**Pittsburgh Crime Data Project**

Scope of Analysis: We will be collecting data on Pittsburgh Crime (2016-2019) and how it relates to:

* Weather (rain, snow, sunny, when is there more crime?)
* Crime frequency vs Incident location (which parts of Pittsburgh have the highest crime rates? Does location matter?)
* Analyzing the relationship between how severe a crime is and where it happens (do more severe crimes happen in the same location(s), crime frequency in city zones vs police zones, which zones are “safer”? is there a pattern?)
* Crime vs time (time of day, time of the week, time of the year, we want to recognize when these crimes are actually happening).
* What percentage of each age group is responsible for crimes in Pittsburgh (what about gender or race?)

Resources

* Pittsburgh Crime Data: <https://catalog.data.gov/dataset/pittsburgh-police-arrest-data>
  + Important columns are age, gender, race, incident location (columns for neighborhood, city zone, district, etc.) date, time, number of and type of offenses.
* Openweather API: <https://openweathermap.org/city/5206379>
  + Provide historical weather data to be matched with each incident of crime
* <https://www.census.gov/topics/population/age-and-sex/data/tables.html>
* <https://www.census.gov/topics/business-economy.html>
  + Possible graphing of economic levels in each neighborhood to determine crime frequency relationship.
* <https://www.census.gov/topics/employment.html>
  + Does unemployment vs employment play a part in crime occuring.
* <https://www.census.gov/topics/families.html>
* <https://www.census.gov/topics/income-poverty.html>
* <https://www.census.gov/topics/housing.html>
* <https://www.census.gov/topics/population/race.html>
  + How does race and population size/density compare to frequency/type of crime
* <https://www.census.gov/programs-surveys/ahs/data/interactive/ahstablecreator.html>

Potential Graphs:

* Weather vs Crime Frequency / Crime Type

1. **Temperature (Hot vs Cold)**
   * + X-Axis: Temperature (Ex. Very Cold, Cold, Moderate, Hot, Very Hot, etc.)
     + Y-Axis: Amount of Crime (measured by incident number)

**2.) Weather Condition:**

* + - X-Axis: Type of weather (ie rain, sun, snow, etc.)
    - Y-Axis: Amount of Crime (measured by incident number)

**3.) Crime Frequency vs Incident Location (Heatmap) (Liam)**

* **Crime Frequency vs Police Zone (Heatmap) (Liam)**

**4.) Crime Severity vs Incident Location (Heatmap)**

* Age, Gender and Race vs Crime (Pie Charts?)

**5.) Crime Frequency**

* + - Grouped by incident location & date

**6.) Crime Type**

**7.) Crime Severity (Felonies, Misdemeanors or Violations/Infractions)**

* Crime vs Time (Bar Graph & Scatter Plots)

**8.) Crime Frequency vs Time of Year (seasonality)**

* + - X-Axis: Months
    - Y-Axis: Amount of Crime (measured by incident number)

**8.5) Crime Severity vs Time of Year (seasonality)**

* + - X-Axis: Months
    - Y-Axis: Severity of Crime (measured by incident number)

**9.) Crime Frequency vs Day of Week**

* + - X-Axis: Day (Monday, Tuesday, Wednesday, etc.)
    - Y-Axis: Amount of Crime (measured by incident number)

**9.5) Crime Severity vs Day of Week**

* + - X-Axis: Day (Monday, Tuesday, Wednesday, etc.)
    - Y-Axis: Severity of Crime (measured by incident type)

**10.) Crime Frequency vs Time of Day**

* + - X-Axis: Time of Day (early morning, morning, early afternoon, noon, etc.)
    - Y-Axis: Amount of Crime (measured by incident number)

**10.5) Crime Severity vs Time of Day**

* + - X-Axis: Time of Day (early morning, morning, early afternoon, noon, etc.)
    - Y-Axis: Severity of Crime (measured by incident type)

DataFrames

* Original Pittsburgh Crime Data
* Clean Pittsburgh Crime Data
  + Split up Offenses column
  + Split up Arrest Time column
  + Drop unnecessary columns not needed in analysis
* Historical Pittsburgh Weather DataFrame
  + Organized with each row showing a days weather.