Low Level Design (LLD)

AirBNB Data Analysis

Revision Number - 1.0

Last Date of Revision - 23/08/2024

AKHIL K N
USHA YADAV

Document Control

Date	Version	Description	Author
23/08/2024	1.0	Introduction, Problem Statement, Dataset Information, Architecture Description	Akhil Usha

Contents

1. Introduction

1.1 What is Low Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the AirBNB Analysis dashboard. LLDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 What is Scope?

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

Since 2008, Airbnb has revolutionized travel by offering guests and hosts a platform that enhances travel experiences with unique, personalized accommodations. This dataset provides a comprehensive overview of Airbnb's listing activities and key metrics in San Diego, California, for the year 2019. It includes detailed information about hosts, the geographical distribution of listings, and other crucial data points.

These metrics are essential for conducting predictive analyses and drawing meaningful conclusions about the short-term rental market in this region. The dataset serves as a valuable resource for exploring trends, understanding host behaviors, and evaluating the overall impact of Airbnb in San Diego.

2. Problem Statement

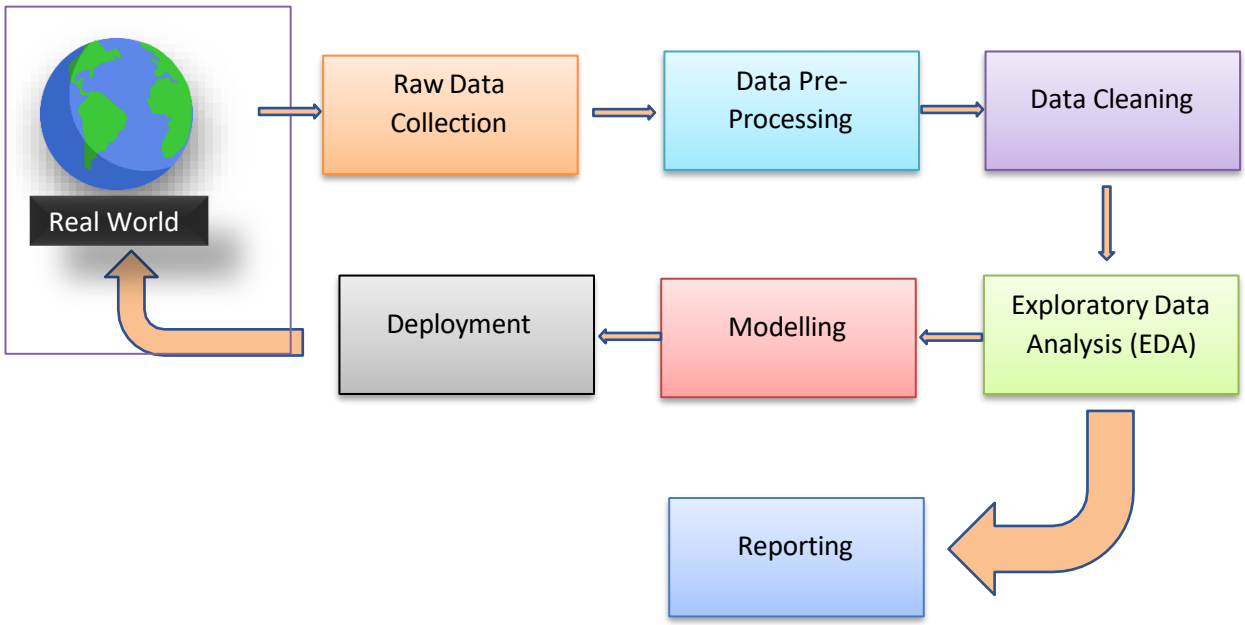
Since 2008, Airbnb has transformed the way people travel by offering guests and hosts a platform to explore diverse and personalized accommodation options. This dataset provides detailed insights into Airbnb's listing activities and key metrics for San Diego, California, for the year 2019.

The dataset contains essential information for understanding host profiles, geographical distribution of listings, and other important metrics. With 18,723 rows and 19 columns, it includes comprehensive data necessary for making predictions and drawing meaningful conclusions about the short-term rental market in San Diego.

