# Final Presentation

## Zillowbnb

*Zestimate for Airbnb Listings*

# Background

People rely on Zillow's Zestimate to estimate the market price of houses. The Zestimate predicts a home's value using publicly available housing data. It serves as a starting point for homeowners who are looking to sell their property and enables buyers to gain a better understanding of the market.

Such a feature would be useful to predict the market rental price for a unit on Airbnb.  Guests could use the tool to determine whether listings are fairly priced and hosts could determine a good listing price.

# Data Used

Inside Airbnb (http://insideairbnb.com/get-the-data.html) uses the publicly available data from Airbnb. The data has been released under a Creative Commons CC0 1.0 Universal (CC0 1.0) "Public Domain Dedication". This site includes multiple csv files for many cities compiled by Airbnb.

**Listings.csv.gz:** Detailed listing data. Some variables of interest include neighborhood, latitude, longitude, room_type, price, minimum_nights, maximum_nights, bathrooms, bedrooms, beds, and specific amenities such as air conditioning, indoor fireplace, washer, drier, and many more.

**Calendar.csv.gz:** Detailed calendar data for each listings

**Reviews.csv.gz:** Detailed reviews data for each listing

# Use Cases

1. Host wants to get a recommended listing price.
    a. User clicks “Host”. Form is displayed
    b. User fills out form about unit specifications, clicks submit
    c. Specifications are inputted into model
    d. Page shows an interval with the predicted rental price of the unit

2. Guest wants to determine if a listing is overpriced
    a. User clicks “Guest”, city buttons are displayed
    b. User clicks the city the listing is in, map displays
    c. User enters address into search bar, map zooms in to address
    d. Map displays estimated rental price. A colour indicating the value of the actual listing (good, average, poor) is displayed

# Demo

# Design - Visualization

**Description:** The visualization is the interface the user uses when deciding if they are using the Guest Visualization or the Host VIsualization. Uses a bash command to run and open

**Interactions:** The visualization uses a dataframe that is constructed by the data gathering and cleaning modules, and the results of the price prediction model.

# Design - Price Prediction Models

**Description:** Machine Learning model used to predict listing prices of Airbnb listings

**Interactions:** Requires a cleaned dataset that is created from running the data cleaning modules

# Design - Data Gathering and Cleaning

**Description:** A group of modules used to read in data from Inside Airbnb and clean and aggregate the data into forms that the machine learning model can use.

**Interactions:** These modules are used to create a cleaned dataframe that can be used by the machine learning model and the visualization/user interface

# Project Structure:

```
Zillowbnb/
  |- data/
     |- Seattle.joblib.dat
     |- Seattle_low.joblib.dat
     |- calendar_price_averages.csv
     |- clean_listings.csv
     |- clean_predicted.csv
     |- reviews_sa_summarized.csv
     |- seattle_merged.csv
  |- docs/
     |- Component_Specification.pdf
     |- Final_Presentation.pdf
     |- Functional_Specification.pdf
     |- Technology Review.pdf
     |- zillowbnb.jpg
  |- examples/
     |- User_Guide.pdf
  |- zillowbnb/
     |- submodule/
        |- __init__.py
        |- constants.py
        |- convert_to_matrix.py
        |- detect_outliers.py
        |- example.py
        |- get_calendar_summary.py
        |- get_cleaned_listings.py
        |- get_data.py
        |- host_predict.py
        |- price_prediction.py
        |- sentiment.py
        |- train_model.py
     |- test/
        |- __init__.py
        |- submodule_path.py
        |- test_convert_to_matrix.py
        |- test_detect_outliers.py
        |- test_get_calendar_summary.py
        |- test_get_cleaned_listings.py
        |- test_get_data.py
        |- test_host_predict.py
        |- test_price_prediction.py
        |- test_sentiment.py
        |- test_train_model.py
     |- __init__.py
     |- zillowbnb.py
  |- .coveragerc
  |- .travis.yml
  |- LICENSE
  |- README.md
  |- requirements.txt
  |- setup.py
```

# Lessons Learned and Future Work

## Lessons Learned

- Bokeh UI
- XGBoost

## Future Work

- Explore more cities
- Merge with buildings data
- Determine weights of amenities