

ProyectoAED2024

- ¹ Universitat de València. ETSE.; alcaluis@alumni.uv.es
² Universitat de València. ETSE.; carigar4@alumni.uv.es
³ Universitat de València. ETSE.; jopebrui@alumni.uv.es

Simple Summary: Estudio de los delitos causados por menores en España entre 2013 y 2023.

Abstract: Proyecto de tratamiento de datos para la asignatura de Análisis Exploratorio de Datos. El tópico es la delincuencia en España provocado por menores de entre 14 y 17 años (ambos inclusivos). En este estudio se analizarán los delitos causados con mayor frecuencia.

Keywords: Delito; Delincuencia; Menores; Datos; Ciencia de Datos;

1. Introducción

1.1. Definición proyecto y planteamiento de preguntas (objetivos)

La delincuencia en España es un tema de conversación latente. Un estudio importante, es tener en cuenta la edad del delincuente. En este estudio analizaremos los diferentes tipos de delito ocasionados por menores de entre 14 y 17 años durante los 10 últimos años. Los objetivos del estudio es responder a las siguientes preguntas:

1. ¿Cuáles son los delitos más frecuentes y qué edades los perpetran? ¿Han aumentado o han disminuido en los últimos años?
2. ¿Cómo evoluciona la cantidad de delitos dependiendo de la edad del delincuente?
3. Analizar los picos de delitos (mínimos y máximos) y valores faltantes, relacionándolos con hechos reales. (Desde sucesos hasta leyes o reformas)

1.2. Estudio y acondicionamiento de los datos

1.2.1. Análisis de la estructura y valores iniciales de los datos.

Comenzamos ejecutando “summary” del conjunto de datos, recién cargado, para analizarlo.

```
##      delitos_n1      delitos_n2      delitos_n3      delitos_n4
## Length:2915      Length:2915      Length:2915      Length:2915
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##      delitos_n5      edad      año      total_delitos
## Length:2915      Length:2915      Min.   :2013      Min.    :  0.0
## Class :character Class :character      1st Qu.:2015      1st Qu.:  17.0
## Mode  :character Mode  :character      Median :2018      Median :  78.0
##                      Mean  :2018      Mean   : 756.2
##                      3rd Qu.:2021      3rd Qu.: 410.0
##                      Max.   :2023      Max.   :26349.0
##                      NA's    :210
```

Citation: Albacete, L; Ribes, C; Pedro, J.
ProyectoAED2024. *Journal Not Specified* **2023**, *1*, 0. <https://doi.org/>

Received:

Revised:

Accepted:

Published:

Copyright: © 2025 by the author.
Submitted to *Journal Not Specified*
for possible open access publication
under the terms and conditions
of the Creative Commons Attribution
(CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Gracias a “summary” podemos apreciar qué columnas numéricas contienen NAs y si los valores (de las columnas numéricas) entran dentro de lo esperado (min, max, median, ...).

Teniendo para la columna **año** valores esperados, desde 2013 hasta 2023 y la media obviamente en 2018. Respecto a la columna **total_delitos**, podemos apreciar que tiene un número relevante de valores faltantes (210) y confirmamos que los valores se han importado correctamente (manteniendo los millares). Además hemos identificado que la columna **edad** se ha importado como tipo “character”. Tras hacer una visualización del archivo nos damos cuenta que los años se han importado como “14 años” agregando el sustantivo para todos los valores y en algunos casos agregando un “Total”. Acondicionamiento de la columna será necesario.

Las columnas restantes, que son de tipo “character”, representan el delito en distintos niveles de clasificación. Habremos de factorizar estas variables categóricas. El uso de la función “str” o “glimpse” en este caso no nos ha aportado información adicional.

A continuación, ejecutaremos “table” por columna para analizar la cantidad de “NA” y si hay variedad en los elementos.

```
## Columna 'delitos_n1':
```

```
##
## Total Infracciones      <NA>
##           2915           0
```

```
## Columna 'delitos_n2':
```

```
##
## A Delitos  B Faltas      <NA>
##    2585      275        55
```

```
## Columna 'delitos_n3,4,5' (últimos 3):
```

```
##
## Faltas contra las personas      R Resto de delitos
##           55                    55
##           <NA>
##           165
```

```
##
##           8.5 Prostitución y corrupción menores
##                                55
## 8.99 Otros delitos contra la libertad e indemnidad sexuales
##                                55
##                                <NA>
##                                1045
```

```
##
## 22.2.1 Atentados contra la autoridad  22.2.2 Resistencia y desobediencia
##                                55                                55
##                                <NA>
##                                2695
```

```
## Columna 'edad':
```

```
##
## 14 años 15 años 16 años 17 años  Total  <NA>
##    583    583    583    583    583    0
```

```
## Columna 'año':
```

```
##
## 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 <NA>
## 265 265 265 265 265 265 265 265 265 265 265 0
```

De las diferentes columnas podemos decir que:

- **delitos_n1**: la primera columna no nos aporta valor alguno al solo tener un valor. Más adelante será eliminada.
- **delitos_n2**: esta columna nos servirá para hacer el primer filtrado y quedarnos con solo los delitos y faltas cometidas. Omitiendo los totales calculados. Sus valores pueden ser reemplazados en valores más claros.
- **delitos_n3, n4, y n5**: ayudan a clasificar los delitos en subtipos de infracción. Sin embargo, podemos ver como a mayor nivel de detalle más valores faltantes. Esto se debe a que un nivel de clasificación tan preciso no es necesario para describir todos los delitos.

La columna **delitos_n5** no será necesaria. Aporta muy poca información y el análisis no requerirá tanto nivel de detalle para una infracción.

- **edad**: todos los datos de esta columna son uniformes y esperados (ningún valor faltante). Finalmente se eliminarán aquellas instancias que representen un total y se convertirá la columna a tipo numérico.
- **total**: respecto a esta columna pese a que no hayamos utilizado ninguna instrucción, hemos mediante visualización del fichero analizado un caso en concreto para definir la naturaleza de esta columna. Ya vimos anteriormente, gracias a otras columnas, que hay agrupaciones internas de los datos calculando los totales. Pero eso no es todo. Podemos apreciar que el total de un delito, es la suma de todas sus subclasificaciones.

Table 1. Caso delitos 'Contra el patrimonio y el orden socioeconómico' en 2013 por delinquentes de 14 años

delitos_n4	delitos_n5	total_delitos
NA	NA	1433
13.1	NA	142
13.2	NA	1108
13.2	13.2.1 Robo con violencia	602
13.2	13.2.2 Robo con fuerza	506
13.4	NA	55
13.5	NA	9
13.6	NA	8
13.9	NA	97
13.1	NA	13
13.9	NA	1

Como podemos observar, el total de "13.2 Robos", es igual a la suma de "13.2.1" y "13.2.2", a su vez el delito contra el patrimonio es la suma de los delitos "13.1", "13.2", ... Este hecho agrega redundancia a los datos. Por ende no solo tendremos que eliminar las filas que representen totales, valor "total" en edad, pero también aquellas filas que son las agrupaciones de otras con el objeto de tener un conjunto **Tidy**.

En resumen, las columnas de delitos n1 y n5 no son necesarias y pueden ser eliminadas sin problemas. Ya que hemos averiguado que los delitos de N5 se suman para los delitos de N4, es decir no perdemos datos. Las columnas de clasificación de delitos restantes serán convertidas a factor. La columna edad será modificada para mayor fidelidad al dato y un adecuado uso más adelante.

1.3. Análisis de ‘outliers’, métodos de imputación aplicados y análisis de datos perdidos

1.3.1. Restricciones de Rango

Primeramente comprobaremos que los valores numéricos se encuentran dentro de los rangos esperados. Utilizando la instrucción ‘stopifnot’ comprobando los mínimos y máximos esperados.

1.3.2. Restricciones de unicidad. Duplicados (enteros y parciales)

En este apartado analizaremos si hay alguna observación faltante o repetida. Es decir, para cada delito (y sus subclasificaciones) tenemos un año y una edad específica. Para ello, vamos a contar cuantas veces aparece cada delito esperando obtener un único valor que es 44 (11 años * 4 edades * cada delito en específico). Tras realizar la prueba efectivamente obtenemos dicho valor, para todos los delitos.

Hemos averiguado que no tenemos observaciones repetidas, pero ello se puede deber a que hayan fallos tipográficos o de ortografía. Es decir, hemos considerado que “13 delito” y “13 delito” no son duplicados. En este apartado trataremos este caso si es necesario.

```
##                                     n
## delito_compl                      44
## Delito 1 Homici ... y sus formas    1
## Delito 10 Contr ... ción de secretos 1
## Delito 10 Contr ... miento de morada 1
## Delito 13 Contr ... 13.5 Usurpación 1
## Delito 13 Contr ... uso de vehículo 1
## Delito 13 Contr ... 6 Defraudaciones 1
## Delito 13 Contr ... mico 13.1 Hurtos 1
## Delito 13 Contr ... n socioeconómico 1
## Delito 13 Contr ... ómico 13.2 Robos 1
## Delito 13 Contr ... ómico 13.9 Daños 1
```

Tras echarle un vistazo a la lista de combinaciones “delito_falta” + “clasi_infrac” + “sub_clasi_infrac” podemos confirmar que no hay errores.

1.3.3. Validar totales (validación cruzada)

Para validar que los datos se han acondicionado correctamente vamos a sacar las filas totales originales y compararlas con nuestro dataframe haciendo agrupaciones de los datos y sumas de los valores.

```
## Comparación infracciones totales por año:
```

```
## año total_proporcionado total_validacion_datos
## 1 2013                25814                25814
## 2 2014                25717                25717
## 3 2015                24005                24005
## 4 2016                21986                21986
## 5 2017                23050                23050
## 6 2018                24340                24340
```

```
## Comparación delitos y faltas totales por año:
```

```
## delitos_n2 año total_proporcionado total_validacion_datos
## 1 A Delitos 2013                16646                16646
## 2 B Faltas 2013                 9168                 9168
## 3 A Delitos 2014                16012                16012
## 4 B Faltas 2014                 9705                 9705
## 5 A Delitos 2015                15779                15779
## 6 B Faltas 2015                 8226                 8226
```

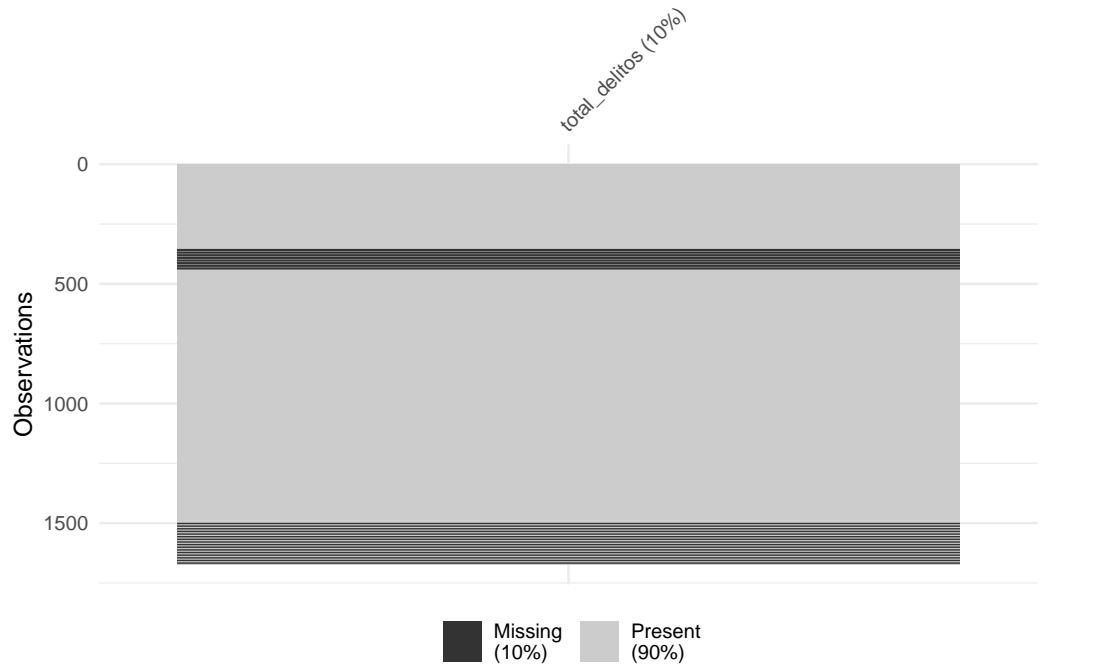
## Comparación delitos totales para el delito 13 año:				165
##	año	total_proporcionado	total_validacion_datos	166
##	1 2013	9198	9198	167
##	2 2014	8449	8449	168
##	3 2015	7809	7809	169
##	4 2016	9548	9548	170
##	5 2017	10075	10075	171
##	6 2018	10346	10346	172

Como podemos comprobar los valores coinciden entre los totales proporcionados por la fuente de datos y los nuestros calculados haciendo agrupaciones de los datos acondicionados.

Para el cálculo hemos omitido todos aquellos valores de la columna “total_delitos” que eran NA. Al obtener los mismos resultados podemos confirmar que han considerado como 0 aquellos valores faltantes.

1.3.4. Explorar valores faltantes

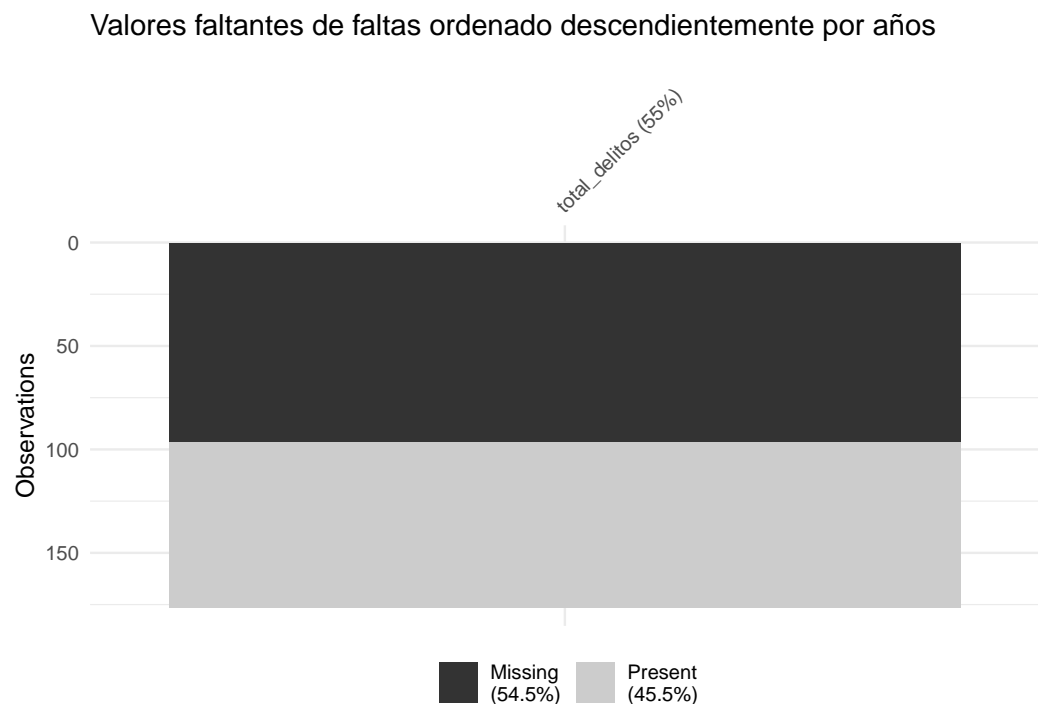
Primeramente vamos a visualizar los datos faltantes de la columna “total_delitos”; pues la otra columna que tiene valores faltantes “sub_clasi_infrac” ya está justificada. A partir de aquí vamos a tratar de razonar si estos datos son faltantes debido a un factor aleatorio o hay un razonamiento detrás.



