

# **LOW LEVEL DESIGN DOCUMENT**

**(Airbnb Data Analysis)**

Akshay Salve

VERSION: 1.0

DATED: 26/06/2022

## **Document Version Control:**

### **Airbnb Data Analysis – Business Intelligence Project**

<b>Version</b>	<b>Date</b>	<b>Author</b>	<b>Change</b>
1.0	27/06/2022	Akshay R. Salve	First version of LLD

## **Abstract**

Airbnb, Inc. is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Based in San Francisco, California, the platform is accessible via website and mobile app. Airbnb does not own any of the listed properties; instead, it profits by receiving commission from each booking. The company was founded in 2008 by Brian Chesky, Nathan Blecharczyk, and Joe Gebbia. Airbnb is a shortened version of its original name, AirBedandBreakfast.com.

This dataset provides huge data on based on the survey conducted in July 2017. Based on this information the ultimate goal is to find the insights highlighting key indicators and metrics that influence the travel market.

## Contents:

<b>1. Introduction.....</b>	<b>05</b>
1.1. What is Low-Level Design Document?.....	05
1.2 Scope.....	05
<b>2. Architecture.....</b>	<b>05-06</b>
<b>3. Architecture Description.....</b>	<b>06</b>
3.1. Data Sourcing.....	06-07
3.2. Data Overview.....	07
3.3. Data Description.....	07-08
3.4. Data loading in Power BI Query Editor.....	08
3.5. Data to Insights through Visualizations and Excel Data Analysis.....	09-10
<b>4. Deployment to Power BI Service.....</b>	<b>10-11</b>

# 1. Introduction

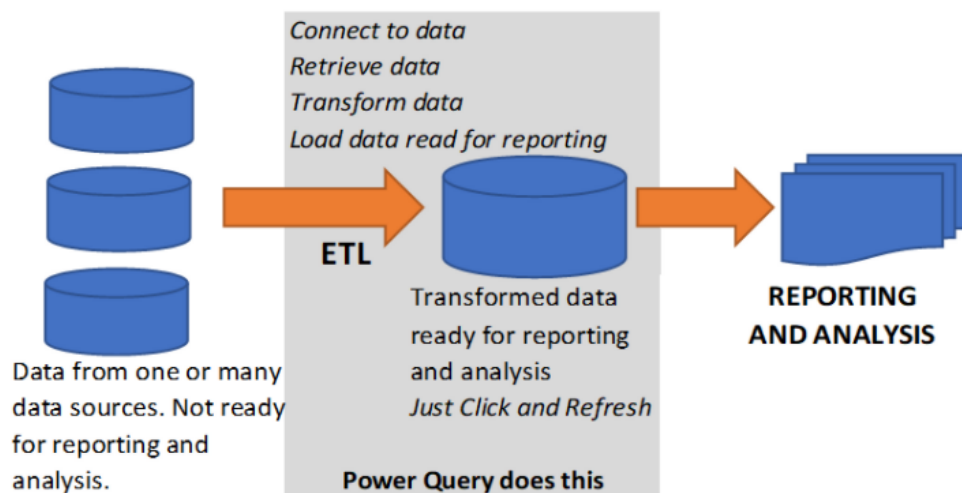
## 1.1. Why this Low-Level Design Document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Bank Marketing Campaign Analysis. LDD 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. 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.

# 2. Architecture:



## Low Level Design (LLD)

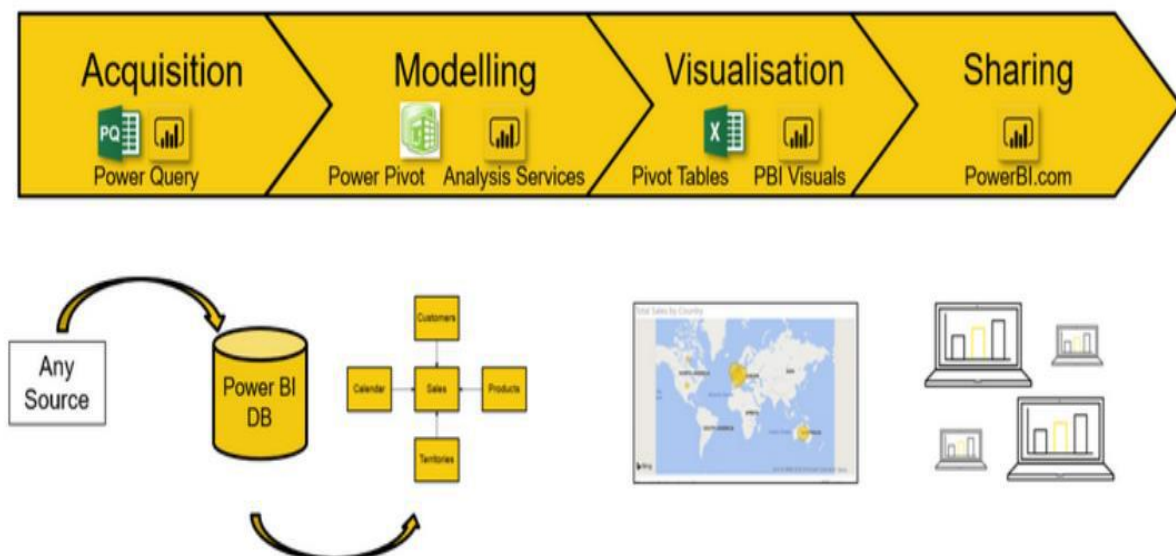
ETL (extract, transform and load) in Power BI uses preparation of data sets for analysis by removing irregularities in the data. It also involves data visualization to draw meaningful patterns and insights.

