

High Level Design

**AIRBNB DATA ANALYSIS**

Rajhans Prasad



Document version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date issued | Version | Description | Author |
| 02/06/2023 | 1.0 | First version of  complete HLD | Rajhans Prasad |
|  |  |  |  |
|  |  |  |  |



**CONTENTS**

**Document Version Control………………………………………………………………………….1**

**Abstract…………………………………………………………………………………………………2**

**1 Introduction .............................................................................................................................3**

**1.1 Why this High-Level Design Document?.........................................................................3**

**1.2 Scope ...............................................................................................................................3**

**2 General Description ................................................................................................................4**

**2.1 Product Perspective & Problem Statement .....................................................................4**

**Since 2008, guests and hosts have used AirBNB to expand on travelling possibilities and present more unique, personalized way of experiencing the world. This dataset describes the listing activity and metrics in San Diego, California for 2019…………………………………….4**

**2.2 Tools used........................................................................................................................4**

**3 Design Details.........................................................................................................................5**

**3.1 Functional Architecture ....................................................................................................5**

**3.2 Optimization .....................................................................................................................6**

**4 KPIs.........................................................................................................................................7**

**4.1 KPIs (Key Performance Indicators) .................................................................................8**

**5 Deployment.............................................................................................................................8**

**6 Power BI Dashboard Introduction………………………………………………………………..10**

**7 AirBnb Data Visualization Dashboard……………………………………………………………11**

**Abstract**

Airbnb is an online marketplace that connects people who want to rent out their property with people who are looking for accommodations in specific locales. Airbnb offers people an easy, relatively stress-free way to earn some income from their property.

These data can be used for various types of analysis depending on the scope of the specific work.

The data-set contains information about the hosts, geographical

Availability.



**1 Introduction**

**1.1**Why this High- Level Design Document ?

The purpose of the high level design document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding and can be used as a reference manual for how the module interacts at a high level.

The HLD will:

● Present all the design aspect and define them in detail

● Describe the user interface being implemented

● Describe the hardware and software interfaces

● Describe the performance requirements

● include design features and the architecture of the project ● list and describe the non functional attributes like

❖ Security

❖ Reliability

❖ Maintainability

❖ Portability

❖ Reusability

❖ Application compatibility

❖ Resource utilisation

❖ Serviceability

1.2 Scope

The HLD documentation presents The structure of the system, such as the database architecture, application architecture(layers), application flow(Navigation), and technology architecture. The HLD use non technical to mildly-technical terms which should be understandable to the administrators of the system.

**2. General Descriptions**

**2.1** Product perspective & problem statement



Since 2008, guests and hosts have used AirBNB to expand on travelling possibilities and present more unique, personalized way of experiencing the world.This dataset describes the listing activity and metrics in San Diego, California for 2019.

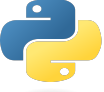
Objective of the project is to perform data visualisation techniques to understand the insight of the data .

This project aims to apply various business intelligence tools such as Tableau or Power Bi to get to a visual understanding of the data.

**2.2** Tools used

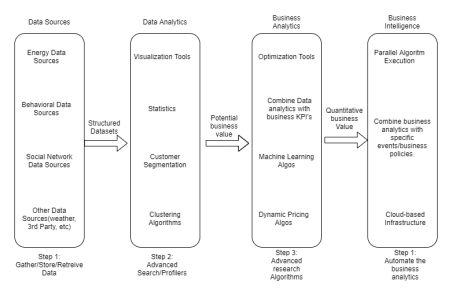
Business intelligence tools and libraries such as

Numpy,Pandas,Excel,R,Tableau,Power BI are used to build the whole framework.

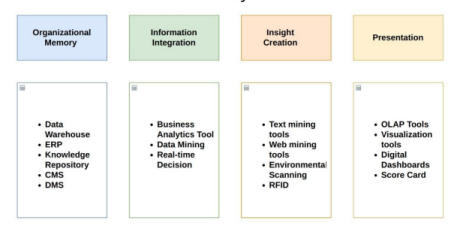


**3. Design Details**

**3.1** Functional Architecture

****

How BI Really Works





3.2 Optimization

**Your data strategy drives performance**

* Removing duplicate records from the dataset.
* Handling null values.
* Performing feature encoding on the dataset.
* Creating groups, hierarchy and calculated fields for easy analysis.
* Building Dashboard using Power BI.

**Reduce The marks(data points) in your view**

**●** Practice guided Analytics .There is no need to fit everything you plan to show in a single view. compiled related views and connected them with an action filter to travel from overview to highly- granular views at the speed of thought.

