encoding=1252, QuoteStyle=QuoteStyle.None]),

#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),

#"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"ID", Int64.Type}, {"COMUNAS", Int64.Type}, {"BARRIOS", type text}, {"", type text}, {"\_1", type text}}),

#"Removed Columns" = Table.RemoveColumns(#"Changed Type",{"", "\_1", "COMUNAS"}),

#"Renamed Columns" = Table.RenameColumns(#"Removed Columns",{{"ID", "id\_comuna"}, {"BARRIOS", "barrios"}})

in

#"Renamed Columns"

**tipo-lesion:** dataset que contiene la tipificación de las lesiones

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla1\_Table = Source{[Item="Tabla1",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla1\_Table,{{"id", type text}, {"n\_victimas", Int64.Type}, {"aaaa", Int64.Type}, {"mm", Int64.Type}, {"dd", Int64.Type}, {"fecha", type date}, {"hora", type datetime}, {"franja\_hora", Int64.Type}, {"direccion\_normalizada", type text}, {"comuna", Int64.Type}, {"tipo\_calle", type text}, {"otra\_direccion", type text}, {"calle", type any}, {"altura", type any}, {"cruce", type any}, {"geocodificacion\_CABA", type text}, {"longitud", type text}, {"latutid", type text}, {"victima", type text}, {"acusado", type text}, {"participantes", type text}, {"moto", type text}, {"auto", type text}, {"transporte\_publico", type text}, {"camion", type text}, {"ciclista", type text}, {"gravedad", type text}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type",null,"SD",Replacer.ReplaceValue,{"direccion\_normalizada"}),

#"Removed Columns" = Table.RemoveColumns(#"Replaced Value",{"calle", "altura", "cruce"}),

#"Replaced Value1" = Table.ReplaceValue(#"Removed Columns",null,"SD",Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1",null,"SD",Replacer.ReplaceValue,{"longitud"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2",null,"SD",Replacer.ReplaceValue,{"latutid"}),

#"Removed Columns1" = Table.RemoveColumns(#"Replaced Value3",{"hora", "tipo\_calle", "acusado"}),

#"Uppercased Text" = Table.TransformColumns(#"Removed Columns1",{{"gravedad", Text.Upper, type text}}),

#"Removed Columns2" = Table.RemoveColumns(#"Uppercased Text",{"id", "n\_victimas", "aaaa", "mm", "dd", "fecha", "franja\_hora", "direccion\_normalizada", "comuna", "otra\_direccion", "geocodificacion\_CABA", "longitud", "latutid", "victima", "participantes", "moto", "auto", "transporte\_publico", "camion", "ciclista"}),

#"Removed Duplicates" = Table.Distinct(#"Removed Columns2"),

#"Removed Top Rows" = Table.Skip(#"Removed Duplicates",1),

#"Added Index" = Table.AddIndexColumn(#"Removed Top Rows", "Index", 1, 1, Int64.Type),

#"Renamed Columns" = Table.RenameColumns(#"Added Index",{{"Index", "id\_tipo\_lesion"}})

in

#"Renamed Columns"

**hechos-ppal-dir-norm:** dataset que se crea para poder hacer union con el dataset “hechos-ppal-otra-dir” y así poder únificar las direcciones de los siniestros en una sola columna

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla1\_Table = Source{[Item="Tabla1",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla1\_Table,{{"id", type text}, {"n\_victimas", Int64.Type}, {"aaaa", Int64.Type}, {"mm", Int64.Type}, {"dd", Int64.Type}, {"fecha", type date}, {"hora", type datetime}, {"franja\_hora", Int64.Type}, {"direccion\_normalizada", type text}, {"comuna", Int64.Type}, {"tipo\_calle", type text}, {"otra\_direccion", type text}, {"calle", type any}, {"altura", type any}, {"cruce", type any}, {"geocodificacion\_CABA", type text}, {"longitud", type text}, {"latutid", type text}, {"victima", type text}, {"acusado", type text}, {"participantes", type text}, {"moto", type text}, {"auto", type text}, {"transporte\_publico", type text}, {"camion", type text}, {"ciclista", type text}, {"gravedad", type text}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type",null,"SD",Replacer.ReplaceValue,{"direccion\_normalizada"}),

#"Removed Columns" = Table.RemoveColumns(#"Replaced Value",{"calle", "altura", "cruce"}),

