Day 24

Project Week

The Coding Bootcamp | Saturday, March 3, 2018

Project Week Overview

Project Week! (This Week)

Today's Class:

- Divide into groups
- Begin researching data sets
- Outline project ideas
- Submit Project Proposal for Approval
- Initial data exploration

Next Class:

- Hardcore Development

Class After Next:

- Hardcore Development

Project Week (Next Week)

Next Week (M/T):

- Hardcore Development
- Presentation Prep

Next Week (W/Th)

- Next Unit -Introduction to SQL

Saturday's Class:

- Presentations!

Teams

Teams - Mon/Wed Class

Team #1
Amy
Ashwini
Gayatri
Jaejun
Opeyemi

Team #2
Diana
Frank
Michael
Trevor
Steven

Team #3
Alfonso
Daniel
Deric
Heather

Team #4
Dylan
Joe
Richard
Vivian

Team #5
Anselmo
Kathleen
Kenneth
Solomon
Oscar

Team #6
Aditi
Christopher
Nishit
Bill
Nelson

Teams - Tue/Thu Class

Team #1
Naveen
Anjali
Kiera
Alan

Team #2
Jen
Chris
Anh
Steven

Team #3
Andrea
David
Luke
Rupali

Team #4
Fernando
Chander
Jeanette
Mark

Team #5
Jasmine
Matt
Abby
Gloria

Team #6
Yizhi
Marko
Madeleine
Sarah

Team #7
Javier
Nick
Jordan
Bryan

Team Roles

Team Lead

- Responsible for reporting status to us and create presentation slides (present with others)

Scrum Master

- Responsible for organizing github Kanban boards and tasks, checking progress, collaborating on issues.

Data Engineer

- Responsible for finding, storing and making data available. Work with Tester (& team) to ensure clean / sanitized data.

Tester

- Responsible for making sure all code works, integration of code, and deployment, no bugs.

Requirements

Development Requirements

- Use **Pandas** to clean and format your data set(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
- Optionally, 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 under each heading, a short description of what you found and any relevant plots.

Presentation Requirement

- **You will also be responsible for preparing a 10 minute presentation.**

- **This will be a formal presentation.**

- **One in which you explain in detail:**

  - **What questions** you and your group 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**. This should include a **numerical summary** as well as **visualizations** of that summary
  - **Discuss** the implications of your **findings**. This is where you get to have an open-ended discussion about what your findings "**mean**".

Be Glam for the Camera

- Presentations will be recorded…

- These can be great **portfolio** pieces for your resume to show when job-seeking if you invest the time.

Suggested Data Sources

Suggestions for Data Sources

Stick to data sources that:

- Are sufficiently large
- Have a consistent format
- Ideally contain more data than needed
- Are well-documented

Feel free to ask your instructors for input!

Example Project Ideas

Private Investigator

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

- Most crime in NYC takes place in the summer. Can you uncover similar patterns in your city of choosing?

- What do your results suggest about how police should plan their patrols? About how best to distribute law enforcement resources over the calendar year?

Uber Rides & Weather

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

- Using Uber ride data from Kaggle 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.

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

Bullying & Crime Rates

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

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

- Do these two activities track each other? What do the results suggest about society and public policy?

Today's Focus

By End of Day - Today

- Brainstorm possible ideas

- Begin Data Research

- Write a description of the scope of your research

- Create a short 1 page proposal listing out each of the following:

1. Project Title
2. Team Members
3. Project Description/Outline
4. Research Questions to Answer
5. Data sources or Data Sets to be Used
6. APIs to be consumed (if any)
7. Rough Breakdown of Tasks

Questions