

# High Level Design

## Travel Data Analysis

### (AirBnB Data Analysis)

Revision number: 1.0

Last date of revision: 15/10/2021

## Document Version Control

Date Issued	Version	Description	Author
15/10/2021	HLD-V1.0	First Version of Complete HLD	Lokesh Attarde

# Contents

Document Version Control.....	2
Abstract .....	4
1. Introduction.....	5
1.1. Purpose of the Document .....	5
1.2. Objective of HLD .....	5
1.3. Scope of HLD .....	5
2. General Description.....	6
2.1 Product Perspective & Problem Statement.....	6
2.2 Data Requirements .....	6
2.3 Tools Used.....	7
2.4 Constraints.....	7
3. Design Details .....	8
3.1 Process Flow.....	8
3.2 Error Handling / Exception Handling .....	8
4. Conclusion .....	09
5. References... ..	10

## Abstract

Airbnb is an American company that facilitates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. It basically connecting travelers with local hosts who want to rent out their homes with people who are looking for accommodations in that locality. On the other hand, this platform enables host to list their available space and earn extra income in the form of rent and it also enables travelers to book unique homestays from local hosts, saving them money and giving them a chance to interact with locals.

In the world of rising new technology and innovation, Travel industry is advancing with the role of Data Science and Analytics. Data analysis can help them to understand their business in a quiet different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates the how different analysis help out to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services. Different analysis performed such as Exploratory Data Analysis and Descriptive Analysis on variety of use cases to get the key insights from this data based on which business decisions will be taken.

# 1. Introduction

This document will be used for documenting High-level designs of project.

## 1.1 Purpose of the Document

The purpose of this plan is to

- Describe different design approaches.
- Describe different analysis approaches based on variety of Use Cases.
- Describe third party components/tools required for the system.
- Present complete Process Flow followed for this project.

## 1.2 Objective of HLD

1. To provide an overview of the entire system.
2. To provide introduction of Problem Perspective & Statement, Data Requirements, Tools used and many more.
3. To provide a module-wise breakup of the entire system.

## 1.3 Scope of HLD

This HLD covers all areas of system.

## 2. General Description

### 2.1 Product Perspective & Problem Statement

Travel industries are having important reflection of the economy from past few decades, and Airbnb housing price ranges are of great interest for both Hosts and Traveler. In this project, we are analyzing the various aspects with different use cases which covers many aspects of airbnb listings. It helps in not only understanding the meaningful relationships between attributes but it also allows us to do our own research and come-up with our findings.

The objective of the project is to perform an exploratory data analysis, data pre-processing, data cleaning & imputation and at the end, apply different Data Visualization techniques to get the meaningful insight from the given data. This project aims apply some amazing Python Libraries such as Plotly and WordCloud which will give a boost to our visual understanding of the data.

### 2.2 Data Requirements

Data Requirement completely depend on our problem.

- In this project, to perform analysis, we are using datasets that are provided by iNeuron Company.
- We make a use of those different datasets as per the requirement and the problem statement.
- The features which are taken into consideration are:
- Some of the important features are:

Name	Description
<b>listing_id</b>	Unique ID of the Airbnb
<b>name</b>	Name of the Airbnb
<b>neighbourhood</b>	Name of the Neighbourhood
<b>city</b>	Name of the City
<b>property_type</b>	Airbnb Property Type
<b>room_type</b>	Airbnb Room Type
<b>amenities</b>	Name of the Amenities provided by Airbnb
<b>price</b>	Airbnb Price per day
<b>monthly_price</b>	Airbnb Monthly Price
<b>comments</b>	Traveller Comments/Reviews

## 2.3 Tools Used

- Jupyter Notebook is used as IDE.
- Pandas and NumPy are used for Data Manipulation & Pre-processing and Mathematical functions respectively.
- Exploratory data analysis is automated by dataprep.
- For visualization of the plots, Matplotlib, Seaborn, Plotly are used.
- WordCloud is used to representing the text data.
- GitHub is used as version control system

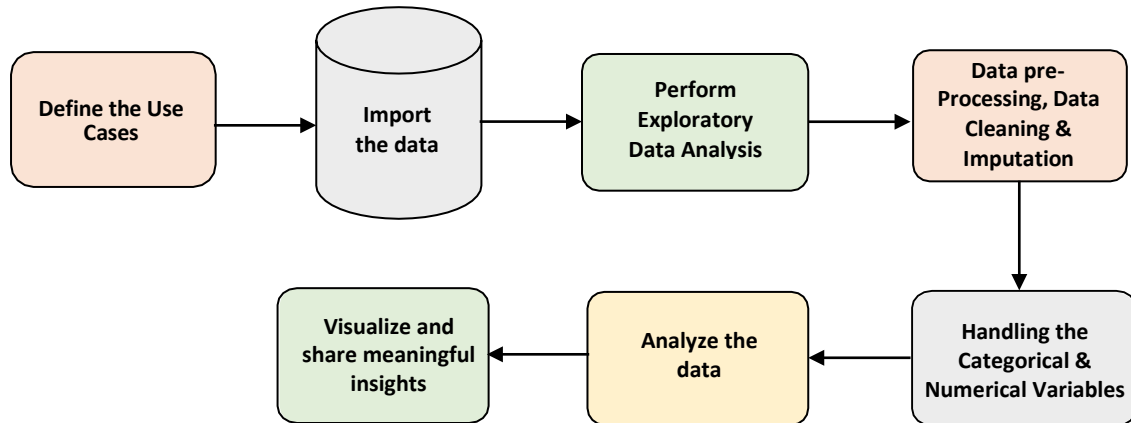


## 2.4 Constraints

The analysis must be user friendly, code must be neat & clean, EDA must be automated as much as possible because it will save huge amount of time. Moreover, users should not be required to have any of the coding knowledge as the insights they are looking for are mentioned in-detail with respective visuals.

## 3. Design Details

### 3.1 Process Flow



### 3.2 Error Handling / Exception Handling:

We have designed this project in such a way that, complete script is tested and runs multiple times to make sure that there is no error occurred during process flow.

Additionally, we have also dismiss the un-necessary warnings to avoid confusion by using `filterwarnings` class from `warnings` module.



## 4. Conclusion

In this analysis project, about fifteen different use cases are analyzed for the given dataset to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services. It has been found that **Most of the Bookings** were taken place for the "**Mission Bay**" of around "**22%**" followed by "**Pacific Beach**", "**La Jolla**" which has "**15.5%**" & "**12%**" respectively. We have also analyzed that, we have **Total "11091"** which are the "Budget Hotel" as well as they are "Affordable". Additionally, we also find-out the Top Earners (Host), relationship between Monthly Earning and Prices, Price comparison in terms of Property Type, and Room Type, Preference of Guests w.r.t. Property Type & Room Type. Furthermore, we have also been analyzed Maximum Number of Bookings, Customer Reviews/Comments, Amenities provided by the Airbnb's and many more.

## 5. References

1. [How Airbnb works, insights into Business & Revenue Model](#)
2. [Airbnb - Wikipedia](#)