● Remove the unneeded dimensions from the detail self.

● Explore. Try displaying your data in different types of views.

**Limit your filters by number and type**

● Reduce the number of filters in use. Excessive filters on a view will create a more Complex query which takes longer to return results. double check your filters and remove any that are necessary.

● use and include filters. exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.

● Use a continuous date filter. Continuous date filters( relative and range-of- date filters) can take advantage of the induction properties in your database and are faster than discrete filters.

● Use Boolean or numeric filters. Computers process integer and booleans(t/f) much faster than strings.

● Use parameters and Action filters.These reduce the query load (and work across data sources).

**Optimise and materialise your calculations** ● Perform calculations in the database

● Reduce the number of nested calculations.

Reduce the granularity of load table calculations in the view. the more granular the calculation, the longer it takes .

❖ LODs - Look at the number of unique dimension members in the calculation.

❖ table calculations - the more marks in the view, the longer it will take to calculate.

● Where possible, use MIN or MAX instead of AVG.AVG requires more processing than MIN or MAX. often rows will be

duplicated and display the same result with MIN,MAX or AVG.

● Make groups with calculations. like include filters, calculated groups load only named members of the domain, whereas tableau’s Group function loads the entire domain.

● Use booleans or numeric calculations instead of string

calculations. Computers can process integer and booleans(t/f) Much faster than strings.

Boolean>int>Float>Date>Datetime>string.

**4. KPIs**

Dashboards Will be implemented to display and indicate certain KPIs and relevant indicators for the AirBnb.





As and when The system starts to capture the historical/

periodic data for a user, the dashboards will be included to

display charts over time with progress on various indicators or factors.

4.1 KPIs (Key Performance Indicators)

Key indicators displaying a summary of the AirBnb data and its relationship with different metrics.

1. Min & Max price of rooms.

2. Number of accomodates in neigherbood.

3. Room type by prices.

4. Top earners by Revenue.

5. Number of Bedrooms

**5.Deployment.**

Prioritising data and Analytics couldn't come at a better time. your company ,No matter what size , is already collecting data and most likely analysing just a person of it to solve business problems, gain competitive advantages, and drive enterprise transformation. with the explosive growth of enterprise data, database technology and the high demand for analytical skills,Today's most effective it organisations have shifted their focus to enabling self service by deploying and operating tableau at scale, as well as organising, orchestration ,And unifying desperate source of data for business users and Experts I like to author and consume content. Tableau prioritizes choice in flexibility to fit to, rather than dictate, your enterprise architecture. tableau server and tableau online leverage your existing Technology investments and integrate into your IT infrastructure to provide or self service, modern analysis platform for your users. with on premises, cloud and hosted options, there is a version of tableau Too much you requirements.Below is a comparison of the three types.

TYPE PROS AND CONS

**Tableau server - on premises**

● Full control of hardware and software.



● infrastructure and data remain behind your firewall.

● Need dedicated administrators to manage hardware and software.

● Additional infrastructure needed to access off

-network(mobile,external)

**Tableau server- Public cloud( laas )**

**●** Full control of software on managed hardware.

**●** puts infrastructure in same place as data**(**For migration to cloud**)**

**●** flexibility to spin up/down hardware as needed.

● need dedicated administrators to manage software.

● additional infrastructure needed to access off

network(mobile,external)

**Tableau online(SaaS)**

**●** Fully hosted solution( hardware, software upgrades)

**●** Fast to deploy

● Easy for external audience to access

● Single site in multi -tenant environment

● cubes are not supported

● no guest account access

Depending on your organisational roles and responsibilities.Tableau server should be installed by a systems administrator and the designated Tableau server administrator in coordination with the appropriate IT roles. For Tableau online, you will integrate with your existing Technology and configure the site settings. The Data & Analytics Survey completed by Business teams, identifies and Prioritizes data use cases, audience size, And users. you will use the information collected in both services to plan your deployment strategy. including sizing installation and configuration of your tableau server or integration and configuration of Tableau online. in addition to installing Tableau server or configuring tableau online.

administrators will also need to plan For the client software installation



of tableau prep builder, Tableau Desktop, Tableau bridge for Tableau online where applicable.

**6.Power BI Dashboard Introduction.**

Power BI is a business analytics service by Microsoft that provides interactive visualizations and business intelligence capabilities with an interface that is simple enough for end users to create their own reports and dashboards.

