**Final Project Notes**

Checklist to do

Transform/Clean

Add table for Average Rates Monthly

Add table for Average Rates Day of Week

Add column for neighborhoods table.

Clean bathroom text to int

Questions

1. How do the listing rates change with amenities & accommodations?
2. Do the listing price change during different times of the year and week?
3. How does the number of unit available within a specific area affect the price of Airbnb?
4. Do good reviews drive up the rates?
5. Why is the location so important to the price?

Model

* Random Forest
* KNN
* FB Prophet

Web App – Tableau Public Outline

1. Dashboard

* Density Map
* Cards
  + Total # Reviews
  + Total # of Superhosts
  + Average Review Score
  + Total # of Neighbourhoods in Data
  + Total # of Listings
  + Average Rates (per night)
* Rates
  + Yearly
  + Monthly
  + Day of Week
* Room Type (Tree Map)
* Top 10 Most Reviews
* Top 10 Review Scores

1. Models Analysis
   1. Random Forest
      * Correlation of Features
      * Features Importance (Pie or Bubble)
      * Price by Room Type
      * Price by Location
      * Accuracy Results
   2. KNN
      * # of Listings per Neighbourhood
      * Prices in that neighbourhood
      * Accuracy Results
   3. FB Prophet
      * Previous Price Trends
      * Current Price Trends

Database Changes

* Added tables for monthly and day of week rates to get a better overview of how to price Airbnb listings (used calendar CSV)

Clean/Transform 🡪 Models

* Delete columns irrelevant to model testing
* Fill in n/a with avg results for columns with blanks to prevent skewed data
* Changed all text to categorical
* Took out a few rows w/o superhost value

Difficulties/Challenges

* Connection limitations with ElephantSQL 🡪 tableau connection slows down
* Storage size limits 🡪 go through process of transforming data with groupby
* Overfitting data

Next Steps

* Move to AWS for ETL
* Explore other models
* Create form to enter different criteriea like neighbourhood, accommodations, amenities number to display price

Presentation Ideas

* Touched on Yearly Rates, Rates by Neighbourhood, Rates by Room Type in each Neighbourhood
* We want to go more in depth to look into what other thing that affect the price.
  + Size
  + Amenities
  + Month
  + Day of Week
  + Events/Holiday of Year

Powerpoint outline

1. Title - Airbnb Price Predictor
2. Trends in NYC AirBNB
3. Exploratory Questions
4. Project process (SQL DB 🡪 Tableau Visualisations 🡪 Dashboard Trends & Price Predictor)
5. Changes to ETL 🡪 Rates Monthly/Yearly

* Predicted Rates
* Talk about transformation/cleanup & loading on the DB

1. Relationship of New Dataset
2. Model Analysis
3. Model Used & Results
4. Model Data Transformation
5. What we found on the Models
   1. Feature Importance
   2. Neighbourhood Clustering
   3. Trends & Seasonality
   4. Predicted prices overtime
6. Final Product
7. Dashboard Demo