Homework Week 9 18/10/2020

Rohan Tan 27804836

Joe Liu - Wed 2pm Clayton

url: https://ailuropanda.github.io/FIT3179/Homework%20Week%209/

Allow for some loading time please as there is currently an issue with the join mechanic.

Domain of Data Visualisation

Washington DC and its various economic variables that affect crime and each other, (Housing, Income, Poverty, Population)

Visualised dataset

The dataset is from https://opendata.dc.gov/ more specifically the data comes from DCGISopendata.

Data transformation

The population data and crime data had a different naming convention between the Census Tracts which was the area of the crime committed to convert this we utilised a python script to get the same naming convention e.g 000100 becomes Census Tract 1 and 001401 becomes Census Tract 14.01. Therefore joining data by specific area.

To further improve this I will need to normalise the crimes by population in each area as higher population will generally give higher crime which is misleading when shown on the current graph.

Data classfication

None Currently

Justification

I utilised two different maps currently as both the maps are not fully completed currently, the first map the dot map shows each crime in it's exact location whilst also showing the color of the specific crime allowing for a more precise analysis of each crime and its location and details (This utilises neighbourhood clusters instead of census clusters for geolocation).

The second map is a chloropleth map which will be used eventually to compare different forms of economic variables and their respective crime rates within those regions.

Screenshots

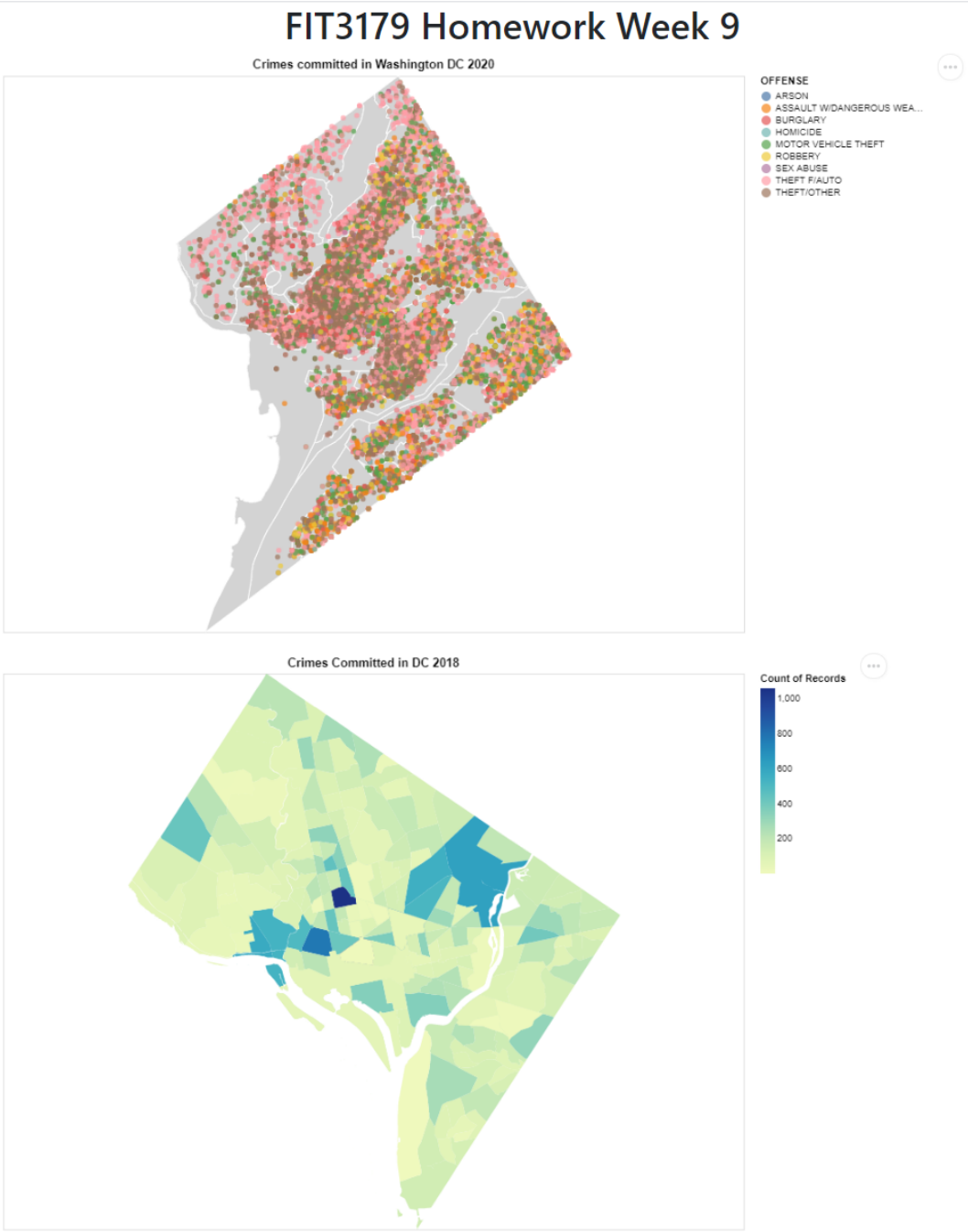

Figure 1: Interval Partioning Flowchart