#"Replaced Value1" = Table.ReplaceValue(#"Removed Columns",null,"SD",Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1",null,"SD",Replacer.ReplaceValue,{"longitud"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2",null,"SD",Replacer.ReplaceValue,{"latutid"}),

#"Removed Columns1" = Table.RemoveColumns(#"Replaced Value3",{"hora", "tipo\_calle", "acusado"}),

#"Uppercased Text" = Table.TransformColumns(#"Removed Columns1",{{"gravedad", Text.Upper, type text}}),

#"Replaced Value4" = Table.ReplaceValue(#"Uppercased Text","GRAVE","1",Replacer.ReplaceText,{"gravedad"}),

#"Replaced Value5" = Table.ReplaceValue(#"Replaced Value4","FATAL","2",Replacer.ReplaceText,{"gravedad"}),

#"Filtered Rows" = Table.SelectRows(#"Replaced Value5", each ([direccion\_normalizada] <> "SD")),

#"Removed Columns2" = Table.RemoveColumns(#"Filtered Rows",{"otra\_direccion"}),

#"Renamed Columns" = Table.RenameColumns(#"Removed Columns2",{{"direccion\_normalizada", "direccion"}})

in

#"Renamed Columns"

**hechos-ppal-otra-dir:** dataset que se crea para poder hacer union con el dataset “hechos-ppal-dir-norm” y así poder únificar las direcciones de los siniestros en una sola columna

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla1\_Table = Source{[Item="Tabla1",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla1\_Table,{{"id", type text}, {"n\_victimas", Int64.Type}, {"aaaa", Int64.Type}, {"mm", Int64.Type}, {"dd", Int64.Type}, {"fecha", type date}, {"hora", type datetime}, {"franja\_hora", Int64.Type}, {"direccion\_normalizada", type text}, {"comuna", Int64.Type}, {"tipo\_calle", type text}, {"otra\_direccion", type text}, {"calle", type any}, {"altura", type any}, {"cruce", type any}, {"geocodificacion\_CABA", type text}, {"longitud", type text}, {"latutid", type text}, {"victima", type text}, {"acusado", type text}, {"participantes", type text}, {"moto", type text}, {"auto", type text}, {"transporte\_publico", type text}, {"camion", type text}, {"ciclista", type text}, {"gravedad", type text}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type",null,"SD",Replacer.ReplaceValue,{"direccion\_normalizada"}),

#"Removed Columns" = Table.RemoveColumns(#"Replaced Value",{"calle", "altura", "cruce"}),

#"Replaced Value1" = Table.ReplaceValue(#"Removed Columns",null,"SD",Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1",null,"SD",Replacer.ReplaceValue,{"longitud"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2",null,"SD",Replacer.ReplaceValue,{"latutid"}),

#"Removed Columns1" = Table.RemoveColumns(#"Replaced Value3",{"hora", "tipo\_calle", "acusado"}),

#"Uppercased Text" = Table.TransformColumns(#"Removed Columns1",{{"gravedad", Text.Upper, type text}}),

#"Replaced Value4" = Table.ReplaceValue(#"Uppercased Text","GRAVE","1",Replacer.ReplaceText,{"gravedad"}),

#"Replaced Value5" = Table.ReplaceValue(#"Replaced Value4","FATAL","2",Replacer.ReplaceText,{"gravedad"}),

#"Filtered Rows" = Table.SelectRows(#"Replaced Value5", each ([otra\_direccion] <> "Comisaria: Vecinal 13A. Calle y altura: Figueroa Alcorta, Pres. Av. 6400 Piso: Dpto: Barrio: Belgrano" and [otra\_direccion] <> "Comisaria: Vecinal 13A. Calle y altura: Pino, Virrey Del 2200 Piso: Dpto: Barrio: Belgrano" and [otra\_direccion] <> "Comisaria: Vecinal 14A. Calle y altura: Santa Fe Av. 3700 Piso: Dpto: Barrio: Palermo" and [otra\_direccion] <> "Comisaria: Vecinal 15B. Calle y altura: Cordoba Av. 4690 Piso: Dpto: Barrio: Villa Crespo" and [otra\_direccion] <> "Comisaria: Vecinal 1A. Calle y altura: Pellegrini, Carlos 1175 Piso: Dpto: Barrio: Retiro" and [otra\_direccion] <> "Comisaria: Vecinal 1C. Calle y altura: San Juan Av. 1800 Piso: Dpto: Barrio: Constitucion" and [otra\_direccion] <> "Comisaria: Vecinal 1E. Calle y altura: Manso Juana 1601 Piso: Dpto: Barrio: Puerto Madero" and [otra\_direccion] <> "Comisaria: Vecinal 7C. Calle y altura: BONIFACIO, JOSE 2948 Piso: Dpto: Barrio: Flores" and [otra\_direccion] <> "Comisaria: Vecinal 8C. Calle y altura: Roca, Cnel. Av. y Larrazabal Av. 0 Piso: Dpto: Barrio: Villa Lugano" and [otra\_direccion] <> "Comisaria: Vecinal 9A. Calle y altura: Cardenas, Cnel. Av. y Tandil 0 Piso: Dpto: Barrio: Mataderos" and [otra\_direccion] <> "SD")),