3. Dataset Information

- **room_id** - ID of the room.
- **survey_id** - ID of the survey.
- **host_id** - ID of the host.
- **room_type** - Type of room (e.g., Shared room).
- **country** - Country where the room is located (seems to be empty in this case).
- **city** - City where the room is located (e.g., Amsterdam).
- **borough** - Borough or district within the city.
- **neighborhood** - Neighborhood within the borough.
- **reviews** - Number of reviews the listing has received.
- **overall_satisfaction** - Overall satisfaction rating (e.g., 4.5, 5).
- **accommodates** - Number of people the room can accommodate.
- **bedrooms** - Number of bedrooms in the listing.
- **bathrooms** - Number of bathrooms in the listing.
- **price** - Price of the listing.
- **minstay** - Minimum stay required.
- **name** - Name or description of the listing.

4. Architecture



4.1 Architecture Description

1. **Raw Data Collection**
   The Dataset was taken from iNeuron's Provided Project Description Document.
   https://docs.google.com/spreadsheets/d/1ebVuM0atieTe5rzRwGZqYkM9clS8RLcT/edit?usp=sharing&ouid=105973225632152334213&rtpof=true&sd=true

2. **Data Pre-Processing**
   Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feeded to the model to train.
   This Process includes-
   a) Handling Null/Missing Values
   b) Handling Skewed Data

3. **Data Cleaning**
   Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.
   a) Remove duplicate or irrelevant observations
   b) Filter unwanted outliers
   c) Renaming required attributes

4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

5. Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self- explanatory report because your model will be used by many stakeholders who are not from technical background.

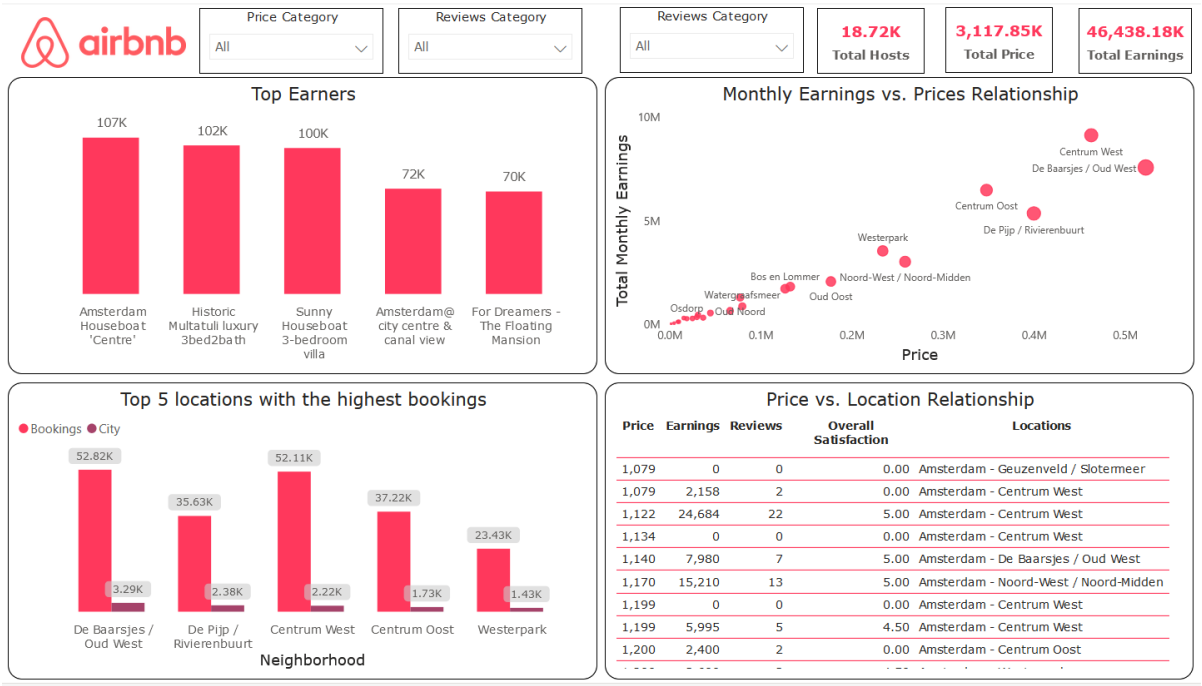
a) High Level Design Document (HLD)
b) Low Level Design Document (LLD)
c) Architecture
d) Wireframe
e) Detailed Project Report
f) Power Point Presentation

6. Modelling

Data Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

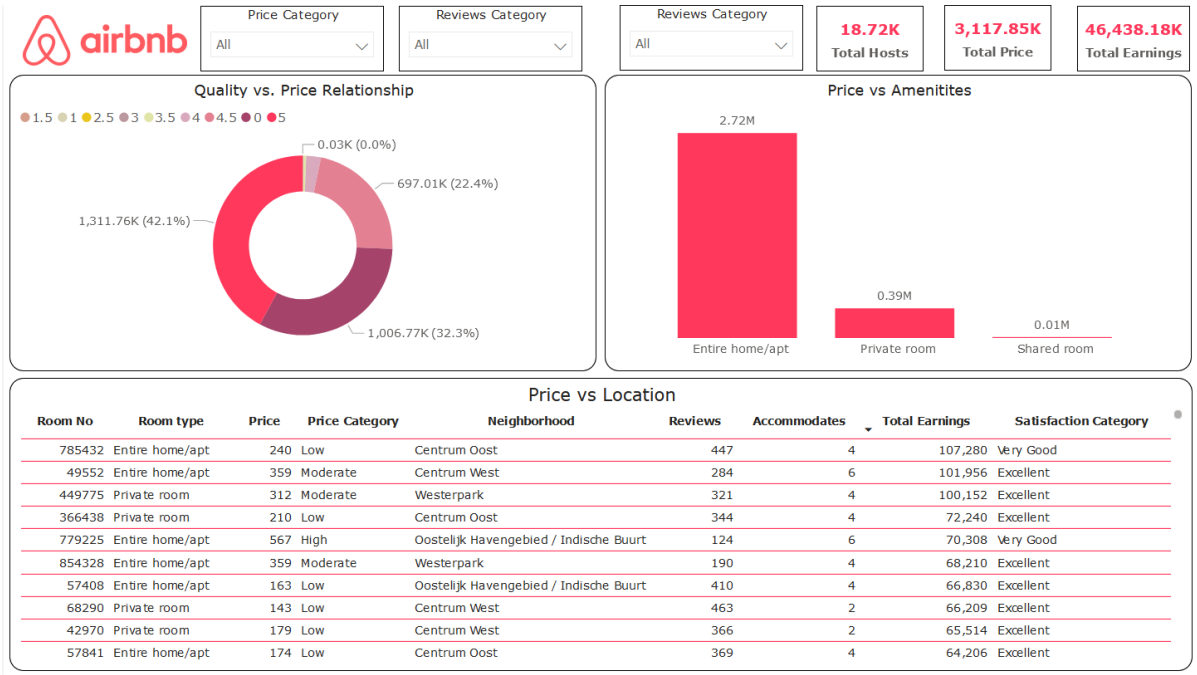
7. Deployment

We created a Power BI Dashboard



Price	Earnings	Reviews	Overall Satisfaction	Locations
1,079	0	0	0.00	Amsterdam - Geuzenveld / Slotermeer
1,079	2,158	2	0.00	Amsterdam - Centrum West
1,122	24,684	22	5.00	Amsterdam - Centrum West
1,134	0	0	0.00	Amsterdam - Centrum West
1,140	7,980	7	5.00	Amsterdam - De Baarsjes / Oud West
1,170	15,210	13	5.00	Amsterdam - Noord-West / Noord-Midden
1,199	0	0	0.00	Amsterdam - Centrum West
1,199	5,995	5	4.50	Amsterdam - Centrum West
1,200	2,400	2	0.00	Amsterdam - Centrum Oost

LOW LEVEL DESIGN (LLD)



Price vs Location

Room No	Room type	Price	Price Category	Neighborhood	Reviews	Accommodates	Total Earnings	Satisfaction Category
785432	Entire home/apt	240	Low	Centrum Oost	447	4	107,280	Very Good
49552	Entire home/apt	359	Moderate	Centrum West	284	6	101,956	Excellent
449775	Private room	312	Moderate	Westerpark	321	4	100,152	Excellent
366438	Private room	210	Low	Centrum Oost	344	4	72,240	Excellent
779225	Entire home/apt	567	High	Oostelijk Havengebied / Indische Buurt	124	6	70,308	Very Good
854328	Entire home/apt	359	Moderate	Westerpark	190	4	68,210	Excellent
57408	Entire home/apt	163	Low	Oostelijk Havengebied / Indische Buurt	410	4	66,830	Excellent
68290	Private room	143	Low	Centrum West	463	2	66,209	Excellent
42970	Private room	179	Low	Centrum West	366	2	65,514	Excellent
57841	Entire home/apt	174	Low	Centrum Oost	369	4	64,206	Excellent