

Projects & Collaboration with Git

Data Boot Camp Lesson 7.1



Project Week Overview

Project Week! (This Week)

Day 1:



Form Groups



Outline Project Ideas



Initial Data Exploration



Begin Research of Datasets

Submit Project Proposal for Approval



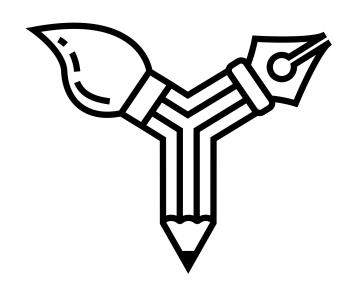


Hardcore Development





Hardcore Development



Project Week! (Next Week)

Day 4:



Hardcore Development

Day 5:



Hardcore Development



Presentation Prep

Day 6:



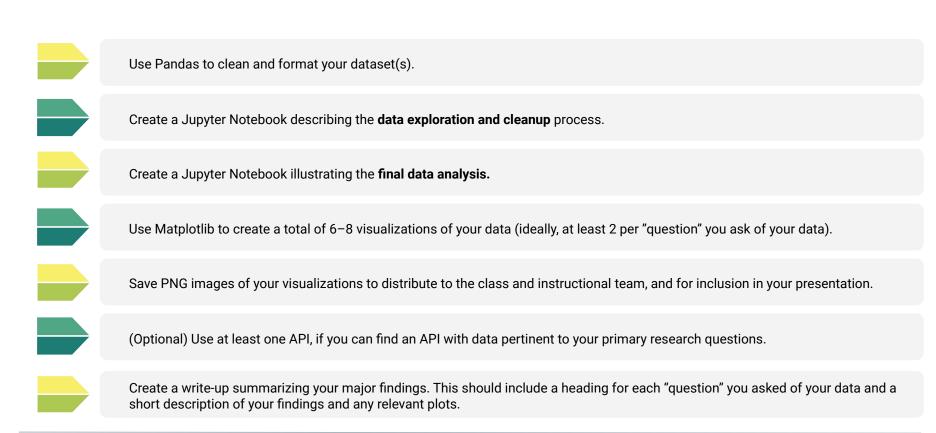
Presentations

Time to divide into teams!



Project Requirements

Development Requirements



Presentation Requirements

You will also be responsible for preparing 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?

Suggested Data Sources

Suggestions for Data Sources

Feel free to ask us (the instructional staff) for input, but our general advice is to stick to data sources that:



Are sufficiently large.



Have a consistent format.



Ideally, contain more data than needed.



Are well-documented.

Example Project Ideas

Private Investigator

01

Use aggregate crime data from different police precincts in a city to uncover patterns in criminal activity.

02

Most crime in NYC takes place in the summer.

Can you uncover similar patterns in your city?



03

What do your results suggest about how police should plan their patrols?

What do your results suggest about how best to distribute law enforcement resources over the calendar year?

www.nydailynews.com

Uber Rides and Weather

01

No one likes to walk in subzero temperatures *or* scorching heat. Do people use Uber more when the weather is uncomfortable?

02

Using <u>Uber ride data</u> from <u>Kaggle</u> and data from a weather API, find out if people take Uber more during summer and winter, and if there are relationships between daily temperature and ride frequency.

03

What do the results tell you about surge pricing strategies and commuter habits?

www.kaggle.com

Bullying and Crime Rates

01

Bullying and violent crime seem like they should be related. Can we find a correlation between frequency of bullying and rates of violent crime?



02

Using <u>Data.gov's data on</u> <u>bullying</u> and data from police districts of your choice, investigate relationships between bullying and violent crime frequency and location (zip code, city, etc.).



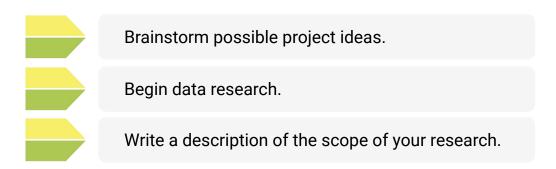
Are these two activities correlated?

What do the results suggest about society and public policy?

www.data.gov

Today's Focus

By the End of Today's Class:



Create a short, 1-page project proposal that covers the following:



