Aaran

Project 2 Instructor Feedback

Problem Statement: 2.5

Glows: Good scoping. It is clear what is being investigated and how it could be valuable.

Grows: Include metrics that will be used and how success will be defined.

Data Cleaning and EDA: 3

Glows: Steps taken were clear. Visuals and stats used to drive and demonstrate decisions. Modeling decisions chosen based on EDA results. Nice work!

Grows:

Preprocessing and Modeling: 2.5

Glows: Variables are encoded properly. Methodical, logical process for building and comparing models. Many models tested.

Grows: Given how close the initial scores were, it might be good to do a little tunining before moving on with just Ridge. It would also be good to compare RMSE for those initial models also.

Evaluation and Conceptual Understanding: 1

Glows: Baseline identified, but non-standard baseline used. Overall, good use of metrics throughout modeling. Good look at model coefficients.

Grows: Be cautious about using accuracy when discussing the performance of regression models, especially calling R2 accuracy. Also, be cautious about comparing coefficients directly for variables on different scales. Compare models to baseline in final evaluation. Make sure to interpret neighborhood features in reference to the baseline neighborhood (the one that was dropped).

Conclusion and Recommendations: 2

Glows: Conclusions follow directly from the problem statement and work done. Good inclusion of future

work.

Grows: Given your resulting model performance and inferences, what recommendations would you give? Are your models performant enough to make strong recommendations, or is further work needed?

Project Organization: 2.5

Glows: Good repo organization. Good use of relative paths. Nice use of markdown and comments.

Grows: Put all imports at the top of the notebooks.

Visualizations: 2

Glows: Visuals are good sizes and nicely formatted. The insights gained are clear.

Grows: Some visualizing of the categorical features, especially neighborhood in eda would be good to see.

Python Syntax and Control Flow: 2

Glows: Code is clean and error free. Good organization into code blocks for readability. Good use of functions.

Grows: Better to use variables than hardcoding numbers for filling the missing values. Some lines of code are a little long and could be split into multiple lines. Move your functions to a .py file and add docstrings to take it to the next level.

Presentation: 2.5

Notes: Clear problem statement with understandable scope. Strong intro orients audience towards project aim. Good approach to include both score and RMSE in pricing terms. Could have used a little more detail on model assesment (scores on train & test, high level note on why you chose different subsets of features). Clean and well formatted presentation. Nice use of scatterplots to highlight relationships of key variables. Good idea to review plot size and neighborhood to demonstrate variance.

Total: 20

Avg: 2.22