Predicting Airbnb prices based on nearby restaurants

Team STAR

Scarlett Dias, Tejas Karangale, Anirudh Rao, Rajat Gupta

Background

- We are analyzing the effect of restaurants and eateries on Airbnb listings in New York City. We have chosen New York City as it is one of the largest and most popular tourist destinations in the world with a large amount of Airbnb listings.

Data Used

- Airbnb Listings data for NY city:

http://insideairbnb.com/get-the-data.html

- Restaurants data -

https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/xx67-kt59/data

- Geocoded restaurant data by Miles Grimshaw:

https://mgrimshaw.carto.com/tables/nytimes_nyc_restaurants/public

1. Search Airbnb listing

- User navigates through our interactive map which maps all Airbnb listings and restaurants in New York using simple clicks
- Interface highlights all restaurants in the region
- A dialog box displays information about the Airbnb listing along with it's price

## 2. Information about restaurants

- User clicks on restaurant on the interactive map
- A dialog box displays information about the restaurant including type of cuisine and health grade

3. Predict the price of an Airbnb listing:

Use Cases

- Created a feature –
  - Used geopy's full circle function to calculate distance between two locations using their geo-coordinates
  - Calculated the number of restaurants in the vicinity of each Airbnb listing considering 1 km diameter around it

_____

Feature: Distance
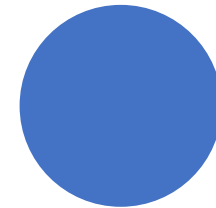
Machine Learning Model Features - Lasso

- bedrooms
- bathrooms
- accommodates
- beds
- number_of_reviews
- minimum_nights
- minimum_nights
- Number of restaurants in 1 km diameter

Demo

Visualization

Design

1

clean_airbnb_map()

clean_restaurants_geo()

- This component would be used to clean the restaurants as well as the Airbnb listing data by retaining only the relevant features for use in our analysis.

2

map_plot()

- Maps all restaurants and Airbnb listings according to their location coordinates

3

dist_calc()

- Counts number of restaurants in the vicinity of each airbnb

4

predict_price()

- Makes a prediction of the price of an Airbnb listing based on specific parameters.

Github Structure

```
uwseds-group-team-star
+-- airbnb_price_prediction
|    +-- __init__.py
|    +-- cleaning_utils.py
|    +-- create_frontpage.py
|    +-- dist_calc.py
|    +-- map_plot.py
|    +-- markers.py
|    +-- ml_model.py
|    +-- tests_clean.py
|    +-- tests_model.py
|    +-- templates
|    |    +-- airnbnb-restaurant.html
|    |    +-- ml_model.html
|    |    +-- new.html
|    +-- Data
|    |    +-- NYC Restaurants Geocoded - Sheet1.csv
|    |    +-- rest_count.csv
+-- Docs
|    +-- component_design.md
|    +-- data.md
|    +-- functional_specification.md
+-- Examples
+-- LICENSE
+-- README
+-- requirements.txt
+-- setup.py
```

Lessons Learnt

- Data cleaning takes time.
- Development is augmented and made easy using the collaborative features of github.

Future Work

- Include more external data like House Prices in order to be able to better predict the relationship of restaurants on Airbnb prices.

- Use kd trees algorithm to calculate restaurants in the nearby radius more accurately.

- Extend this module to include other cities.

Thank You