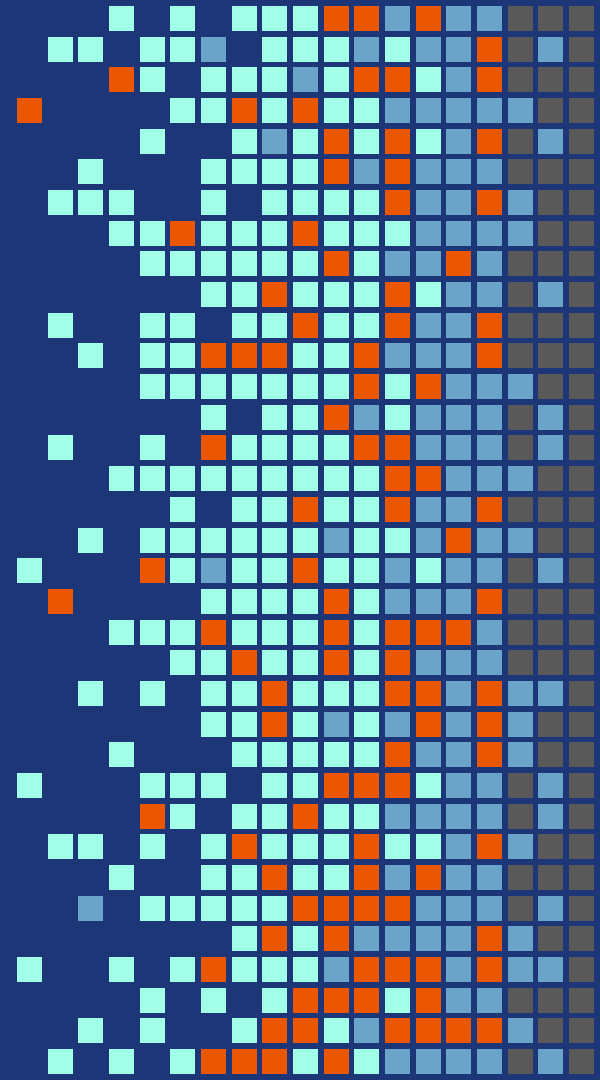# HomeDescription Demo Day Deliverable

datares

March 8

Meet the Team

Gauresh Kapoor


Leo Tien

Ananya Garg

Jane Zou

Joanna Wang

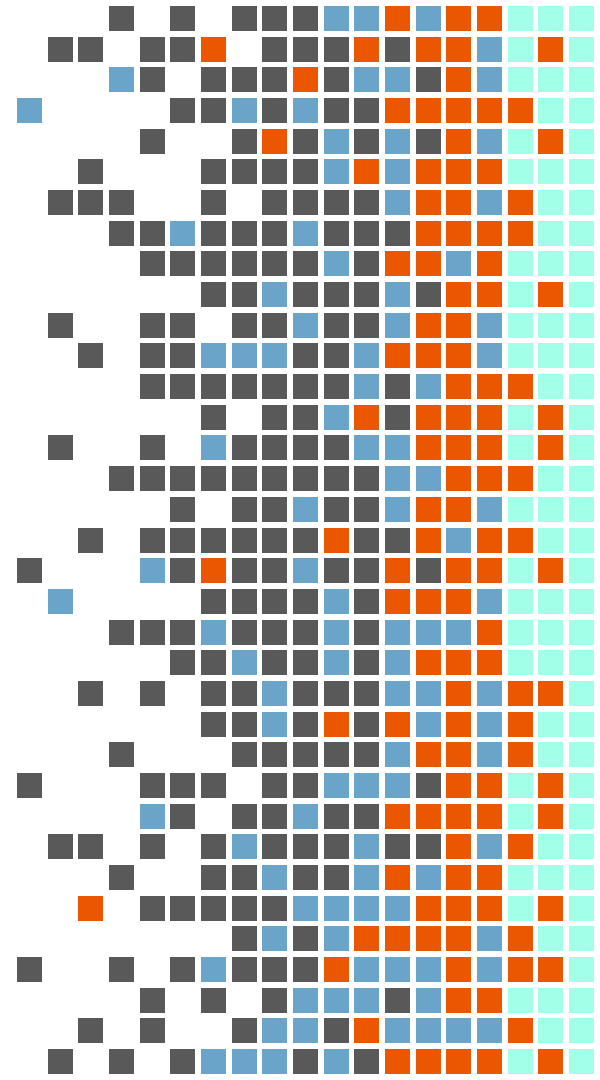# Introduction - Project Scope and Web Scraping

❖ Focused on creating a "sellability" index for real estate descriptions on zillow.

