

Assignment Brief: Airbnb business analysis using data science approach

Background

Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present more unique, personalised ways of experiencing the world. The Airbnb dataset (AB_NYC_2019.csv) along with its description can be accessed through the Kaggle (<https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data>). This dataset describes the listing activity and metrics in NYC, NY for 2019, which includes all needed information about hosts, geographical availability, necessary metrics to make predictions and draw conclusions.

This assignment is aimed at assessing your ability to pose interesting questions relevant to Airbnb business, process the data using the key steps of big data analytics, such as, data pre-processing, analysing, and eventually preparing an analytical report. In order for your analysis to be compelling, it must address a substantive issue rather than a trivial one.

TASK 1 (Business analytic question)

Propose an interesting business analytic question that can be answered using the given Airbnb dataset. The proposed question should be useful for Airbnb. An example question could be: is there any noticeable difference in bookings among different areas and what could be the reasons for it?

TASK 2 (Data analysis)

In this task, use your data analytics skills to answer the question posed in the Task 1. Depending upon your chosen question, you will typically have to perform Exploratory Data Analysis (EDA), data pre-processing, statistics-based data analysis, data visualisation and use unsupervised machine learning algorithms (e.g., clustering).

TASK 3 (Analytical report)

In this task, you are expected to prepare an analytical report, which can easily be interpreted by the executive board members of the Airbnb. There is no fixed specification for this report but in general, it must contain an adequate number of visualisation charts/graphs with a lucid description. The total number of pages of the analytical report must not exceed beyond 2 pages (1000 words equivalent).

Please note: You are required to submit your Python code as an appendix for this assessment. Other appendices should not be used to extend the core report as reports should stand alone, complete, and concise, without the appendices. They should only be used if required, and only for supplementary and/ or supporting information. One key part of the exercises in this module is the need to be able to express ideas succinctly, concisely and with necessary brevity.