Group 6: House Hunter Project Project

Theme: In the US the housing market has recently become a source of contention with many finding homes unaffordable. This project seeks to better understand what affects the affordability of homes in states of the US that interest us as a group.

Powerpoint presentation:

<https://docs.google.com/presentation/d/1lZUSUH1wPR_7J_dUdGGs_k9zkZGvK4MB/edit?usp=sharing&ouid=109685559995624925295&rtpof=true&sd=true>

# USA Real Estate Dataset

<https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset>

1. Four min. Questions:
   1. Which state/territory has the most affordable 3 bed 2 bath?

Hypothesis: Puerto Rico will have the most affordable 3 bed 2 bath home.

Null hypothesis: There is no difference in pricing for 3 bed 2 bath homes regardless of the state the home is located in.

Steps:

* + - 1. Identify all of the 3 bed 2 bath homes in every state
      2. Separate homes by state
      3. Sort homes by price (in decreasing order)
      4. Identify home with the least price in each state
  1. Based on this data, which attribute (number of bedrooms/baths, square footage) most affects the selling price of a home?

Hypothesis:

-The higher the house size, the more expensive the house is.

- The number of bedrooms does not result in lower house prices.

- The number of bathrooms results in a higher house price.

Null Hypotheses:

* The house size does not increase the price of the house
* The number of bedrooms does not affect the house price
* The number of bathrooms does not affect the house price

Steps:

* + - 1. Drop outliers
      2. Generate a scatter plot of Price vs. Bedrooms
      3. Generate a scatter plot of Price vs. Bathrooms
      4. Generate a scatter plot of Price vs. House Size
      5. Analysis using regression for all 3 attributes
  1. What is the average sales price for each state and how does the number of rooms compare to the sales price for that state?

Hypothesis: The higher number of rooms, the more expensive the home will be.

Null Hypothesis:

Steps:

1. Grouped the data by the states and territories and called it group\_by\_states
2. Used the .mean() function on the bedrooms, bathrooms, and price columns
3. Created a bar graph to show the average sale price for each state and sorted the states from highest to lowest
4. Created a bar graph to show the average number of bedrooms and bathrooms for each state
5. Created a scatter plot to show how the average number of rooms relate to the sales price for each state with blue circles being bedrooms and orange triangles being bathrooms
   1. For the all states, how does the average home size impact the average selling price per state?

Hypothesis: Homes that are of larger size will have a higher sales price - positive correlation.

Null Hypothesis:

Steps:

1. Used the .mean() function on the home size column
2. Created a scatter plot to show how average home size (square footage) relate to the average sales price for each state with stars representing each state and where they fall on the chart
   1. For the state with the most homes for sale (data), how does the home size impact the selling price for that state and what is the correlation?

Hypothesis: Homes that are of larger size will have a higher sales price - positive correlation.

Null Hypothesis:

Steps:

1. Used the .value\_count function to determine the number of homes in each state by counting each time the state showed up in the dataset.
2. Created a pie chart to show the amount of homes in each state
3. Created a dataframe to locate all information for the state of Pennsylvania, which had the most number of homes
4. Created datasets for the price & size of homes in Pennsylvania
5. Created a scatter plot for the Pennsylvania sale price vs Pennsylvania home square footage.
6. Calculated the correlation coefficient for the Pennsylvania sale price & Pennsylvania home size.
   1. How does the acreage impact the selling price per state?

Hypothesis: Homes with more acreage will have a higher selling price per state than homes with less acreage.

Null Hypothesis: Acreage will have no impact on the selling price of homes.

Steps:

1. Rough breakdown of tasks
   1. Zemi- create powerpoint, A
   2. Lakshmi-Data cleaning
   3. Ben- C, D, & E
   4. Corine- B
   5. Sarah- Project manager, F

We all pick 1 question to code and create 2 graphs

Presentation Doos and Don'ts

* Don’t present a lot of your code (flowcharts instead only)
* Do make it colorful (Seths secretly artsy fartsy)
* Don’t make it boring
* Do make charts easy to read
* Don’t make chart axis’ misleading
* Do add a note if axis starts in a weird spot
* Do make sure charts are well labeled
* Do make sure everyone talks equally
* Do keep text per page low