Requirements

color-key

Things that went well
My notes
Post presentation questions
Things to reflect on

For next time

- ❏ Look into collinearity between rooms and sqft_living
- ❏ Add some graphs to show how the data is correlated
- ❏ Remove OSEMiN slide
- ❏ Iterative process to modeling
- ❏ Add questions to notebook (bigger markdown)
- ❏ Investigate zipcode and age using modeling
- ❏ Look into bathrooms (pvalue is very high)
- ❏ Explain 3 variables
- ❏ Look into lat, long

This section outlines the rubric we'll use to evaluate your project.

- ❏ What did you learn the most?

- ❏ If given more time, what would you explore further?

- ❏ Things to think about

Technical Report Must-Haves

For this project, your Jupyter Notebook should meet the following specifications:

Organization/Code Cleanliness

- ❏ The notebook should be well organized, easy to follow, and code should be commented where appropriate.
    - ❏ Level Up: The notebook contains well-formatted, professional looking markdown cells explaining any substantial code. All functions have docstrings that act as professional-quality documentation

- ❏ The notebook is written for a technical audiences with a way to both understand your approach and reproduce your results. The target audience for this deliverable is other data scientists looking to validate your findings.

- ❏ Feedback
    - ❏ Very clean
    - ❏ Very organized
    - ❏ Really love your markdown cells and images
    - ❏ Could use a bit more markdown headings, but nothing too serious

Visualizations & EDA

- ❏ Your project contains at least 4 _meaningful_ data visualizations, with corresponding interpretations. All visualizations are well labeled with axes labels, a title, and a legend (when appropriate)
    - ❏ Histograms over all dataframe variables
        - ❏ Looked at range and distribution of features
        - ❏ Found which variables were continuous and categorical
        - ❏ Looked at normality
    - ❏ df.describe()
        - ❏ Looked at the 5 point statistics across all features
        - ❏ Looked at median and mean
    - ❏ Correlation heatmap
        - ❏ Used to find correlation and remove collinearity
        - ❏ Also, used this to find what was correlated with price
        - ❏ Great job!
    - ❏ Scatterplot of features vs price
        - ❏ Living area
        - ❏ Sqft_above
        - ❏ Grade
        - ❏ All positively correlated

- ❏ You pose at least 3 meaningful questions and aswer them through EDA. These questions should be well labled and easy to identify inside the notebook.
    - ❏ Level Up: Each question is clearly answered with a visualization that makes the answer easy to understand.

    - ❏ Feedback
        - ❏ Do people consider zipcodes?
            - ❏ No, they do not

- ❏ Need to check this using modeling
- ❏ Based on the initial conditions what you can purchase in King County
- ❏ Age of house affect price?
    - ❏ No
    - ❏ Need to check this using modeling
- ❏ Your notebook should contain 1 - 2 paragraphs briefly explaining your approach to this project through the OSEMN framework.

Model Quality/Approach

- ❏ Your model should not include any predictors with p-values greater than .05.
    - ❏ Bathrooms has high pvalue
- ❏ Your notebook shows an iterative approach to modeling, and details the parameters and results of the model at each iteration.
    - ❏ Level Up: Whenever necessary, you briefly explain the changes made from one iteration to the next, and why you made these choices.

    - ❏ Feedback
        - ❏ You iterated through this, but didn't catch some pvalues
        - ❏ I think it's also worth experimenting with more variables

- ❏ You provide at least 1 paragraph explaining your final model.
- ❏ You pick at least 3 coefficients from your final model and explain their impact on the price of a house in this dataset.

Non-Technical Presentation Must-Haves

- ❏ The second deliverable should be a Keynote, PowerPoint or Google Slides presentation delivered as a pdf file in your fork of this repository with the file name of `presentation.pdf` detailing the results of your project. Your target audience is non-technical people interested in using your findings to maximize their profit when selling their home.

- ❏ Your presentation should:

- ❏ Contain between 5 - 10 professional-quality slides.
    - ❏ Level Up: The slides should use visualizations whenever possible, and avoid walls of text.
- ❏ Take no more than 5 minutes to present.

- ❏ Avoid technical jargon and explain the results in a clear, actionable way for non-technical audiences.

- ❏ Feedback
    - ❏ Love the design! Love the color and the simplicity, no slide is too wordy! :)
    - ❏ OSEMiN slides not necessary in non technical
        - ❏ Very technical language around this slide
    - ❏ Findings slide is very good.
        - ❏ Discussed bedrooms vs bathrooms
        - ❏ Sqft living is high predictor
        - ❏ Condition of house/ low required maintenance affects price
        - ❏ Number of rooms effects price
    - ❏ Is there a way to present some of this data and what you found specifically regarding these features?

- ❏ Your presentation should contain at least 2 concrete recommendations for how to improve the selling price of a home.
    1. Make sure all components are functional
    2. Love that your recommendations were split between old/new homes
    3. If you can add more rooms do so. Add more space if possible.
    4. Use high quality materials.