

Case Study 2

BSDS Spring 2021

Instructions

- You must work in your assigned group on Monday, 4/5 in a breakout room during the live lecture (unless you have made prior arrangements).
- On Friday, 4/9 each group must present one or two of the open-ended problems in the assignment.
- Each group will submit a Jupyter Notebook file (.ipynb) and an associated .html file via Canvas. Each group member will submit the same two files, except for the “equal work pledge”.
- The submitted notebook must start with a Markdown cell header indicating the Case Study number, dataset, group members, and any external sources. Everyone will additionally include an “equal work pledge” which states that they understand all code/answers in the assignment and that all group members contributing equally. If this is not the case please indicate.
- Each question must be annotated appropriately with Markdown cells. The Notebook file should be written in a way that a third party with no knowledge of the questions can read it.
- All group members will receive the same grade unless the “equal work pledge” is violated.

The dataset for this assignment is available [HERE](#).

- You will have to make a free Kaggle account to access the data.
 - You can either download the csv files as needed, or download the whole dataset at once.
 - ZHVI stands for Zillow Housing Value Index, a measure of typical house value in a given location calculated by Zillow.
 - ZRI stands for Zillow Rental Index, a measure of typical market rent in a given location calculated by Zillow.
1. Compare the median rental price across 3 different types of residences from Jan. 1, 2000 until the most recent date. You can either choose a specific State/City or average over all States/Cities.
 - You must include a single visualization of the price over time for all three residences.

- You must formulate a hypothesis about the data you’ve cleaned.
- 2.** Compare the median rental price for all homes across 5 geographic locations (different states, metros, etc.) since Jan. 1, 2009.
- You must include a single visualization of the price over time.
 - You must formulate a hypothesis about the data you’ve cleaned.
- 3.** Pick one geographic location and compare the ZHVI for all homes to the sale price over time in this location.
- You must include a single visualization of the data.
 - You must formulate a hypothesis about the data you’ve cleaned.
- 4.** Analyze the data, similar to the tasks above, including one or more visualizations. (Suggestions: Take a look at ZRI or “DaysOnZillow”)
- BONUS:** Incorporate data from another data source (outside of the Kaggle page) into your answer for number four.