Falconine Falcons

Team Members: Ivy Chen, Mussie Engdashete, Ed Gomez, Jeff Stevens

Crime Stats (Project 1)

July 8, 2019

PROJECT DESCRIPTION

Research of calls for service identified as burglaries within the Los Angeles Police Dept. jurisdiction for the calendar year of 2018. Analyze the correlation of burglaries which occur at specific times of days and how they compare to same time on different days.

RESEARCH QUESTIONS TO ANSWER

1. Determine the number of burglaries in five randomly selected jurisdictions within the Los Angeles Police Dept.?
2. What time of day did the burglary occur in each of the five jurisdictions?
3. Is there a correlation for the time of actual incidents and the time of reporting?
4. Is there a trend across the five major jurisdictions?
5. Hypothesis: the number of instances of burglary increase during the summer season.
6. Hypothesis: most burglaries occur during daytime hours.
7. Hypothesis: most burglaries occur during weekdays.

DATASETS INCLUDED

* data.lacity.org (api-based json data)
* https://crime-data-explorer.fr.cloud.gov/api (FBI Crime Data Explorer)
* Crime blotter3
* Crime blotter4

TASKS / MILESTONES

Day 1 (Mon1)

* Decompose the research questions (consider the interpretations of the data and how that influences collection and analysis, is the scope manageable and reasonable for adding value to the question)
* Identify the Data Sources
* Define the strategy and metrics (create the blueprint for the data being targeted, e.g. how many cities, random sampling)
* Build data retrieval plan (e.g. Yelp Fusion for querying the API, Python script to randomly select 700 US cities)

Day 2 (Wed)

* Retrieve the data
* Assemble and validate the data

Day 3 (Sat)

* Analyze for trends
* Draft findings
* Acknowledge limitations

Day 4 (Mon2)

* Final QA of report
* Finalize presentation

DELIVERABLES

1. GitHub repository
2. Project documentation
3. Final presentation

**Development Requirements**

* Use Pandas to clean and format your dataset(s).
* Create a Jupyter Notebook describing the data exploration and cleanup process.
* Create a Jupyter Notebook illustrating the final data analysis.
* Use Matplotlib to create a total of 6–8 visualizations of your data (ideally, at least 2 per ”question” you ask of your data).
* Save PNG images of your visualizations to distribute to the class and instructional team, and for inclusion in your presentation.
* (Optional) Use at least one API, if you can find an API with data pertinent to your primary research questions.
* Create a write-up summarizing your major findings. This should include a heading for each “question” you asked of your data and a short description of your findings and any relevant plots.

**Presentation Requirements**

Prepare a formal, 10-minute presentation that covers

* Questions you found interesting and what motivated you to answer them
* Where and how you found the data you used to answer these questions
* The data exploration and cleanup process (accompanied by your Jupyter Notebook)
* The analysis process (accompanied by your Jupyter Notebook)
* Your conclusions, which should include a numerical summary and visualizations of that summary
* The implications of your findings: what do your findings mean?