#"Removed Columns2" = Table.RemoveColumns(#"Filtered Rows",{"direccion\_normalizada"}),

#"Renamed Columns" = Table.RenameColumns(#"Removed Columns2",{{"otra\_direccion", "direccion"}}),

#"Reordered Columns" = Table.ReorderColumns(#"Renamed Columns",{"id", "n\_victimas", "aaaa", "mm", "dd", "fecha", "franja\_hora", "direccion", "comuna", "geocodificacion\_CABA", "longitud", "latutid", "victima", "participantes", "moto", "auto", "transporte\_publico", "camion", "ciclista", "gravedad"})

in

#"Reordered Columns"

**hechos-ppal-normalizada:** dataset final de siniestros es el dataset de hechos del modelo estrella

let

Source = Table.Combine({#"hechos-ppal-dir-norm", #"hechos-ppal-otra-dir"}),

#"Merged Queries" = Table.NestedJoin(Source, {"direccion"}, direcciones, {"direccion"}, "direcciones", JoinKind.LeftOuter),

#"Renamed Columns" = Table.RenameColumns(#"Merged Queries",{{"direcciones", "direcciones\_id"}}),

#"Expanded direcciones\_id" = Table.ExpandTableColumn(#"Renamed Columns", "direcciones\_id", {"direcciones\_id"}, {"direcciones\_id.direcciones\_id"}),

#"Removed Columns" = Table.RemoveColumns(#"Expanded direcciones\_id",{"moto", "auto", "transporte\_publico", "camion", "ciclista"}),

#"Split Column by Delimiter" = Table.SplitColumn(#"Removed Columns", "participantes", Splitter.SplitTextByDelimiter("-", QuoteStyle.Csv), {"participantes.1", "participantes.2"}),

#"Changed Type" = Table.TransformColumnTypes(#"Split Column by Delimiter",{{"participantes.1", type text}, {"participantes.2", type text}}),

#"Removed Columns1" = Table.RemoveColumns(#"Changed Type",{"victima"}),

#"Renamed Columns1" = Table.RenameColumns(#"Removed Columns1",{{"participantes.1", "victima"}, {"participantes.2", "acusado"}}),

#"Merged Queries1" = Table.NestedJoin(#"Renamed Columns1", {"victima"}, tipo\_vehiculo, {"tipo\_vehiculo"}, "tipo\_vehiculo", JoinKind.LeftOuter),

#"Expanded tipo\_vehiculo" = Table.ExpandTableColumn(#"Merged Queries1", "tipo\_vehiculo", {"id\_tipo\_vehiculo"}, {"tipo\_vehiculo.id\_tipo\_vehiculo"}),

#"Renamed Columns2" = Table.RenameColumns(#"Expanded tipo\_vehiculo",{{"tipo\_vehiculo.id\_tipo\_vehiculo", "id\_tipo\_vehiculo\_victima"}}),

#"Merged Queries2" = Table.NestedJoin(#"Renamed Columns2", {"acusado"}, tipo\_vehiculo, {"tipo\_vehiculo"}, "tipo\_vehiculo", JoinKind.LeftOuter),

#"Expanded tipo\_vehiculo1" = Table.ExpandTableColumn(#"Merged Queries2", "tipo\_vehiculo", {"id\_tipo\_vehiculo"}, {"tipo\_vehiculo.id\_tipo\_vehiculo"}),

#"Renamed Columns3" = Table.RenameColumns(#"Expanded tipo\_vehiculo1",{{"tipo\_vehiculo.id\_tipo\_vehiculo", "id\_tipo\_vehiculo\_acusado"}}),

