Falconine Falcons

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

Crime Stats (Project 1)

July 8, 2019

PROJECT DESCRIPTION

Research of crime statistics for five major cities within Los Angeles County and analyze the correlation of residential burglaries which occur at a specific time of day and how it compares to same time on different days.

RESEARCH QUESTIONS TO ANSWER

1. What is the burglary rate for each of the five major cities?
2. What time of day did the burglary occur in each of the five major cities?
3. Is there a trend for what time of day burglaries occur?
4. Does this trend change across the five major cities?

DATASETS INCLUDED

* Crime blotter1
* Crime blotter2
* Crime blotter3
* Crime blotter4
* Crime blotter5

TASKS / MILESTONES

Day 1 (Mon1)

* Decompose the ask (Consider the interpretations of the terms 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 (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 with code samples
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?