

Capstone Project - The Battle of Neighborhoods (Week 1)

2. Data

The data used in this project was obtained from an official page of Mexico City, the database used is "pgjcdmx.csv", which contains the Street Level Crime Investigation Folders for the first half of 2019.

- **año_hechos:** crime year
- **mes_hechos:** crime month
- **fecha_hechos:** crime date
- **delito:** specific crime
- **categoria_delito:** crime classification
- **fiscalía:** place of investigation
- **agencia:** agency where the investigation was conducted
- **unidad_investigacion:** investigation unit
- **colonia_hechos:** neighborhood where the crime occurred
- **alcaldia_hechos:** borough where the crime occurred
- **fecha_inicio:** investigation start date
- **mes_inicio:** investigation start month
- **ao_inicio:** investigation start year
- **calle_hechos:** street where the crime occurred
- **calle_hechos2:** second street where the crime occurred
- **longitud:** longitude of the crime
- **latitud:** latitude of the crime
- **Geopoint:** the latitude and longitude of the crime

This database will seek to find the boroughs with the lowest number of crimes. Subsequently, a borough will be selected to find the neighborhoods with the highest number of crimes, the neighborhoods with the lowest number of crimes and the category of crimes that occurred most frequently.

Finally, the coordinates of the neighborhoods will be obtained using Google Maps API geocoding to create a new dataset that will be used to generate the venues for each neighborhood using the Foursquare API.