De la visualización podemos apreciar dos grupos de datos faltantes en la columna “total_delitos”. De los casos, al estar el dataframe inicialmente organizado por los delitos, podemos decir que:

- El grupo de datos faltantes del final corresponden a las “faltas” que se convirtieron en “delitos”, hecho que analizaremos más adelante en el análisis. Si recogemos solo las faltas y organizamos por año confirmaremos este caso.
- Pero tenemos otro caso extra que analizaremos a continuación.

Primero la confirmación de las faltas:



A continuación vamos a tratar el otro grupo de datos faltantes, ignorando las “faltas” y agrupando por delitos:

## # A tibble: 6 x 2	
## clasi_infrac	na_delitos
## <fct>	<int>
## 1 8 Contra la libertad e indemnidad sexuales	72
## 2 1 Homicidio y sus formas	0
## 3 10 Contra la intimidad, derecho a la propia imagen	0
## 4 13 Contra el patrimonio y el orden socioeconómico	0
## 5 17 Contra la seguridad colectiva	0
## 6 18 Falsedades	0

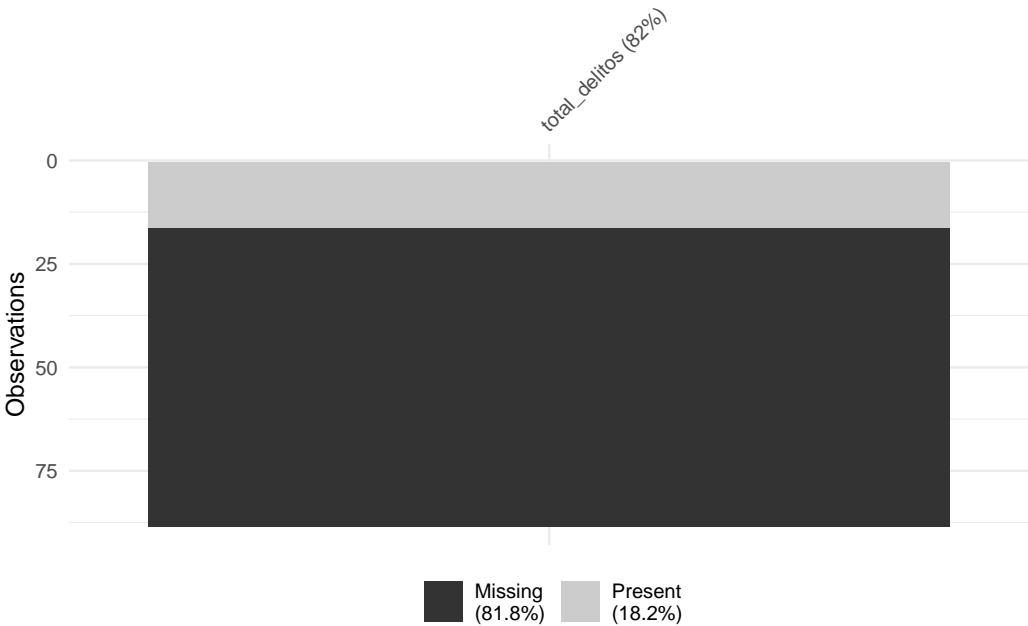
Podemos ver que todos los datos faltantes corresponden al delito “8 Contra la libertad e indemnidad sexuales”. Y dentro de los delitos contra la libertad sexual tenemos solo datos faltantes en dos delitos en concreto “BIS Abusos y agresiones sexuales a menores de 16 años” y “Prostitución y corrupción menores”

## # A tibble: 6 x 2	
## sub_clasi_infrac	na_delitos
## <fct>	<int>
## 1 8.2 BIS Abusos y agresiones sexuales a menores de 16 años	36
## 2 8.5 Prostitución y corrupción menores	36
## 3 18.2 Falsedades documentales	0
## 4 18.99 Otros delitos de las falsedades	0
## 5 8.1 Agresiones sexuales	0
## 6 8.2 Abusos sexuales	0

Como vemos hay 36 datos faltantes por cada delito. Que si los dividimos entre 4 edades, nos dan 9 años con valores faltantes.

Si ahora cogemos y los ordenamos estos dos por años podemos ver el efecto de la Ley Orgánica 10/2022 donde se incluyen en estos delitos a los menores de edad.

Valores faltantes del delito 8 ordenado descendientemente por años



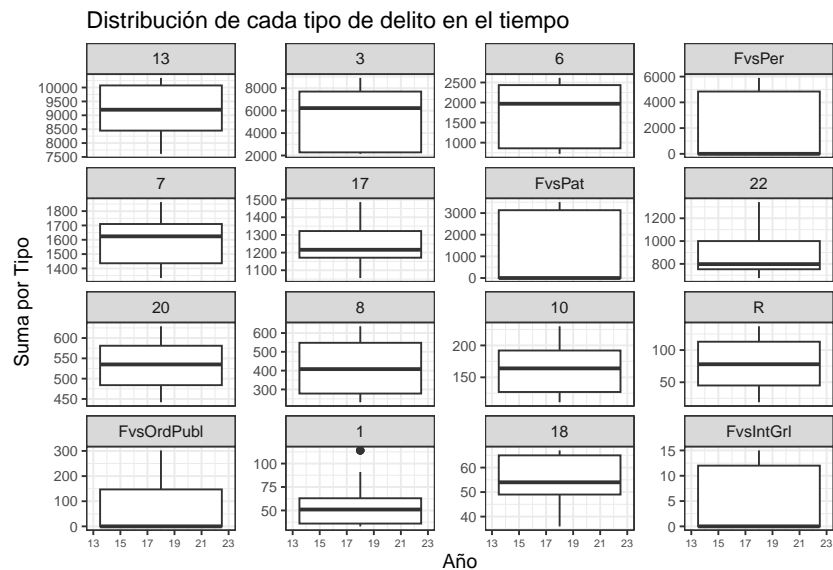
Finalmente podemos confirmar que los dos casos de faltos faltantes están justificados. Uno debido a la reforma relacionada con la clasificación de las infracciones en faltas o delitos, y el otro caso debido a LO 10/2022, donde a partir de 2022 se consideran nuevos delitos en menores, faltando datos previamente.

2. Análisis

2.1. Análisis de las variables

En primer lugar hacemos para cada categoría de la variable *clasi_infrac* (clasificación de infracción) una primera exploración de los estadísticos principales de la variable *total_delitos*: media, varianza, cuantiles, máximos y mínimos. Seguidamente dibujamos un boxplot para visualizar en primera instancia la forma de nuestro conjunto de datos (a pesar de que no hemos comprobado que la distribución de las variables sea normal):

```
## # A tibble: 6 x 6
##   id_clasi d.mean  d.sd   d.Q1 d.Q2_median d.Q3
##   <chr>    <dbl> <dbl> <dbl>      <dbl>  <dbl>
## 1 1         55.7  25.5   37         51  57.5
## 2 10        164   38.8  134        164  190
## 3 13       9111.  919.  8506       9201 9812.
## 4 17       1249.  122.  1182.       1216 1300.
## 5 18         55.4   9.79  49.5         54  63.5
## 6 20         526.   58.3  488.         535 562
```



En la tabla se puede observar que hay mucha disparidad en la escala de *total_delitos* segun la clasificación, cosa que habrá que tener en cuenta cuando se hagan las gráficas.

En lo referente al boxplot, destacan por su anormalidad (seguramente no sigan una distribución normal ni nada que se le parezca) las faltas:

- Faltas contra las personas
- Faltas contra el patrimonio
- Faltas contra el orden público
- Faltas contra el interés general

Todos estos diagramas de caja muestran su mediana en el 0, cosa que indica que almenos en la mitad de los años hubo 0 casos de este tipo de delito.

También podemos comentar que solo un diagrama muestra un outlier, y este es el asociado a 1. *Homicidio y sus formas*.

Empezamos ahora con una visualización más elaborada según la clasificación viendo su evolución en el tiempo:

2.1.1. Análisis univariante

Observamos aquí una visualización de la evolución de la cantidad de delitos a lo largo del tiempo según tipo de delito. Se pueden destacar la existencia de un valle (descenso puntual) de varios tipos de delito entre el 2020 y 2021. Estos tipos son:

- 3. Lesiones
- 6. Contra la libertad
- 7. Torturas e integridad social (++)
- 8. Contra la libertad e indemnidad sexuales
- 13. Contra el patrimonio y el orden socioeconómico
- 17. Contra la seguridad colectiva
- 18. Falsedades (++)
- 20. Contra la administración de Justicia
- R. Resto de delitos

Este decrecimiento (especialmente en el tipo 7 y 18) viene probablemente provocado por el contexto social contemporáneo a los eventos: la crisis del COVID-19 y la cuarentena. Estos sucesos provocaron la cancelación de la mayor parte del sector laboral español y de la totalidad de las escuelas, que pasaron a un formato online.

Es fácil imaginar así que estos tipos de delitos, de naturaleza más escolar y/o física, tengan un descenso en número ante la situación pandémica.

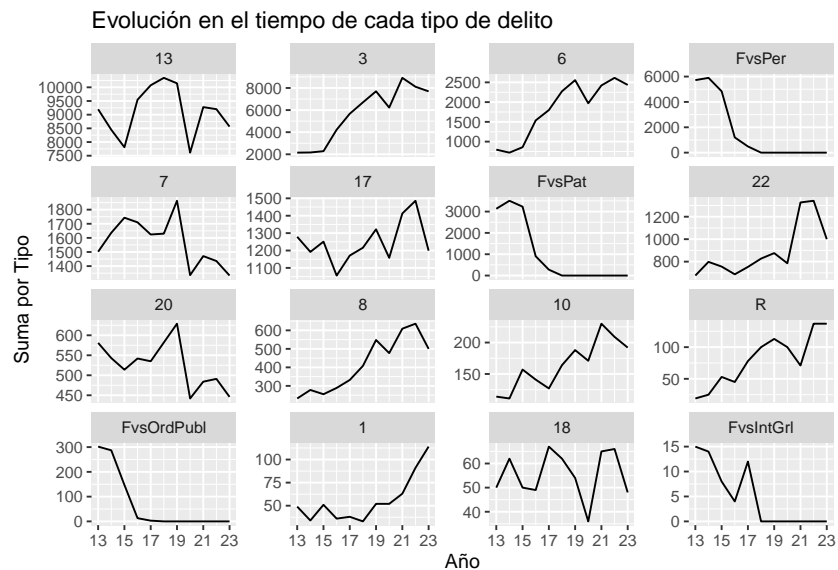


Figure 1. Evolución en el tiempo de cada tipo de delito

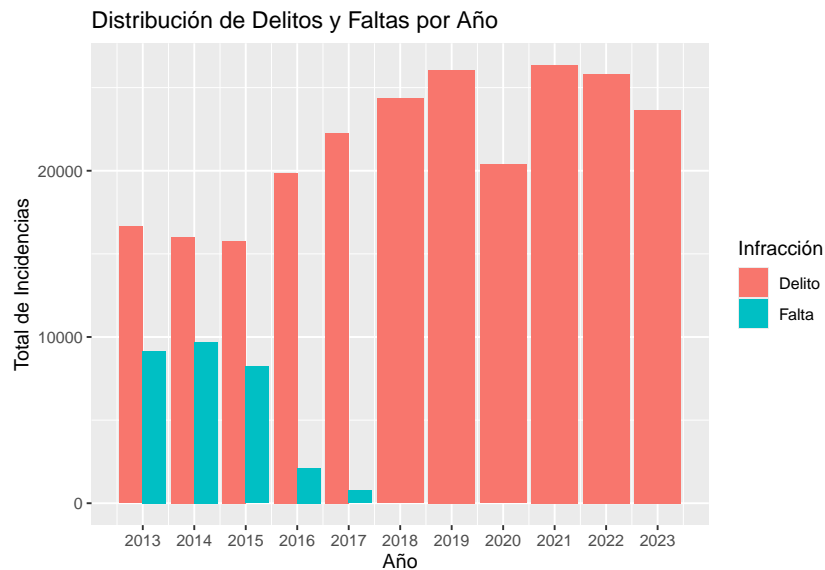
Además, apreciamos que los mas frecuentes son los delitos contra el patrimonio y el orden socioeconómico, junto con los delitos por lesiones. Además se puede apreciar un aumento en estos últimos a partir del 2016, una tendencia que también se refleja en los delitos contra la libertad. Más adelante analizaremos este fenómeno con más detenimiento.

También se observa un aumento alarmante en los delitos de homicidio en los últimos dos años. Esta tendencia podría explicarse por varias razones. Una de ellas es la mayor exposición de los menores a situaciones violentas, que puede estar relacionada con problemas sociales, la influencia de entornos familiares problemáticos o la fácil accesibilidad a contenido violento en línea. Otra posible explicación son los cambios en la legislación y el enfoque de la justicia juvenil, ya que algunos sostienen que la percepción de impunidad puede motivar a los jóvenes a cometer delitos más graves.

Asimismo, destaca el aumento de los delitos contra la intimidad y el derecho a la propia imagen. Esta tendencia probablemente está vinculada al mayor uso de dispositivos digitales y redes sociales entre menores. Se aprecia un pico en 2020, lo cual es comprensible, ya que durante el confinamiento se incrementó el uso de redes sociales, lo que pudo haber llevado a más incidentes de violación de la privacidad digital, como la difusión no consentida de contenido privado o el ciberacoso.

Es importante mencionar que hay ciertos delitos que no son significativos para su análisis debido a su escaso número de registros, lo que les resta importancia, como sucede con las falsedades y las faltas contra intereses generales.

Hemos realizado una primera exploración visual de los datos, observando cómo se distribuyen los distintos tipos de delitos según su clasificación. Ahora procederemos a analizar la evolución de las infracciones según su nivel más alto de clasificación (delito o falta), centrándonos en la variable *delito_falta*. El gráfico de barras siguiente muestra la cantidad total de incidentes registrados entre 2013 y 2023, diferenciados según si se clasificaron como delitos o faltas.

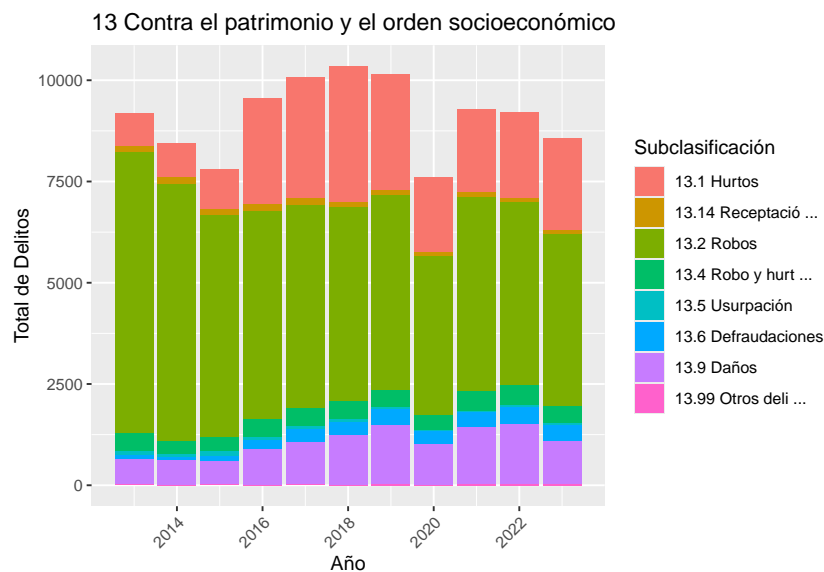


Como podemos observar, hay una clara tendencia creciente de los delitos de menores desde el 2016 hasta el 2019. En el 2020 hay una gran disminución de delitos. Esto se debe, sin duda, a la cuarentena que hubo durante la pandemia del COVID-19. Pero, a partir del 2021, notamos que la cantidad de delitos vuelve a subir, tomando unas cifras similares a las del 2019. En general, podemos observar un notable aumento de delitos juveniles en los últimos diez años. Por otra parte, notamos que el número de faltas va disminuyendo hasta el año 2017, y a partir del 2018 no hay registros sobre faltas.