#"Replaced Errors" = Table.ReplaceErrorValues(#"Renamed Columns3", {{"franja\_hora", 0}}),

#"Renamed Columns4" = Table.RenameColumns(#"Replaced Errors",{{"gravedad", "id\_tipo\_lesion"}, {"direcciones\_id.direcciones\_id", "direccion\_id"}}),

#"Replaced Value" = Table.ReplaceValue(#"Renamed Columns4","SD",null,Replacer.ReplaceValue,{"longitud"}),

#"Changed Type1" = Table.TransformColumnTypes(#"Replaced Value",{{"longitud", Int64.Type}}),

#"Replaced Value1" = Table.ReplaceValue(#"Changed Type1","sd",null,Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1","sd",null,Replacer.ReplaceValue,{"latutid"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2","SD",null,Replacer.ReplaceValue,{"geocodificacion\_CABA"}),

#"Replaced Value4" = Table.ReplaceValue(#"Replaced Value3","SD",null,Replacer.ReplaceValue,{"latutid", "victima", "acusado", "id\_tipo\_lesion"}),

#"Changed Type2" = Table.TransformColumnTypes(#"Replaced Value4",{{"latutid", Int64.Type}, {"victima", type text}, {"acusado", type text}, {"id\_tipo\_lesion", Int64.Type}, {"direccion\_id", type any}}),

#"Merged Queries3" = Table.NestedJoin(#"Changed Type2", {"direccion"}, direcciones, {"direccion"}, "direcciones", JoinKind.LeftOuter),

#"Removed Columns2" = Table.RemoveColumns(#"Merged Queries3",{"direccion\_id"}),

#"Expanded direcciones1" = Table.ExpandTableColumn(#"Removed Columns2", "direcciones", {"id\_direccion"}, {"direcciones.id\_direccion"}),

#"Renamed Columns7" = Table.RenameColumns(#"Expanded direcciones1",{{"direcciones.id\_direccion", "id\_direccion"}})

in

#"Renamed Columns7"

**direcciones:** se separa éste dataset para tener las direcciones separado de la tabla de hechos

let

Source = Table.Combine({#"hechos-ppal-dir-norm", #"hechos-ppal-otra-dir"}),

#"Removed Columns" = Table.RemoveColumns(Source,{"id", "n\_victimas", "aaaa", "mm", "dd", "fecha", "franja\_hora", "victima", "participantes", "moto", "auto", "transporte\_publico", "camion", "ciclista", "gravedad"}),

#"Added Index" = Table.AddIndexColumn(#"Removed Columns", "Index", 1, 1, Int64.Type),

#"Reordered Columns" = Table.ReorderColumns(#"Added Index",{"Index", "direccion", "comuna", "geocodificacion\_CABA", "longitud", "latutid"}),

#"Renamed Columns" = Table.RenameColumns(#"Reordered Columns",{{"Index", "id\_direccion"}}),

#"Removed Columns1" = Table.RemoveColumns(#"Renamed Columns",{"id\_direccion"}),

#"Removed Duplicates" = Table.Distinct(#"Removed Columns1", {"direccion"}),

#"Added Index1" = Table.AddIndexColumn(#"Removed Duplicates", "Index", 1, 1, Int64.Type),

#"Renamed Columns1" = Table.RenameColumns(#"Added Index1",{{"Index", "id\_direccion"}, {"comuna", "id\_comuna"}}),

#"Replaced Value" = Table.ReplaceValue(#"Renamed Columns1","SD",null,Replacer.ReplaceValue,{"direccion", "geocodificacion\_CABA", "longitud", "latutid"})

in

#"Replaced Value"

**tipo\_vehiculo\_victima:** se separa éste dataset para tener las direcciones separado de la tabla de hechos

let

Source = Table.Combine({#"hechos-ppal-dir-norm", #"hechos-ppal-otra-dir"}),

#"Merged Queries" = Table.NestedJoin(Source, {"direccion"}, direcciones, {"direccion"}, "direcciones", JoinKind.LeftOuter),

#"Renamed Columns" = Table.RenameColumns(#"Merged Queries",{{"direcciones", "direcciones\_id"}}),

#"Expanded direcciones\_id" = Table.ExpandTableColumn(#"Renamed Columns", "direcciones\_id", {"direcciones\_id"}, {"direcciones\_id.direcciones\_id"}),

