DATA VISUALISATION 2 REPORT

Glen Chen 30572355 gche0026 – FIT3179 Wednesday 5-7pm Studio – Tutor Joe Liu

URL: <https://pathfinder4.github.io/FIT3179-Visualisation-2/>

Words:

# Introduction and Background

\*\*Insert Picture here \*\*

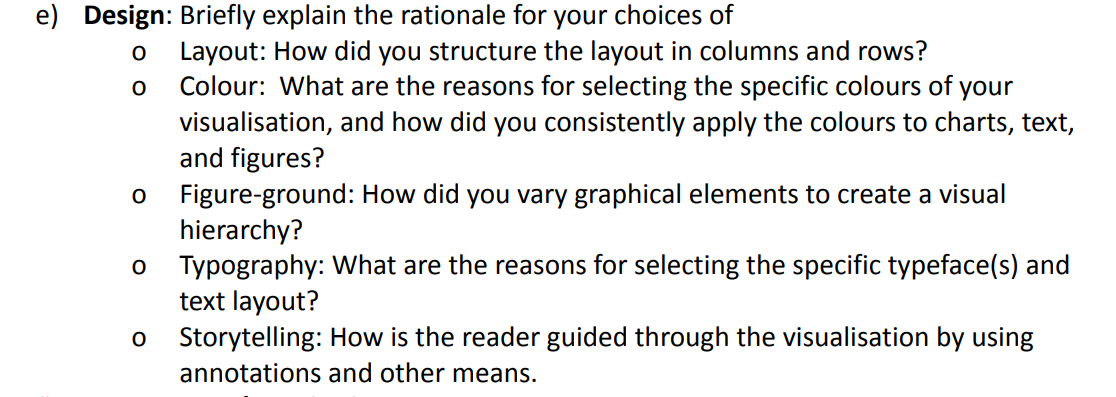
The above infographic is focused around the topic of crime, particularly violent crime in the US. The data was chosen from a range of sources, including a compilation of national police jurisdiction and FBI reports. These datasets were chosen to display trends and changes over time and compare different states using a range of long term publicly available datasets.

The design of the graphic outlines a general overview to any non-expert in the field, with the overview statistics on the top half of the graphic being easily digestible with the attached explanations. In addition, more detailed comparisons are made, through highlighted states chosen to show both the expected violent crime rate of a state, as well as bring a spotlight on certain states which perform as outliers. This is chosen to be used as reference material if required by enthusiasts or specialists wishing to identify relative crime rate between certain states.

The data processing used was fairly complex and involved significant use of the pandas python library to process the large data sets (up to 100s of MB) as can be seen in Appendix A. In addition a range of manual data cleaning methods on excel and python scripting were often used to merge different data sets together. This was necessary to keep the datasets small for quick processing by Vegalite. Additionally, it allows minimal transforming and layering on Vegalite when not required to prevent errors and ensure clean code. If interested, all unprocessed data is publicly listed on the Github, under “data/uprocessedData”.

# Design Choices

\*\*change all graphic titles to include dates as well as key jurisdictions label if limited states are shown (change Fatalitites and go into the datat and change the state names to their long names so its less confusing)



Refer to last report I did well in explaning

### Map Overview

### Yearly Crime Progression In Key Jurisdictions

### Fatality Statistics

### Crime Type And Offence Per Capita Comparison

### Juvenile Violent Crime Comparison

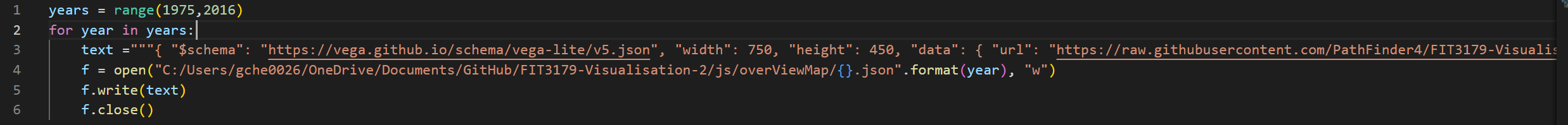
# References

1. Mapshaper.org (2022). Mapshaper is an editor for map data. Retrieved from: <https://mapshaper.org/>
2. United States Census Bureau. (2018). Cartographic Boundary Files - Shapefile. Retrieved from: <https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html>
3. Blue7red. (2021). arrests national juvenile dataset study. Retrieved from: <https://www.kaggle.com/code/rhythmcam/arrests-national-juvenile-dataset-study/data>
4. Kareem, H.F. (2021). fatalities from 2000 to 2016. Retrieved from: <https://www.kaggle.com/datasets/hamdallak/police-fatalities-from-2000-to-2016>
5. Berkay, A. (2019). USArrest Crime - Kmeans Clustering - Elbow. Retrieved from: <https://www.kaggle.com/code/berkayalan/usarrest-crime-kmeans-clustering-elbow/data>

# Appendix A – Data Processing Scripts

### Overview map and yearly line graph data processing





### Fatalities statistics processing



### Juvenile crime statistics

### 

# Appendix B – 5 Design Sheets

### 