Esta tendencia de aumento de delitos y disminución de faltas puede deberse a diferentes motivos. El cambio mas notable se observa entre el 2015 y el 2016. Este fenómeno se debe a la reforma del Código Penal en España del 2015, en la cual se llevó a cabo la supresión de muchas infracciones consideradas faltas, que pasaron a ser clasificadas como delitos leves. Esto explicaría el aumento de delitos en contraposición con la disminución de faltas a partir del año 2015.

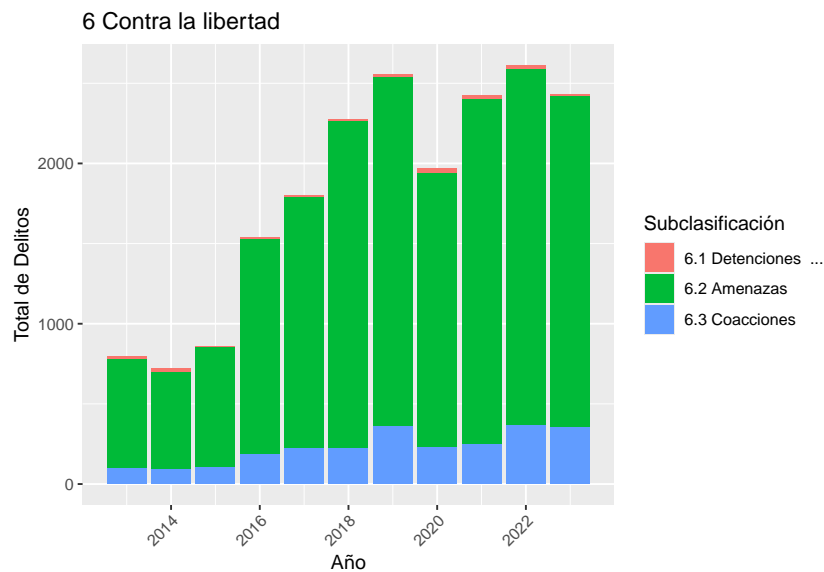
Si nos enfocamos nuevamente en el gráfico 1, se puede apreciar cómo todas las faltas disminuyen de manera drástica hasta desaparecer por completo de los registros a partir del año 2015. Por otro lado, se observa un aumento significativo en delitos como los relacionados con lesiones o contra la libertad. Esto podría deberse a que, tras la reforma mencionada anteriormente, muchas conductas como agresiones físicas leves, amenazas o coacciones, que previamente se consideraban faltas contra las personas, pasaron a clasificarse como delitos por lesiones o contra la libertad, respectivamente. Un razonamiento similar puede aplicarse a la disminución de faltas contra el patrimonio y el orden público, frente al incremento de delitos en estos ámbitos. Este cambio también explica la anomalía observada previamente en los gráficos boxplots de las faltas.

Como mencionamos anteriormente, algunos delitos se dividen en subclasificaciones debido a la amplitud de ciertas categorías. Por ejemplo, aunque el robo y el hurto son diferentes (el primero implica el uso de fuerza o amenazas, mientras que el segundo no), ambos se engloban dentro de los delitos contra el patrimonio y el orden socioeconómico. Por ello, resulta fundamental analizar cómo se distribuyen estas subcategorías. Para ello, emplearemos gráficos de barras apiladas que nos permitirán visualizar las distribuciones de las subclasificaciones más relevantes. La selección se ha realizado considerando aquellas categorías con mayor número de subclasificaciones o las que presentan índices de frecuencia más altos.



Como podemos observar, los delitos contra el patrimonio más frecuentes son los robos y los hurtos.

Conviene también analizar de forma más detallada los delitos contra la libertad, pues son de los más perpetrados y cuentan con diversas subcategorías.



Como podemos observar, los más comunes son los delitos por amenazas, los cuales van claramente en aumento a lo largo de los años, probablemente por el motivo que ya hemos expuesto anteriormente.

Finalmente, resulta relevante analizar los delitos contra la libertad e indemnidad sexuales, dada su importancia y evolución en los últimos años.

Se puede observar que este tipo de delitos ha ido claramente en aumento en los últimos diez años, aunque parece que en 2023 se frenó un poco esta tendencia. Este hecho puede analizarse desde dos puntos de vista: en primer lugar, podría indicar un incremento real en la frecuencia de agresiones sexuales, abusos, ... Sin embargo, este crecimiento en los registros podría explicarse por un mayor número de víctimas que deciden denunciar estas agresiones, lo cual reflejaría un avance en la visibilización del problema y una creciente confianza en las instituciones responsables de atender estas denuncias. Observamos que las subcategorías de agresiones sexuales y abusos sexuales tienen cifras más o menos

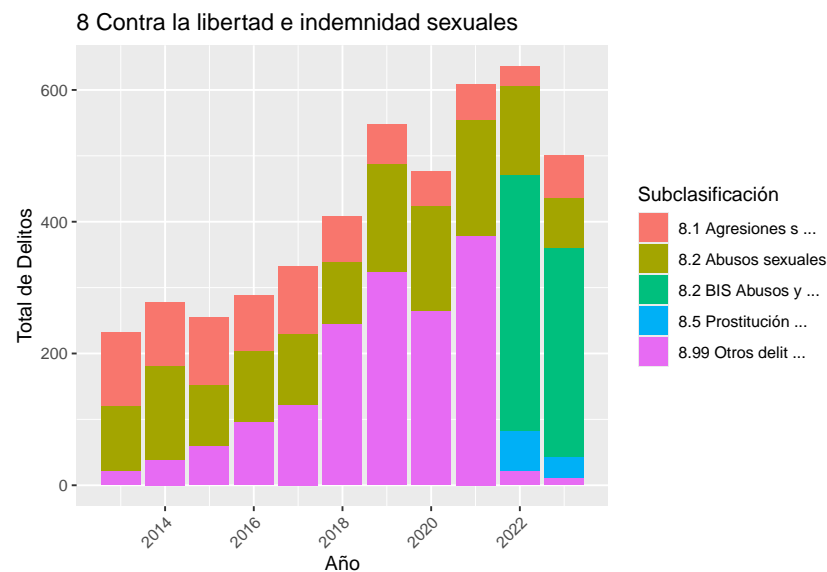
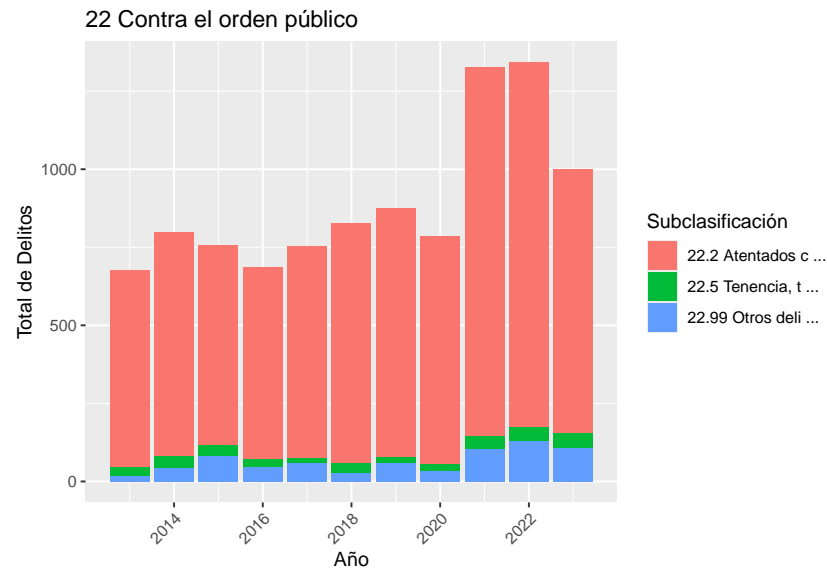


Figure 2. 8 Contra la libertad e indemnidad sexuales

similares todos los años. En cambio, la subcategoría de color rosa ha ido en aumento, hasta que en el año 2022 se añadió una subcategoría para clasificar los abusos y agresiones sexuales a menores de 16 años. Por tanto, podemos concluir que ha habido un aumento muy significativo de las denuncias de abusos y agresiones sexuales por parte de menores a menores de 16 años.