#"Removed Columns" = Table.RemoveColumns(#"Expanded direcciones\_id",{"moto", "auto", "transporte\_publico", "camion", "ciclista"}),

#"Split Column by Delimiter" = Table.SplitColumn(#"Removed Columns", "participantes", Splitter.SplitTextByDelimiter("-", QuoteStyle.Csv), {"participantes.1", "participantes.2"}),

#"Changed Type" = Table.TransformColumnTypes(#"Split Column by Delimiter",{{"participantes.1", type text}, {"participantes.2", type text}}),

#"Removed Columns1" = Table.RemoveColumns(#"Changed Type",{"victima"}),

#"Renamed Columns1" = Table.RenameColumns(#"Removed Columns1",{{"participantes.1", "victima"}, {"participantes.2", "acusado"}}),

#"Removed Other Columns" = Table.SelectColumns(#"Renamed Columns1",{"victima"}),

#"Renamed Columns2" = Table.RenameColumns(#"Removed Other Columns",{{"victima", "tipo\_vehiculo"}})

in

#"Renamed Columns2"

**tipo\_vehiculo\_acusado:** se separa éste dataset para tener las direcciones separado de la tabla de hechos

let

Source = Table.Combine({#"hechos-ppal-dir-norm", #"hechos-ppal-otra-dir"}),

#"Merged Queries" = Table.NestedJoin(Source, {"direccion"}, direcciones, {"direccion"}, "direcciones", JoinKind.LeftOuter),

#"Renamed Columns" = Table.RenameColumns(#"Merged Queries",{{"direcciones", "direcciones\_id"}}),

#"Expanded direcciones\_id" = Table.ExpandTableColumn(#"Renamed Columns", "direcciones\_id", {"direcciones\_id"}, {"direcciones\_id.direcciones\_id"}),

#"Removed Columns" = Table.RemoveColumns(#"Expanded direcciones\_id",{"moto", "auto", "transporte\_publico", "camion", "ciclista"}),

#"Split Column by Delimiter" = Table.SplitColumn(#"Removed Columns", "participantes", Splitter.SplitTextByDelimiter("-", QuoteStyle.Csv), {"participantes.1", "participantes.2"}),

#"Changed Type" = Table.TransformColumnTypes(#"Split Column by Delimiter",{{"participantes.1", type text}, {"participantes.2", type text}}),

#"Removed Columns1" = Table.RemoveColumns(#"Changed Type",{"victima"}),

#"Renamed Columns1" = Table.RenameColumns(#"Removed Columns1",{{"participantes.1", "victima"}, {"participantes.2", "acusado"}}),

#"Removed Other Columns" = Table.SelectColumns(#"Renamed Columns1",{"acusado"}),

#"Renamed Columns2" = Table.RenameColumns(#"Removed Other Columns",{{"acusado", "tipo\_vehiculo"}})

in

#"Renamed Columns2"

**tipo\_vehiculo**: dataset que contiene la tipificación de los vehiculos intervinientes en los siniestros

let

Source = Table.Combine({tipo\_vehiculo\_victima, tipo\_vehiculo\_acusado}),

#"Removed Duplicates" = Table.Distinct(Source),

#"Removed Top Rows" = Table.Skip(#"Removed Duplicates",1),

#"Added Index" = Table.AddIndexColumn(#"Removed Top Rows", "Index", 1, 1, Int64.Type),

#"Renamed Columns" = Table.RenameColumns(#"Added Index",{{"Index", "id\_tipo\_vehiculo"}}),

#"Reordered Columns" = Table.ReorderColumns(#"Renamed Columns",{"id\_tipo\_vehiculo", "tipo\_vehiculo"})

in

#"Reordered Columns"

**victimas-ppal-normalizada:** dataset principal de victimas, contiene todas las modificaciónes necesarias para la obtencion de datos

let

Source = Excel.Workbook(File.Contents("D:\coderhouse\data analytics flex\ds-lesiones-v3 - Juan Pablo Ludueña.xlsx"), null, true),

Tabla2\_Table = Source{[Item="Tabla2",Kind="Table"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Tabla2\_Table,{{"ID hecho", type text}, {"AAA", Int64.Type}, {"MM", Int64.Type}, {"DD", Int64.Type}, {"FECHA ", type date}, {"VEHICULO\_VICTIMA", type text}, {"SEXO", type text}, {"EDAD\_VICTIMA", type any}, {"GRAVEDAD", type text}}),