❖ *Web scraped* property description data from Zillow and store the data sourced in a file named final_zillow.csv.

❖ Worked on researching real estate literature.

❖ Engaged in *Hypothesis testing* that text description impacts sale time to sell.
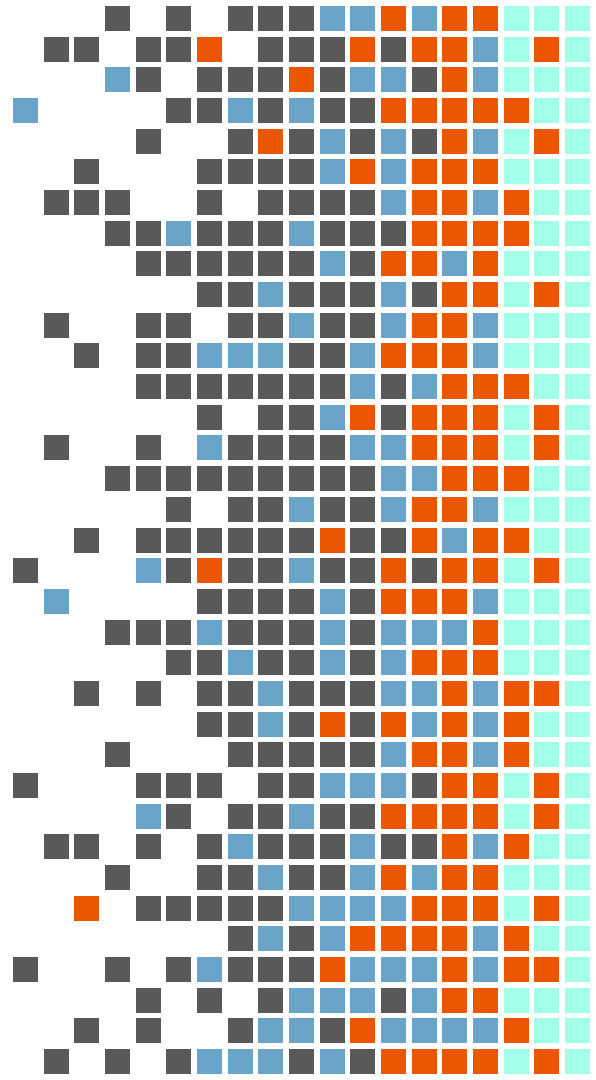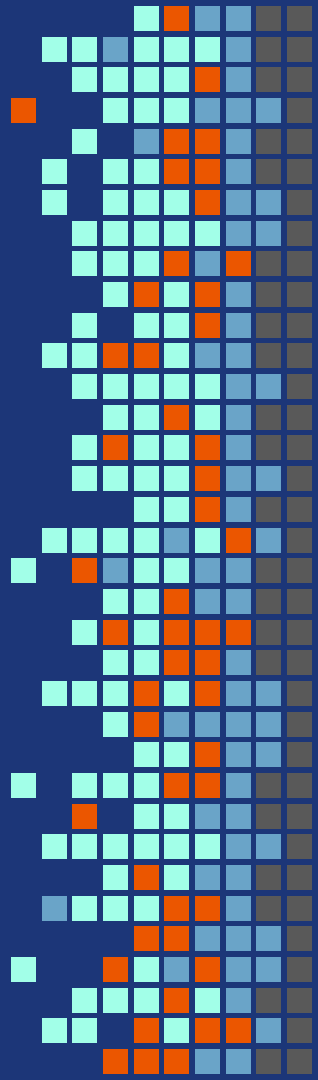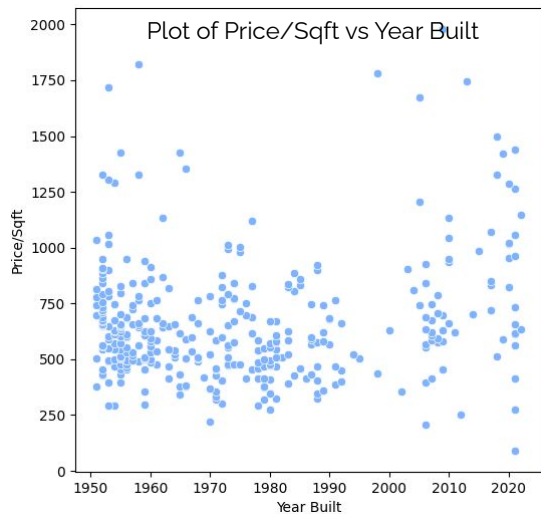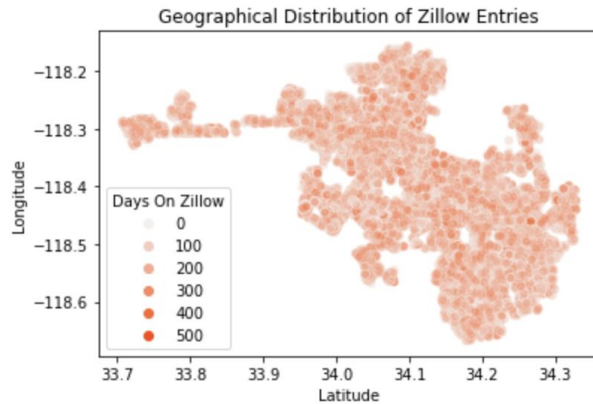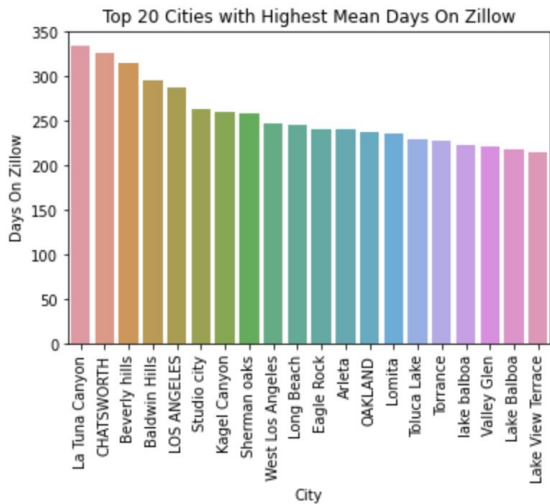
Exploratory Data Analysis