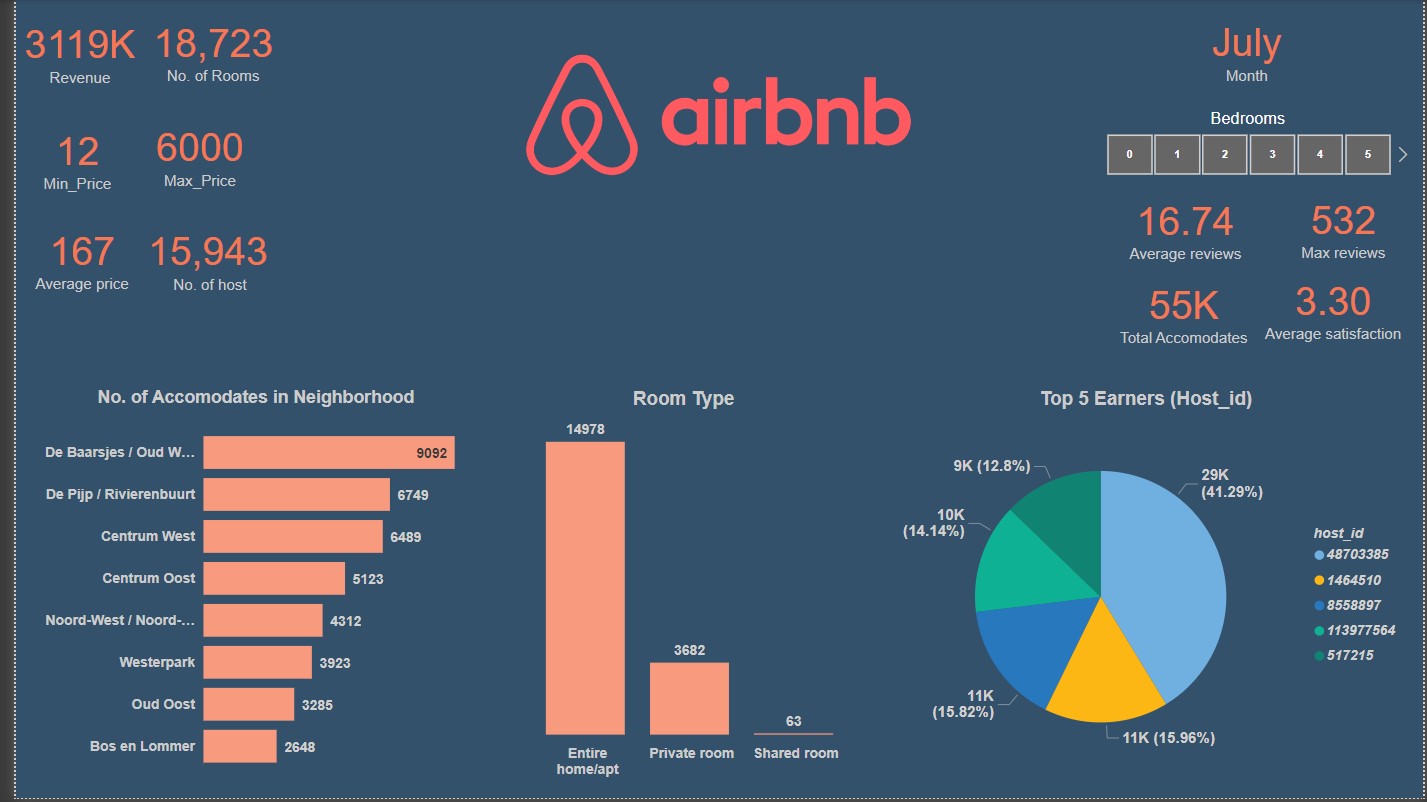
Power BI dashboards can be created using a variety of data sources, such as Excel spreadsheets, SQL databases, and cloud-based applications like Salesforce and Google Analytics. Power BI also provides connectors to many other data sources, making it easy to bring in data from multiple sources. Once data is connected, Power BI allows users to create visually appealing and interactive dashboards.

Users can drag and drop visualizations onto the canvas, customize the appearance of the dashboard, and add filters and slicers to allow for interactive exploration of the data.

Power BI dashboards also have a variety of sharing options. Users can share dashboards with others within their organization, publish dashboards to the web, or embed them into websites and other applications.

Power BI also has many advanced features, such as the ability to perform complex data modeling and calculations using DAX formulas, and the ability to create and share reports with others.Overall, Power BI dashboards are a powerful tool for organizations to gain insights into their data, improve decision-making, and drive business success.

**7.AirBnb Data Visualization Dashboard.**

****