Sprints 1-3 Feedback

Prof GS

SPRINT 1-2

- Good job following the instructions from the video!
- You say you are going to use Git Flow, and that was reflected in your
- I'm glad you have talked about each other's abilities, experiences, and skills, and you are being accepted and understanding; that's great!
- You all seem to have good communication and a solid plan in place. Although I don't think you are taking some of these questions seriously. Unless you plan on showing me your carrier pigeon as part of your grade, I suggest you re-think some of these questions and rephrase them for the next assignment.

SPRINT 3

Git/GitHub

- Again, good job on working as a team via git/github, your git log reflects it!
- For the next assignment I do want to see more contributions from all members, there's two members that are very active compared to the third one.
- Your README is well formatted and organized, good job!
- You have multiple datasets in your project, and it's hard to keep track of all these downloads. Please upload your datasets to a gdrive and share the link in the README for the next assignment.

Feedback on questions:

- How does the frequency and severity of climate indicators correlate with real estate investment patterns in different regions?
- Target variable: Real estate investment volume and value in specific regions in the United States.

- Task: Regression analysis to predict investment patterns based on historical weather data.

  - Ok this sounds good when I look at the target variables and the task. The question itself looks like you are going to do a correlation analysis rather than a model.

- What is the relationship between rising temperatures and changes in property prices in urban heat island areas?

- Target variable: Property prices or property value indices in urban heat island areas in the United States?

- Task: Correlation analysis to determine the impact of rising temperatures on property prices.

  - Every question should have a different target variable, and you'll answer it using modeling. Therefore, this question needs to be changed, right now its an EDA question or a subquestion, but not a modeling question.

- How do changes in temperature and precipitation levels correlate with changes in housing market characteristics, such as listing pricing trends, market duration, and inventory levels?

- Target variable: Housing market indicators (listing prices, how long houses are on the market, inventory levels).

- Task: Time-series analysis to examine the impact of weather patterns on housing market dynamics.

  - We are not doing time-series analysis in this class, nor forecasting. You need to turn this into a classification or regression question.
  - Every question should have a different target variable, and you'll answer it using modeling. Therefore, this question needs to be changed, right now its an EDA question or a subquestion, but not a modeling question.

- All three question can be answered with a single model and you only have 1 target variable.

- You should have a different target variable for each question, which will allow you to try at least three different models for each question.

A few initial concerns:

- Your data has time-series, therefore you need to be very thoughtful in how you are going to split it for a supervised ML problem. You can split the data like we did in HW1 and turn it into a regression problem, but not a time-series or forecasting problem

- You have a lot of moving pieces in your project and its not 100% clear how they are going to come together.

Feedback on Cleaning:

- Are you planning on keeping your datasets separately? or will you merge them later on?
- Please include markdown cells to explain the steps for cleaning.
- What did you learn from your visualizations?
- Your visualization on "Climate*Data*Cleaning" does show seasonality on the data.
- There needs to be more markdown cells explaining your reasoning and interpreting your results.

Steps moving forward

For the next sprint:

- Rethink Q2 and Q3 for ML.
- Revise your README with the feedback given.
- Share the datasets with us via gdrive link.
- Revisit your team contract and check if your strategies are working. Feel free to change them if they are not.
- Make sure everyone is equally contributing to the repo.
- Add explanations to your notebooks (analysis, interpretations, considerations to your cleaning).
- Meet up with your assigned TA.
- Assigned TA: Miguel