- Low and steady mortgage rates are encouraging the LA real estate market.

- Popular home locations are generally along the ocean or near mountains, or else for social life and atmosphere

- Key words: natural beauty, stimulating lifestyle, investment, exclusive

Visualizing Summary Statistics

Top 20 Cities with Highest Mean Days On Zillow
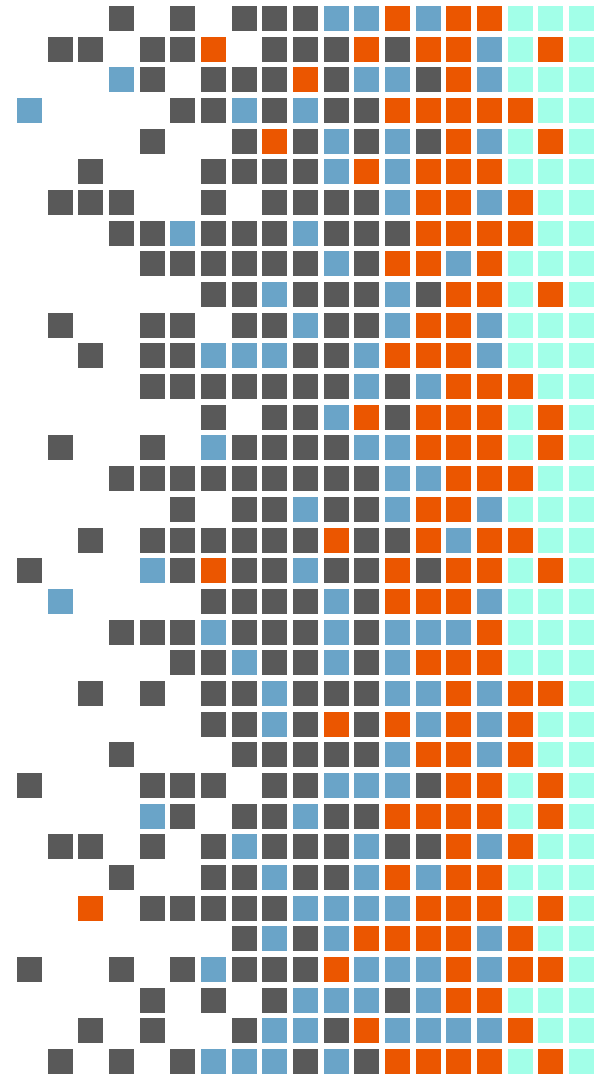
Geographical Distribution of Zillow Entries
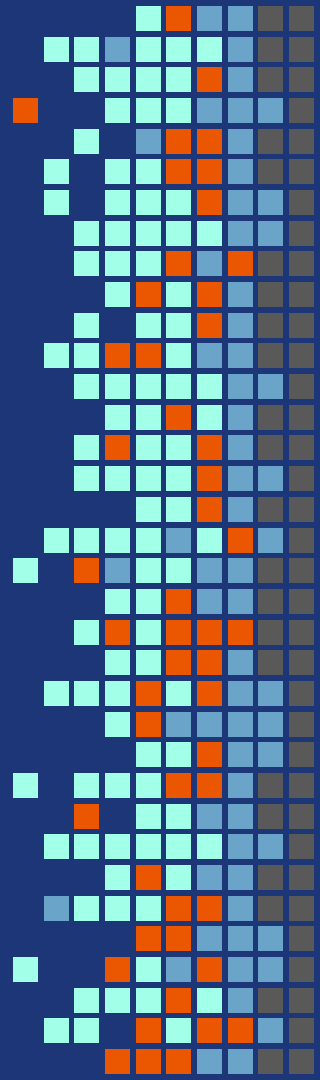
Plot of Price/Sqft vs Year Built