Otros delitos frecuentes entre menores son aquellos clasificados como delitos contra la seguridad colectiva. La subcategoría más común, con diferencia, son los delitos contra la seguridad vial. Ejemplos de delitos contra la seguridad vial son la conducción bajo los efectos del alcohol, la conducción temeraria y el uso de vehículos sin licencia.

Por otra parte, en los delitos contra el orden público, se puede observar un pico en los años 2021 y 2022.



Estos picos podrían estar vinculados a varios factores derivados del contexto post-pandemia. Durante el confinamiento, las restricciones de movimiento y otras medidas sanitarias generaron tensiones entre autoridades y ciudadanos, con numerosos episodios de resistencia frente a las normativas restrictivas. La relajación de las restricciones en 2021 llevó a un incremento generalizado de la actividad social, lo cual también se pudo

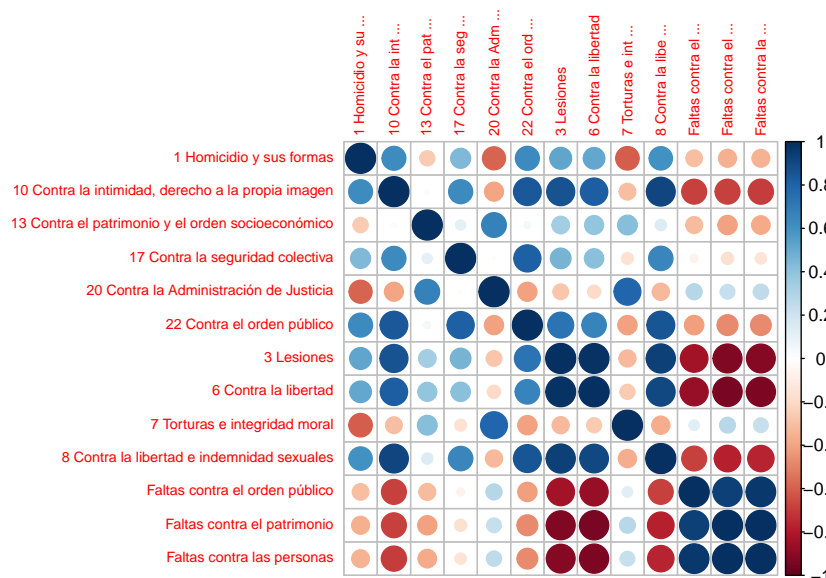


Figure 3. Correlaciones delitos más relevantes

traducir en un aumento de ciertos tipos de delitos, como altercados y confrontaciones con las autoridades.

2.1.2. Analisis bivariante

A continuación, nos enfocaremos en el análisis bivariante. Es decir, analizaremos las relaciones entre las variables para detectar patrones que nos permitan describir la evolución de la delincuencia juvenil a lo largo del tiempo. Para ello, realizaremos visualizaciones que ilustrarán estas relaciones facilitando su interpretación.

En primer lugar, realizaremos un análisis de las relaciones entre los tipos de delitos por clasificación. Tenemos un total de 16 clasificaciones diferentes, algunas de las cuales cuentan con subclasificaciones. Como se observa en las gráficas anteriores, la mayoría de las clasificaciones con subcategorías tienen un número reducido de estas, y suele predominar una subclasificación en particular. Por ello, no resulta particularmente relevante realizar el análisis usando las subclasificaciones. Sin embargo, en el caso de los delitos contra el patrimonio y el orden socioeconómico, hay un total de 8 subcategorías, y las frecuencias de cada una de ellas están más distribuidas. Dado que esta es, además, la clasificación con mayor índice de frecuencia, resulta interesante realizar un análisis bivalente con las subclasificaciones de este tipo de delito. Por consiguiente, primero llevaremos a cabo un análisis bivalente entre clasificaciones para examinar la correlación entre los distintos tipos de delitos, y posteriormente haremos lo mismo para las subcategorías de los delitos contra el patrimonio.

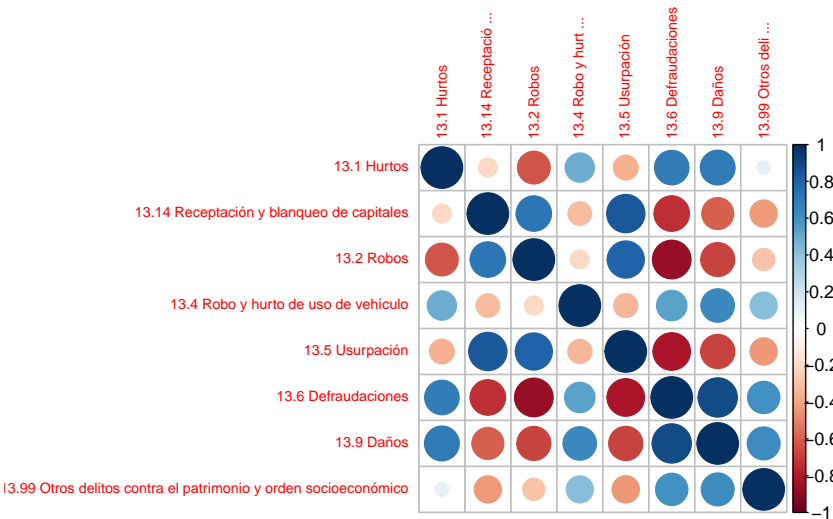
Para llevar a cabo el análisis por clasificaciones, hemos optado por eliminar las categorías menos relevantes. En el gráfico 1 se puede observar que tanto los delitos de falsedades como las faltas contra intereses generales presentan una frecuencia muy baja en los últimos diez años, lo que reduce su relevancia para el análisis. Además, la categoría “Resto de delitos” también muestra una baja frecuencia y aporta poca información al no corresponder a un tipo específico de delito. Por estos motivos, hemos decidido excluir estas categorías, y así poder centrar el análisis en los tipos de delito con mayor impacto y especificidad.

Mediante el cálculo de una matriz de correlación hemos realizado el gráfico 3, que muestra todas las correlaciones entre cada tipo de delito.

Notamos que las correlaciones más significativas se presentan entre las faltas. Esto tiene mucho sentido, ya que, como mencionamos anteriormente, las faltas han ido disminuyendo desde la reforma del código penal en 2015 hasta llegar a desaparecer en los

años posteriores. También es notable la alta correlación entre los delitos por lesiones y los delitos contra la libertad. Estos dos tipos de delitos también muestran una correlación muy baja con las faltas contra las personas, lo que refuerza la idea de que las faltas contra las personas se reclasificaron como delitos por lesiones y contra la libertad tras la reforma. Esto explicaría su aumento paralelo junto con la disminución de las faltas. Además, es relevante señalar la fuerte correlación entre los delitos por lesiones y los delitos contra la libertad e indemnidad sexuales. Esto puede deberse a la naturaleza de estos delitos. Ambos tipos de delito a menudo involucran violencia física, lo que puede llevar a una relación directa en su ocurrencia. Por ejemplo, los delitos de lesiones pueden ocurrir en el contexto de delitos sexuales, donde la violencia puede ser un factor predominante.

