**Project title:** Florida Gun Violence incidents by city / county and income, unemployment, and poverty statistics

**Group members:** Jason Amaya, Andres Sosa, Iñaki Domaica

**Project Description. Draft**

Perform an analysis of incidents of gun violence according to official statistics throughout the country (239,677), filter them for the state of Florida (15,029) and relate them geospatially (cities, towns and counties) to variables such as income, unemployment rate and poverty.

**SOURCES - Datasets to be used**

Kaggle:

<https://www.kaggle.com/gunviolencearchive/gun-violence-database>

https://www.kaggle.com/ericking310/us-gun-violence

US Census Bureau:

Median Household Income by Town: <http://data.ctdata.org/dataset/median-household-income-by-town>

U.S. DEPARTMENT OF AGRICULTURE, Economic Research Service:

<https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/>

Poverty Statistics:

<https://www.ers.usda.gov/webdocs/DataFiles/48747/PovertyEstimates.xls?v=1975.4>

U.S. DEPARTMENT OF LABOR, Bureau of Labor Statistics

Unemployment Labor force data by county, 2019:

<https://www.bls.gov/lau/>

<https://www.bls.gov/lau/#tables>

https://www.ers.usda.gov/webdocs/DataFiles/48747/Unemployment.xls?v=1975.4

CITIPY:

<https://pypi.org/project/citipy/>

**EXTRACT**

By Jupyter Notebook and Pandas dataframe functions

Check dependencies as:

* Datetime. Formatting Dates
* SQLAlchemy (create\_engine). Load Dataframes

**TRANSFORM**

Cleaning of fields or variables not relevant

Dropping

Duplicates

NaN

Changes of type format

Re-indexing, etc

**LOAD**

Probably 2 different databases with different loads in tables, collections:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name of table / collection** | **Tables / Collections** | **Type of DB** | **Used** | **Reason** |
| Florida\_Gun\_Violence\_detail | By incident | SQL | Postgres | Structured data |
| Florida\_Gun\_Violence | By county or city | SQL | Postgres | Structured data |
| Florida\_Detailed\_incidents | By incident | nonSQL | MongoDB | At least some unstructured data |

**POTENTIAL ENHANCEMENT PROJECT TASKS. Self-BONUS**

1. Flask with app to enter detailed info per incident
2. Plot map with top five shooting incidents in Florida

**ROUGH BREAKDOWN OF TASKS:**

1. Preprocessing data
2. Extract
3. Transform
4. Load