- Favorite Count vs. Price/Sqft r-value: 0.10
- View Count vs Price/Sqft r-value: 0.11
- Year Built vs Price/Sqft r-value: 0.16

# Analyzing Real Estate Descriptions

Description Length vs. Favorite Count

- After removing the NAs from the Description column, we have 34190 observations.

- Fitting a linear model:
  $R^2 = 0.0008$

Description Sentiment Scores

Conclusion - Next Steps

Plan for Spring Quarter:

❖ Work on creating a dashboard (using Tableau, Streamlit, etc.) for an easy to view a summary of the real estate sale data.

❖ Research further into the hypothesis that text description impact the property sellability.

❖ Building a NLP model to see how likely it is that the property will sell, and assign scores based on it.
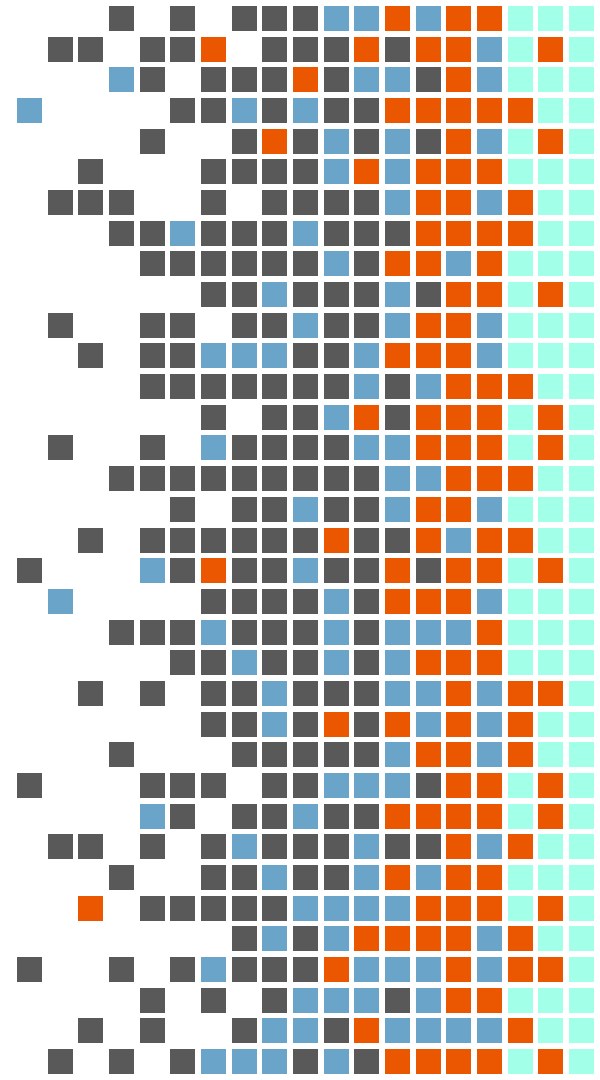
THANKS FOR LISTENING!

Any questions?