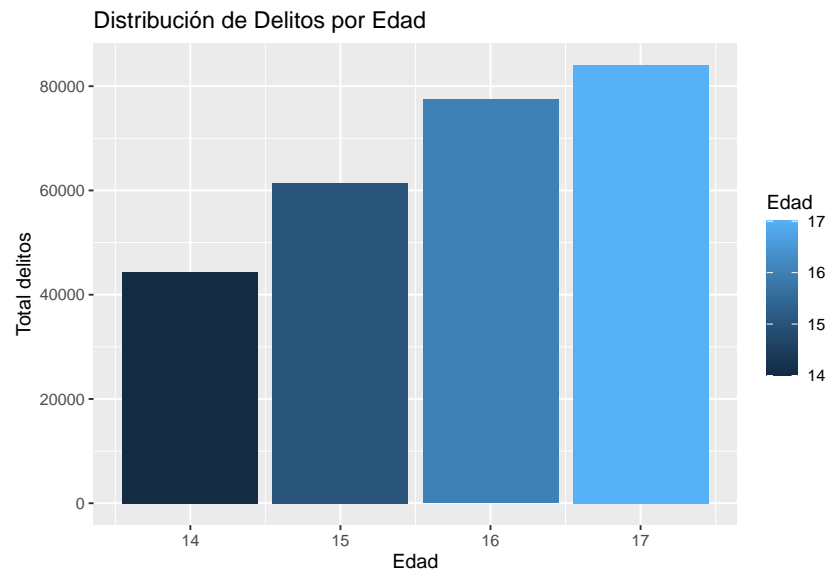
A continuación, realizaremos el mismo análisis para las diferentes subclasificaciones de los delitos contra el patrimonio y el orden socioeconómico.



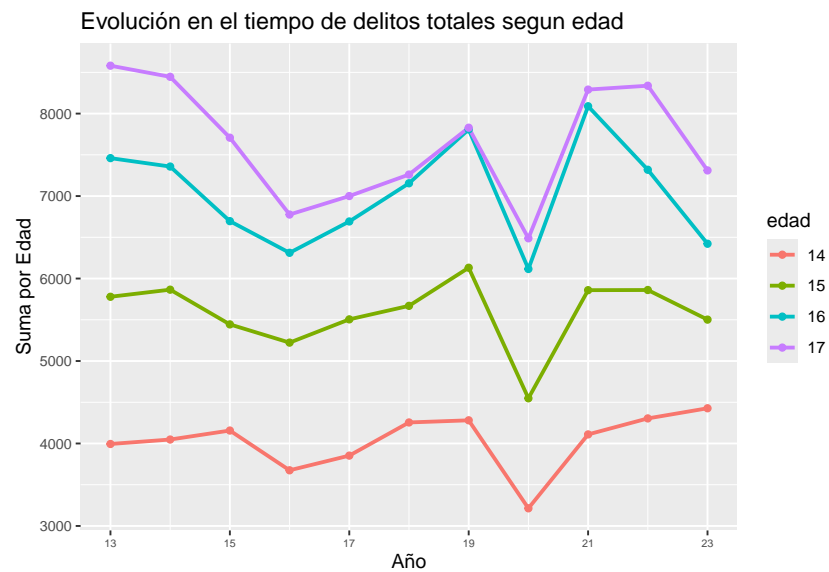
Por lo general, no se observan correlaciones tan altas como en el gráfico anterior. Podemos destacar una alta correlación entre delitos de usurpación y robos. Esto puede deberse a que la usurpación (ocupación ilegal de propiedades) y los robos suelen ocurrir en contextos donde hay vulnerabilidades en la seguridad de viviendas, locales u otros espacios. Es posible que en algunos casos, los delitos de usurpación estén relacionados con robos, ya que la ocupación ilegal de un espacio podría facilitar el acceso a bienes o recursos presentes en el lugar. Por ejemplo, una persona que usurpa una propiedad podría incurrir también en el robo de bienes de ese inmueble. También podemos destacar una alta correlación entre daños y hurtos. Estos delitos a menudo son cometidos por jóvenes en situaciones de rebeldía o frustración social. Ambos delitos pueden estar vinculados a un patrón de comportamiento de vandalismo o robo.

A continuación, analizaremos las relaciones entre las variables de edad y tipo de delito, con el objetivo de identificar qué delitos son más comunes en función de las distintas edades. Para llevar a cabo este análisis, nos centraremos en los delitos clasificados por subcategorías. Es decir, en aquellos delitos que presentan subclasificaciones, utilizaremos estas para el estudio, mientras que para los delitos que no las tienen, emplearemos su clasificación general.

En primer lugar, resulta interesante observar en qué rangos de edad se registra un mayor número de delitos. Para ello, podemos consultar el siguiente gráfico de barras.



Observamos que el gráfico revela una clara tendencia creciente, lo que indica que, a medida que los menores se acercan a la mayoría de edad, tienden a delinquir con más frecuencia. Podemos también visualizar la distribución en el tiempo de esta distribución: Dicho esto, un primer gráfico que nos puede dar un buen entendimiento de cómo se evoluciona el total de delitos en el tiempo, pero teniendo en cuenta la edad es el siguiente:



De nuevo se hace evidente el valle ocasionado por la pandemia y también una aparente correlación lineal positiva entre la edad y el total de delitos. Asimismo, podemos analizar qué delitos son más frecuentes en función de la edad de los delincuentes. A continuación, presentaremos un gráfico que ilustra los diez delitos más frecuentes en cada grupo de edad y su respectiva frecuencia.

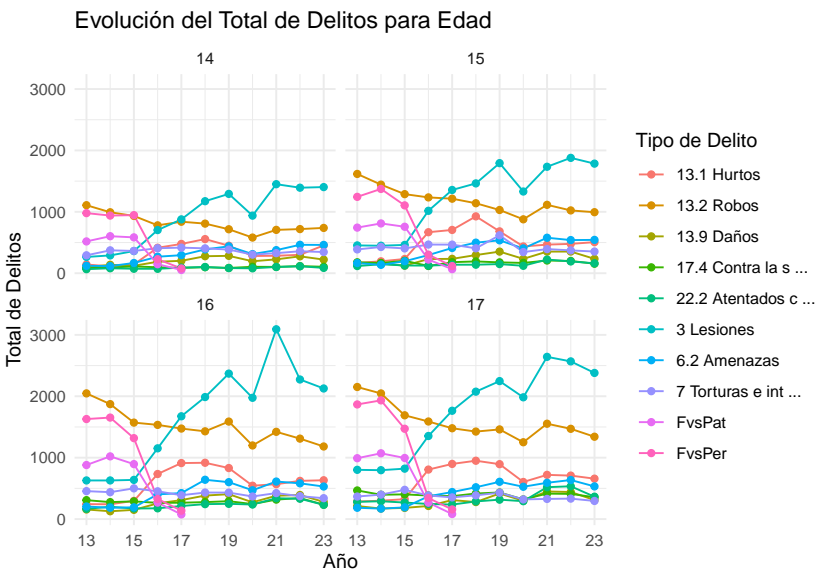
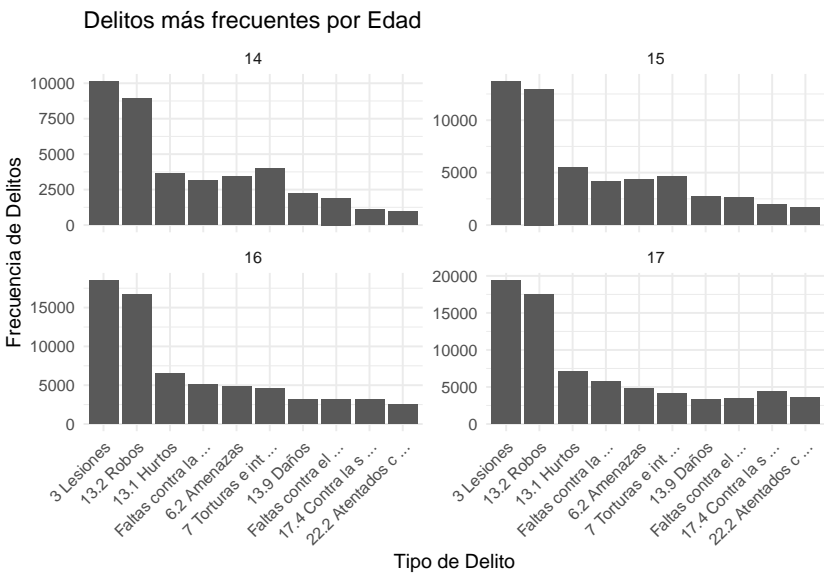


Figure 4. Evolución del Total de Delitos para Edad



Notamos que la mayoría de delitos siguen una distribución similar. Además, se evidencia un aumento en la frecuencia de los delitos conforme avanzamos en los grupos de edad. Los delitos más comunes en todas las edades son, sin duda, los delitos por lesiones y los robos, seguidos por los hurtos.

