San Francisco Crime: Descriptive Statistics and Geographic Visualization Rubric

Legal Studies 123, UC Berkeley

February 11, 2022

3 Descriptive Statistics	criteria	points
3.1 Plot the number of incidents per year from 2018 to present (choose the appropriate type of plot). Have crime rates increased or decreased in general?	plot conclusion	1
3.2 To get a more granular look, plot the number of incidents per month per year from 2018 to present. How does added granularity change your previous analysis of crime rate increase or decrease?	plot conclusion	1
3.3 Check to see if these relationships change when looking at particular types of crime. Plot and explain your findings.	meaningful categories of crime trend over time	2
3.4 Looking just at 2019, what proportion of the total does each type of crime constitute? Use at least one table and at least one plot to support your answer.	table & plot	1
3.5 Is there a relationship between day of week, time, and whether an incident occurs?	relationship supporting data (visualization)	1
3.6 Is there a relationship between day/time and particular types of incidents? What about time of year?	relationship supporting data (visualization)	1
3.7 What neighborhoods experience the most crime? Do different neighborhoods experience different types of crimes at different rates, or is the distribution of crime spatially consistent across neighborhoods? (Note: You can use the "Police District" column for neighborhood information.)	thoughtful response to each question, with ref. to evidence	2
3.8 Discuss two other interesting findings from your data.	two or more other findings and why they are interesting	2

4 Geographic Data		
4.1 Plot individual incidents in 2019 as points on a map of San Francisco 1. Does crime seem randomly distributed in space, or do incidents tend to cluster close together? 2. Shade the points by type of crime and analyze whether certain neighborhoods experience certain types of crime more often. 3. Propose social scientific explanations for the patterns that you find.	be sure to give thoughtful text answers in a markdown cell for each of the three prompts	4
4.2 Merge the incidents data with the GeoJSON file which contains the information on the boundaries of neighborhoods in San Francisco.	main output: successfully merging data	5
5 Mapping Incidents		
5.1 Construct a choropleth map, coloring in each neighborhood by how many incidents it had in 2019. Then, construct several maps that explore differences by day of week, time of year, time of day etc.	at least three different and interesting choropleth maps of neighborhoods	2
5.2 Do you notice any patterns? Are there particular neighborhoods where crime concentrates more heavily?	text explanation of patterns in choropleth maps	1
5.3 Construct a heat map of crime. How does the heat map compare to the choropleth map? Are neighborhoods a reasonably good proxy for the actual concentration of crime?	heat map discussion of neighborhood as unit	2
6 Discussion Questions		
6.1 Based on the evidence from this lab assignment, why do you think “hotspots” policing became more popular in the last few decades? What are the pros and cons to this kind of approach?	narrative evaluation of hotspot policing with reference to at least one reading responses to each prompt	2
6.2 Comment on what sorts of incidents get reported in this database. For instance, do you see a lot of reports about things like white collar crime? How do you think incident categories are selected? As data scientists, what kinds of ethical and legal concerns should we be aware of when we construct these sorts of datasets?	narrative reflection on dataset and processes that generated it responses to each prompt	2
6.3 What other sorts of data would help improve your analysis?	open ended (can include data that may be difficult to collect)	1