

Python Queens

Joseph meeting:

- get a linear line or logistic regression
- look at the data and refine questions based on what the data shows up
- github set up
- submit the github links so he can look at it himself

Question for Assignment

1. How is having Airbnb internationally fair against nationally used Airbnb?
2. Do sales in tourist spots jump up if there is an Airbnb present?
3. What drives the cost for Airbnb rentals?
4. What do most people prefer when booking Airbnb properties?
5. Do cities have a peak booking period?
6. Do they have a low booking period?
7. Do reviews affect bookings?
8. Do amenities affect bookings?
9. Do attractions, events and surroundings have an impact on Airbnb?
10. Does booking closer to date increase cost?
11. Does Holidays have an impact when booking Airbnb?

12. How's Covid had an impact on pricing with Airbnb?
13. What are the age's of consumers booking Airbnb?
14. Does the number of travelers determine if you book Airbnb
15. Do certain areas offer different amenities with Airbnb?

Analyses

Bhapkar Chi-Square and MANCOVA. Possibly the Spearman Rank Correlation.

International Cities

1. Sydney, Australia
2. Athens. Greece
3. Tokyo, Japan
4. Barcelona, Spain
5. Brussels
6. Belize
7. Berlin
8. Toronto
9. Mexico city
10. Bangkok
11. Paris, France
12. Milan, Italy
13. Tel Aviv, Israel
14. Dubai
15. Singapore
16. Rio de Janeiro Brazil
17. Amsterdam, Holland

National Cities

1. Orlando, FL
2. Los Angeles, CA
3. New Orleans, LA
4. Hawaii, Hawaii
5. San-Francisco
6. Seattle
7. Portland
8. Chicago
9. Austin
10. New York, NY
11. Miami, FL
12. Boston, MA

Data Science Program Final Project Executive Summary

At the end of the Data Science program, we are required to complete a final project of our choice. We have six weeks to work on the project. Our team for the final project is made up of Reina, Racquel, Elana and Evelyn.

We are focusing on Airbnb listings in a couple of cities both National and International. We will explain the purpose and scope for the project.

Business Objectives

To showcase our skills, we have acquired them through the Data Science program. We will be using R, Python, Tableau, and other programs to wrangle, analyze, and visualize Airbnb dataset made available on <http://insideairbnb.com/get-the-data.html> and on Kaggle.

At the end of the project, we should be able to explain our work in layman's terms, and present the findings to the faculty, staff, and potential employers, along with other interested parties via Zoom.

Background

As a way to activate and put practical use to what the students have learned, doing a final project is a good way to demonstrate that.

We have chosen the "Airbnb" dataset because we both are interested in traveling both local and international. We hope to glean insight from this document to make actionable suggestions on how to identify Airbnb with excellent rating

Scope

We will be using the software we have learnt in the program to complete the project. We will be using tools of our interest and tools that may aid finding a job.

Functional requirements

Data Wrangling: The downloaded dataset should be successfully cleaned up for analyzing. Columns and unusable columns should be removed. The dataset is fairly large, so we will consider sub-setting the dataset in a proper manner, meaning the subset should be a random selection of the data. The data types for each column should also be converted to a usable format for the needed analysis.

Data Analysis: We will familiarize ourselves with the dataset. Then we should have a good understanding of what each column means, and how the values are measured. We will then brainstorm on questions to ask, and what we might have gathered from the dataset. Then we will identify the proper functions to create models, predictions.

Data Visualization: Once we have a comprehensive understanding of and insight gathered from the dataset, we will work on visualizing the findings. We will use Tableau or other graphing programs, and compile the visuals and texts in a PowerPoint slideshow.

Week 2: Study the dataset and ask questions. 1. What are some possible correlations? 2. Is the data normally distributed? 3. What are some predictive models we can make from it? 4. Visualize the data to see if there are any interesting findings.