High Level Design (HLD)

AirBNB Data Analysis

Author	Hemant Sharma, Anurag Gupta.
Document Version	HLD-V1.0
Last Revised Date	25/08/2023

Document Version Control

Date Issued	Version	Description	Author
25/08/2023	HLD-V1.0	First Version of Complete HLD	Hemant Sharma, Anurag Gupta.

Contents

Abstract.

Travel has always been a very unique experience to every individual. Expanding the possibilities of these unique experiences has been the goal of every travel company throughout the years. Aiding in travel and tourism, for individuals to explore cultures, traditions and beauty of places they visit, in the most customer - oriented way possible, has been the goal of Air BNB.

Knowing their customers is an essential in the process of achieving the said goal. Accurately predicting the and understanding customer requirements is not a perfectly attainable task. With the variations in demand of the customers and the locations they want to tour, many factors with trends do come into picture.

.

1. Introduction

1.1 Why this High-Level design document?

The purpose of this High-Level Design (HLD) 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 modules interact at a high level.

The HLD will:

- Present all of the design aspects 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 utilization
  - 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 uses non-technical to mildly-technical terms which should be understandable to the administrators of the syste

2. General Description

Product Perspective & Problem Statement.

Air BNB has aimed to expand on travelling possibilities and present more unique, personalized ways of experiencing the world. This project aims draw insights from the data obtained from San Diego, California, 2019. Some research questions need to be answered with respect to all the listings of past booking information.

Objectives: Research Questions include:

Regarding the Host,

- Who are top earners
- Is there any relationship between monthly earning and prices

Regarding the Neighborhood,

- Any particular location getting maximum number of bookingsPrice relation
- with respect to location

Regarding the reviews,

- Relationship between Quality and Price

Regarding Price

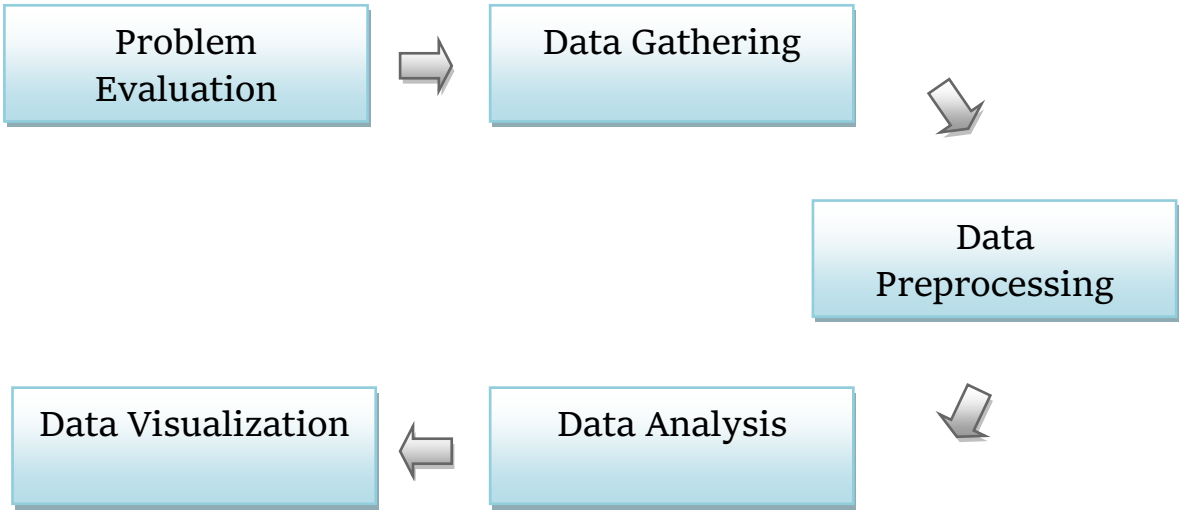
- Price v/s amenities
- Price v/s location

2.2 Tool Used

Business Intelligence tool Power BI is used to build the whole framework.

3. Design Details

3.1 Functional Architecture.



How BI Really works

Organization Memory	Information Integration	Insight Creation	Presentation
• Data Warehouse. • ERP • Knowledge Repository • CMS • DMS	• Business Analytics Tool • Data Minning • Real-Time Decision	• Text Mining Tools • Web Mining Tools • Environmental Scanning • RFID	• OLAP Tools • Visualization Tools • Digital Dashboards • Score Card

3.2 Optimization

Your data strategy drives performance

- Minimize the number of fields
- Minimize the number of records
- Optimize extracts to speed up future queries by materializing calculations,removing columns and the use of accelerated views

Reduce the marks (data points) in your view

- Practice guided analytics. There's no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly granular views at the speed of thought.
- Remove unneeded dimensions from the detail shelf.
- 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 aren't necessary.
- Use an include filter. 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 indexing properties in your databaseand are faster than discrete date filters.
- Use Boolean or numeric filters. Computers process integers and Booleans(t/f) much faster than strings.
- Use parameters and action filters. These reduce the query load (and work across data sources).

Optimize and materialize your calculations

- Perform calculations in the database Reduce
  the number of nested calculations.
- Reduce the granularity of LOD or table calculations in the view. Themore granular the calculation, the longer it takes.
- LODs - Look at the number of unique dimension members in the calculation. o 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 integers and Booleans (t/f) much faster thanstrings. Boolean>Int>Float>Date>Date Time>String

4. KPI's

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the food sales, the dashboards will be included to display charts over time with progress on various indicators or factors.

Key indicators displaying a summary of the Air BNB customers and withdifferent metrics

- Total Neighbourhoods
- Count of total number of hosts
- Neighbourhood with maximum bookings
- Neighbourhood with minimum bookings

Charts

Top hosts with maximum bookings (Column chart)

Average price by reviews
Average Price by room types
Bookings in Neighbourhood groups (Donut chart)
Top earning Hosts (Horizontal Bar chart)
Preferred Room Types (Pie chart)

5. Deployment.

Prioritizing data and analytics couldn't come at a better time. Yourcompany, no matter What size, is already collecting data and most likelyanalyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today's most effective IT organizations have shifted their focus to enabling self-service by deploying and operating Power BI at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

Air BNB data has been cleaned using ETL tool Power Query and analyzed on Power BI which will provide a better key insight for the data and tell you a better story of the raw data. The Power BI report is published on workspace where you can play with the data for necessary insights.