

Realizamos una revisión de las variables con datos vacíos y encontramos que la variable de grupo etario tenía 375 valores faltantes, los cuales imputamos utilizando la moda. Además, identificamos 121,850 valores vacíos en la variable de fecha.

En cuanto a la variable de municipio, detectamos inconsistencias en la denominación de Bogotá, donde aparecían distintas variantes del nombre ("**BOGOTÁ D.C. (CT)**" y "**Bogotá D.C. (CT)**"). Para estandarizar los datos, unificamos ambos en "**Bogotá D.C. (CT)**". De manera similar, encontramos diferencias en la escritura de **Soacha**, con registros como "**Soacha**" y "**SOACHA**", por lo que unificamos todos bajo el mismo nombre.

También hallamos inconsistencias en la codificación de las variables de género y grupo etario, donde los casos no reportados se registraban con múltiples variantes, como "**NO REPORTA**", "**NO REPORTADO**", "**NO RESPORADO**", "**-**", "**NO REPOTADO**", y "**NO-RESPORTADO**". Estas diferencias fueron corregidas para mantener la uniformidad en los datos.

Clasificamos los diez municipios con mayor número de incidentes y encontramos que **Bogotá** lidera con más de **80,000 casos**, seguido de **Soacha** con cerca de **20,000 casos** y **Fusagasugá** en tercer lugar con aproximadamente **4,000 casos**. Los siguientes municipios presentan entre **1,500 y 3,000 incidentes**, en orden descendente: **Facatativá**, **Zipaquirá**, **Girardot**, **Chía**, **Mosquera**, **Funza** y **Madrid**.

En cuanto a los incidentes por género, identificamos que **los hombres** representan la mayor cantidad de casos, con aproximadamente **120,000 registros**, mientras que **las mujeres** suman poco más de **80,000**. Los casos sin información sobre el género son mínimos y prácticamente insignificantes.

El análisis por grupo etario muestra que los **adultos** son los más afectados, con **81%** del total de incidentes. Les siguen los **adolescentes**, con **13%**, y los **menores**, con **6%**. Los casos sin información sobre el grupo etario representan menos del **1%**.

Analizando la evolución de los incidentes a lo largo de los años, observamos que entre **2010 y 2015**, los casos oscilaron entre **6,000 y 8,000 anuales**. Durante **2016, 2017 y 2018**, hubo un incremento significativo, con cifras entre **18,000 y 20,000 casos por año**. En **2019**, los incidentes disminuyeron a aproximadamente **15,000 casos**, mientras que en **2020** la cifra cayó a **10,000**. En **2021 y 2022**, se presentó un repunte, con **17,000 incidentes anuales**, y **2023** se convirtió en el año con el mayor número de incidentes hasta la fecha del estudio, con más de **22,000 casos**.

Por último, el análisis de conductas reportadas muestra que las principales categorías son **sin empleo de armas** (28%), **armas contundentes** (25%),

arma blanca o cortopunzante (21%) y en vehículos (11%). Otras categorías menos frecuentes incluyen **casos sin reporte de conducta, moto, arma de fuego, alimentos vencidos, perro y bicicleta**, las cuales representan menos del **10%** del total de incidentes.