Por último, examinaremos la evolución de cada uno de estos delitos a lo largo de los últimos diez años mediante el gráfico 4, para cada grupo de edad.

Se observa que todos los delitos que se repiten para cada grupo de edad, que son la mayoría, siguen una tendencia similar. Se puede observar el aumento de delitos por lesiones en todos los grupos. Además, los robos tienden a disminuir de manera constante en todas las edades. Los hurtos, en cambio, alcanzan sus picos entre 2016 y 2019, para luego descender. En cuanto al resto de delitos, tienen menor frecuencia y se mantienen bastante estables a lo largo del tiempo.

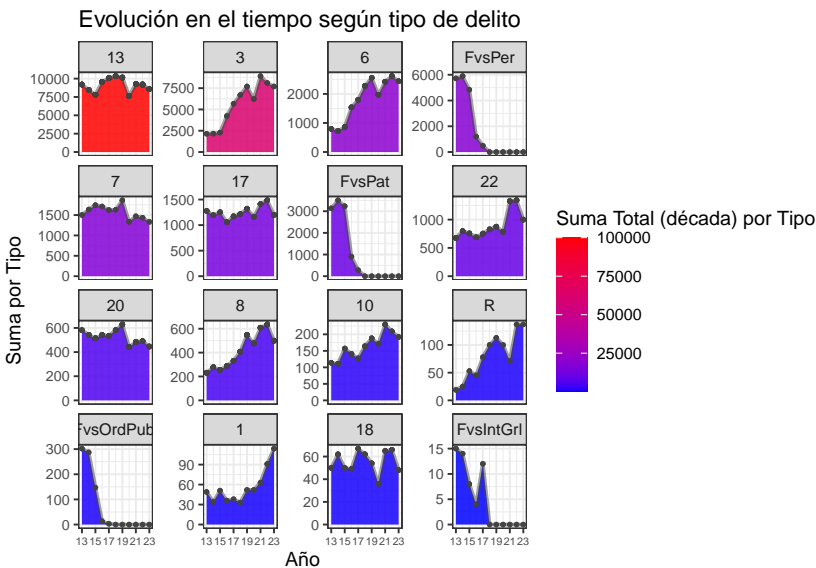
2.2. Visualización Explicatoria y conclusiones

Durante este apartado resolveremos aquellas preguntas que nos planteamos al principio del proyecto.

2.2.1. 1. ¿Cuáles son los delitos más frecuentes y qué edades los perpetran? ¿Han aumentado o han disminuido en los últimos años?

Analizando primeramente la cantidad de delitos sin centrarnos en la edad, podemos ver que los delitos más relevantes son los siguientes:

- 13. Contra el patrimonio y el orden socioeconómico. Que incluyen, robos, hurtos, defraudaciones, ... Es un delito bastante amplio.
- 3. Lesiones. Este incluye todo tipo de lesiones ocasionadas por el menor.
- 6. Contra la libertad. Incluyendo coacciones, amenazas y secuestros.
- Faltas contras las personas. Este, como ya hemos analizado, carece de relevancia actual así que nos centraremos en los otros 3. Ya que su importancia se ve reflejada en los delitos 3 y 6.



Gracias al siguiente gráfico animado (comprobar carpeta “figuras/delitos_relevantes.gif”) podemos comprobar que los delitos más ocasionados a lo largo de los años son el 13 y el 3. Siendo más frecuentados a mayor edad del delincuente, teniendo prácticamente la misma cantidad de delitos entre delincuentes de 16 y 17 años.

Si evaluamos los últimos 5 años, podemos ver el efecto del covid y parece que tiene una tendencia a descender. Sin embargo, si comprobamos con el 2013 vemos que prácticamente se han duplicado la cantidad de delitos.

2.2.2. 2. ¿Cómo evoluciona la cantidad de delitos dependiendo de la edad del delincuente?

Si echamos un vistazo a la gráfica 4, podemos apreciar que las edades parecen dividirse en dos grupos según la edad, 14 y 15 por un lado; y 16 y 17 por otro. Aunque para todos la distribución de los delitos es similar, no lo es la cantidad. Siendo superior en 16 y 17 años.

Sin embargo, podemos apreciar unas tendencias:

- Las lesiones han crecido debido a las faltas. Pero en los últimos dos años están disminuyendo.
- Los robos están disminuyendo.
- Los hurtos que habían crecido en 2016-2017 han disminuido y se han estabilizado.
- Los demás delitos parecen estar estabilizados.

Podemos concluir de esta cuestión que la edad del delincuente no es un factor determinante en qué delitos se cometen o dejan de cometer pero sí en la cantidad.

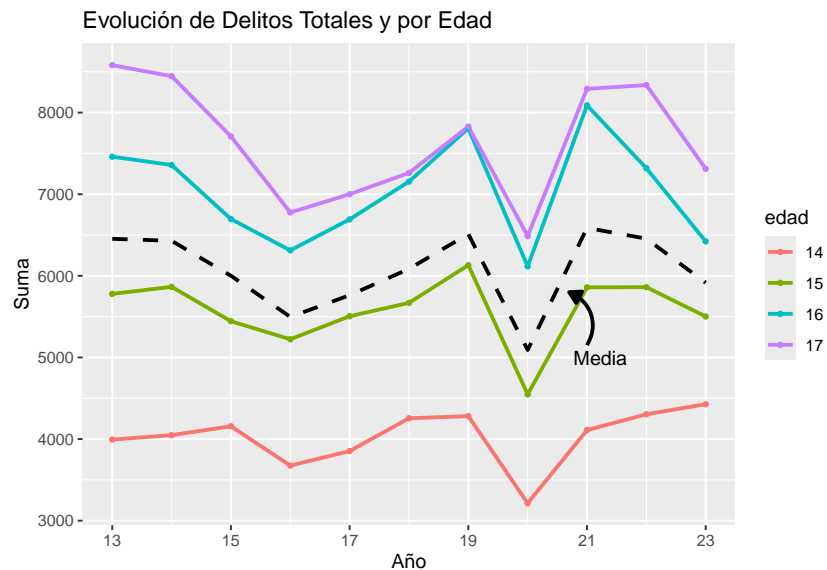


Figure 5. Evolución Delitos totales

2.2.3. 3. Analizar los picos de delitos (mínimos y máximos) y valores faltantes, relacionándolos con hechos reales. (Desde sucesos hasta leyes o reformas)

Tras el estudio exploratorio hemos podido diferenciar claramente un suceso y dos leyes o reformas.

- Covid-19. Teniendo una gran decaimiento en la cantidad de delitos en 2020-2021, como podemos apreciar en la siguiente gráfica 5.
- Reforma del código penal el año 2015. Retirando las Faltas y considerandolas delitos. Apareciéndose claramente al analizar el gráfico de correlaciones 3. Donde están fuertemente correlacionadas, de manera inversa, las faltas con los delitos 3 y 6.
- Ley Orgánica 10/2022. Agregando dos nuevos delitos “Contra la libertad e indemnidad sexuales” debido al crecimiento de los delitos sexuales 2 en los últimos años requiriendo una clasificación y tratamiento personalizado.

Para finalizar y con el objetivo de un mayor entendimiento del tema, proporcionamos un archivo R (en /shiny/shiny_ProyectoAED2024.R) con una app realizada mediante SHINY. Ejecutando la aplicación, podréis interactuar con dos gráficos. El primero os permitirá graficar los delitos más frecuentes a partir de la edad del infractor. El segundo os ayudará a visualizar todos los delitos y sus clasificaciones a lo largo de los años.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.