

Introducción

El presente análisis exhibe una proyección de delitos en base a los datos del 2016 a 2019 que se publican en el portal de datos del Gobierno de la Ciudad de Buenos Aires. A su vez, a fin de incorporar una dimensión que permita enriquecer dicha proyección se decidió incorporar una clasificación de los delitos por radio censal. Esto nos permitió obtener un mayor nivel de precisión sobre el análisis realizado.

En el análisis de los datos de los periodos elegidos, se pudo observar que los delitos que mayor ocurrían eran Robo y Hurto. Se profundizó en identificar cuál era la comuna con mayor cantidad de delitos, siendo la comuna 1 y ubicando con el más del doble que la media y la mediana de C.A.B.A. por comuna. También realizamos un top 10 por de barrios

63.208

Delitos cometidos
2016-2019 en la
comuna 1

Análisis Exploratorio de datos



Los barrios con más delitos son Palermo, que registra 38.338, y Balvanera, con 33.749 (más del triple que la media registrada en C.A.B.A. por barrio que es de 10.004).

Análisis y visualizaciones

El barrio con mayor incidencia de delitos/población es San Nicolás. Esto puede deberse a la elevada circulación de personas que hay en ese barrio durante los días de semana, personas que trabajan allí y que provienen de otros barrios y de otras ciudades del AMBA.



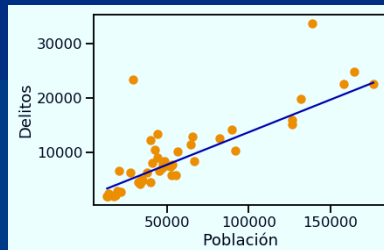
Conclusión

Para concluir el análisis, se realizó un algoritmo de **k-means** para identificar **clusters** utilizando como features **tipo de delitos**, **franja horaria** y **población por fracción censal**. Es decir, se asignó un cluster a cada una de las fracciones censales. A partir de realizar el método del codo, establecemos $k=7$.

El Silhouette index que se obtiene es de 0.73, lo que significa que la similitud entre muestras intra-cluster es alta y que la similitud inter-cluster es baja.

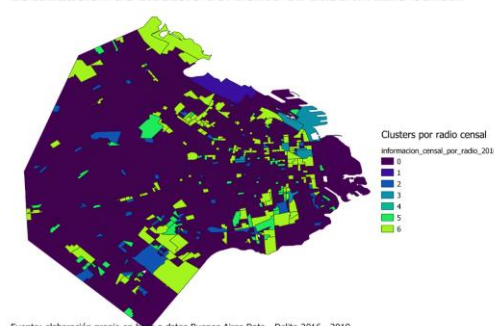
Modelos

- Se hizo un **scatterplot** para ver si hay una relación entre la población de cada barrio y los delitos



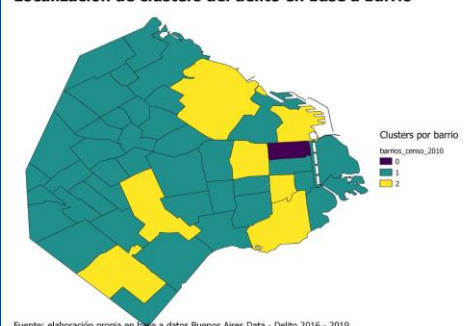
- Realizamos una **regresión lineal** entre ambas variables obteniendo un R^2 de 0.62

Localización de clusters del delito en base a radio censal



Fuente: elaboración propia en base a datos Buenos Aires Data - Delito 2016 - 2019

Localización de clusters del delito en base a barrio



Fuente: elaboración propia en base a datos Buenos Aires Data - Delito 2016 - 2019