#"Merged Queries" = Table.NestedJoin(#"Changed Type", {"VEHICULO\_VICTIMA"}, tipo\_vehiculo, {"tipo\_vehiculo"}, "tipo\_vehiculo", JoinKind.LeftOuter),

#"Expanded tipo\_vehiculo" = Table.ExpandTableColumn(#"Merged Queries", "tipo\_vehiculo", {"id\_tipo\_vehiculo"}, {"tipo\_vehiculo.id\_tipo\_vehiculo"}),

#"Renamed Columns" = Table.RenameColumns(#"Expanded tipo\_vehiculo",{{"tipo\_vehiculo.id\_tipo\_vehiculo", "id\_tipo\_vehiculo"}}),

#"Removed Columns" = Table.RemoveColumns(#"Renamed Columns",{"AAA", "MM", "DD", "FECHA ", "GRAVEDAD"}),

#"Added Index" = Table.AddIndexColumn(#"Removed Columns", "Index", 1, 1, Int64.Type),

#"Renamed Columns1" = Table.RenameColumns(#"Added Index",{{"Index", "id\_victima"}}),

#"Reordered Columns" = Table.ReorderColumns(#"Renamed Columns1",{"id\_victima", "ID hecho", "VEHICULO\_VICTIMA", "SEXO", "EDAD\_VICTIMA", "id\_tipo\_vehiculo"}),

#"Changed Type1" = Table.TransformColumnTypes(#"Reordered Columns",{{"EDAD\_VICTIMA", type number}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type1","SD",null,Replacer.ReplaceValue,{"SEXO"}),

#"Replaced Value1" = Table.ReplaceValue(#"Replaced Value","sd",null,Replacer.ReplaceValue,{"VEHICULO\_VICTIMA"}),

#"Replaced Errors" = Table.ReplaceErrorValues(#"Replaced Value1", {{"EDAD\_VICTIMA", null}}),

#"Filtered Rows" = Table.SelectRows(#"Replaced Errors", each true),

#"Replaced Value2" = Table.ReplaceValue(#"Filtered Rows","sd",null,Replacer.ReplaceValue,{"SEXO"}),

#"Changed Type2" = Table.TransformColumnTypes(#"Replaced Value2",{{"EDAD\_VICTIMA", Int64.Type}}),

#"Replaced Value3" = Table.ReplaceValue(#"Changed Type2",null,"No Definido",Replacer.ReplaceValue,{"SEXO"})

in

#"Replaced Value3"

# **Tablas auxiliares:**

* **Tabla de medidas**
  + cantidad\_hechos =

// contamos la cantidad de registros en la tabla hechos-ppl-normalizada

COUNT('hechos-ppal-normalizada'[id])

* + cantidad\_hechos\_fatales =

// cantidad de hechos que tuvieron victimas fatales

CALCULATE(COUNT('hechos-ppal-normalizada'[id\_tipo\_lesion]), 'tipo-lesion'[gravedad] = "FATAL")

cantidad\_hechos\_graves =

// cantidad de hechos que tuvieron severidad grave

CALCULATE(COUNT('hechos-ppal-normalizada'[id\_tipo\_lesion]), UPPER('tipo-lesion'[gravedad]) = "GRAVE")

* cantidad\_hechos\_sin\_lesiones =

// cantidad de hechos con lesiones leves/nulas

CALCULATE([cantidad\_hechos] - [cantidad\_hechos\_fatales] - [cantidad\_hechos\_graves])

* cantidad\_personsas\_en\_siniestros =

// cantidad total de victimas registradas en todos los siniestros

COUNT('victimas-ppal-normalizada'[id\_victima])

* cantidad\_tipo\_vehiculo\_acusado =

// cantidad por tipo de vehiculo del acusado

COUNT('hechos-ppal-normalizada'[id\_tipo\_vehiculo\_acusado])

* cantidad\_tipo\_vehiculo\_victima =

// cantidad por tipo de vehiculo de la victima

COUNT('hechos-ppal-normalizada'[id\_tipo\_vehiculo\_victima])

* **Tabla calendario**

tabla\_calendario =

CALENDAR(

MIN('hechos-ppal-normalizada'[fecha]),

MAX('hechos-ppal-normalizada'[fecha])

)

# 

# **Modelo de datos en Power Bi**