Based on the results of ETL, companies also make business decisions, which can have repercussions later.

- If ETL is not done properly then it can damage the business a lot in many ways such as loss of client which we are working for, the decision making will go completely wrong and many more issues.
- If done well, it may improve the efficacy of everything we do next.

Below are following steps to follow for ETL:

1. Data Sourcing
2. Data Cleaning
3. Data Modelling
4. Data Visualization



## 3. Architecture Description:

### 3.1 Data Sourcing:

The dataset is in csv (comma separated values) format. MS Excel is used to load the data.

### Citation Request:

This dataset is provided by the client.

1.Title - Airbnb\_prices.csv.

### 3.2. Data Overview –

- ❖ The data includes single csv file with all examples, ordered by date (22 July 2017 to 23 July 2017).
- ❖ The Number of Instances – 18724 for Airbnb\_price.csv
- ❖ Number of Attributes – 19 Attributes

### 3.3. Data Description –

- ❖ room\_id – Ids for rooms [10176931, 8935871, 14011697, ..., 19859427, 17132164, 7605782]
- ❖ survey\_id – Id of survey conducted [1476]
- ❖ host\_id – Ids for hosts [49180562, 46718394, 10346595, ..., 139135665, 1501422, 29724632]
- ❖ room\_type – Type of rooms ['Shared room', 'Entire home/apt', 'Private room']
- ❖ country – empty
- ❖ city – Amsterdam
- ❖ borough – empty
- ❖ neighbourhood – Names of neighbourhoods offering services ['De Pijp / Rivierenbuurt', 'Centrum West', 'Watergraafsmeer', 'De Baarsjes / Oud West', 'Oostelijk Havengebied / Indische Buurt', 'Westerpark', 'Oud Oost', 'Centrum Oost', 'Geuzenveld / Slotermeer', 'Buitenveldert / Zuidas', 'Oud Noord', 'Bos en Lommer', 'Slotervaart', 'Bijlmer Centrum', 'Ijburg / Eiland Zeeburg', 'Noord-West / Noord-Midden', 'De Aker / Nieuw Sloten', 'Noord West', 'Bijlmer Oost', 'Osdorp', 'Noord Oost', 'Gaasperdam / Driemond', 'Westpoort']
- ❖ reviews – numeric
- ❖ overall\_satisfaction – numeric
- ❖ accommodates – numeric
- ❖ bedrooms – numeric
- ❖ bathrooms – empty
- ❖ price – numeric
- ❖ mainstay – empty
- ❖ name – names of stays
- ❖ last\_modified – Date of survey
- ❖ latitude – North-South coordinates of the room location

## Low Level Design (LLD)

- ❖ longitude – East-West coordinates of the room location
- ❖ location – Exact location of the room

### 3.4 Data loading in Power BI Query Editor

Power Query is the data connectivity and data preparation technology that enables end users to seamlessly import and reshape data from within a wide range of Microsoft products, including Excel, Power BI, Analysis Services, data verse, and more with the following characteristics:

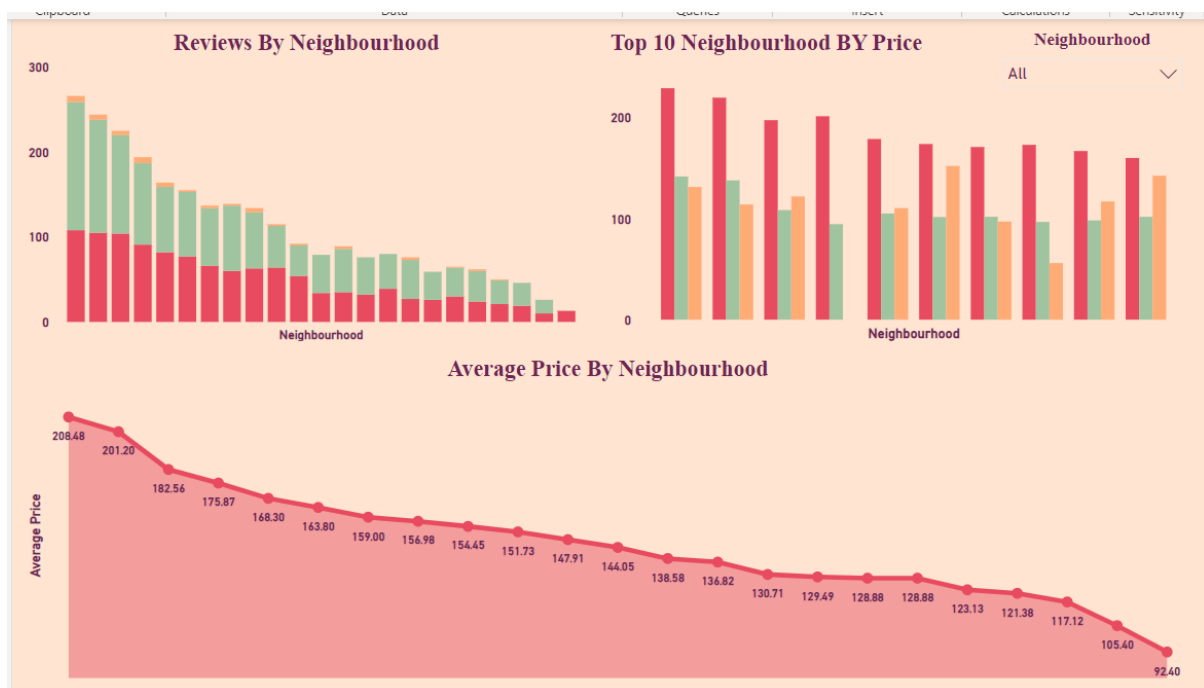
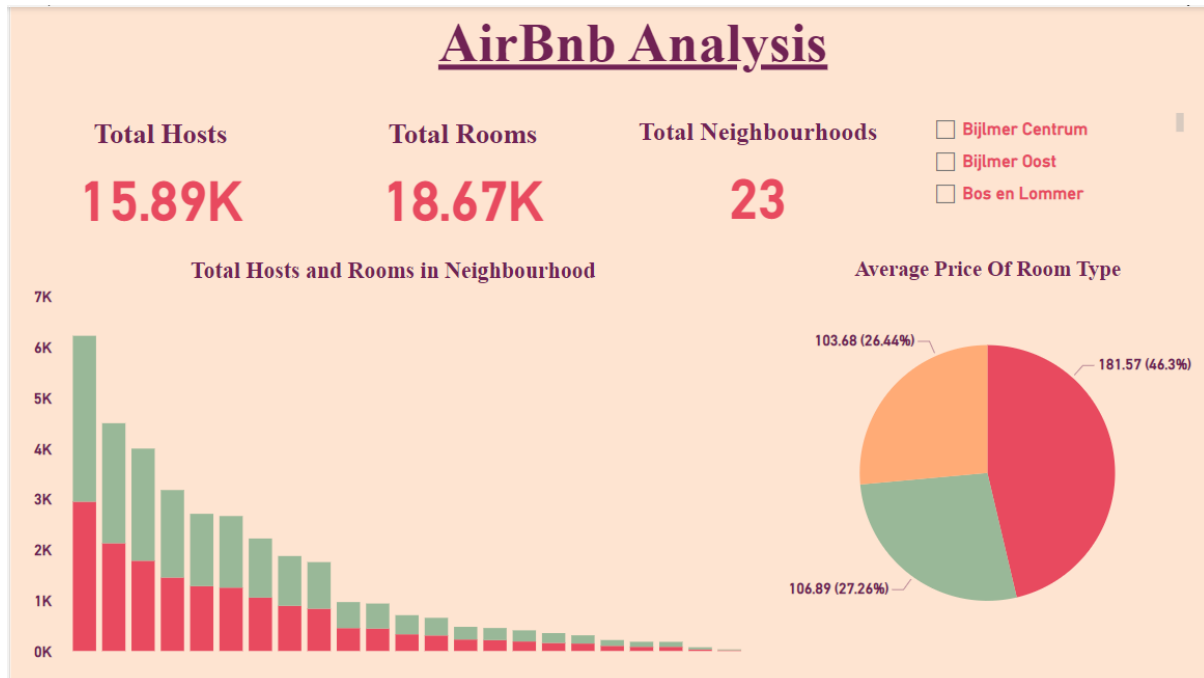
- ❖ There can be multiple rows and columns in the data.
- ❖ Each row represents a sample of data,
- ❖ Each column contains a different variable that describes the samples (rows).
- ❖ The data in every column can be a different type of data – e.g. numbers, strings, dates, Boolean etc.

The screenshot displays the Power BI Query Editor interface. The main area shows a table with the following columns: `room_id`, `survey_id`, `host_id`, `room_type`, `city`, `neighborhood`, and a status column. The table contains 25 rows of data. The right sidebar shows the 'APPLIED STEPS' pane with a list of transformations: Source, Promoted Headers, Changed Type, Removed Errors, Filtered Rows, Filtered Rows1, Changed Type1, Filtered Rows2, Added Conditional Column, Removed Columns, Added Conditional Column1, Changed Type2, Removed Columns1, Added Conditional Column2, and Removed Columns2.

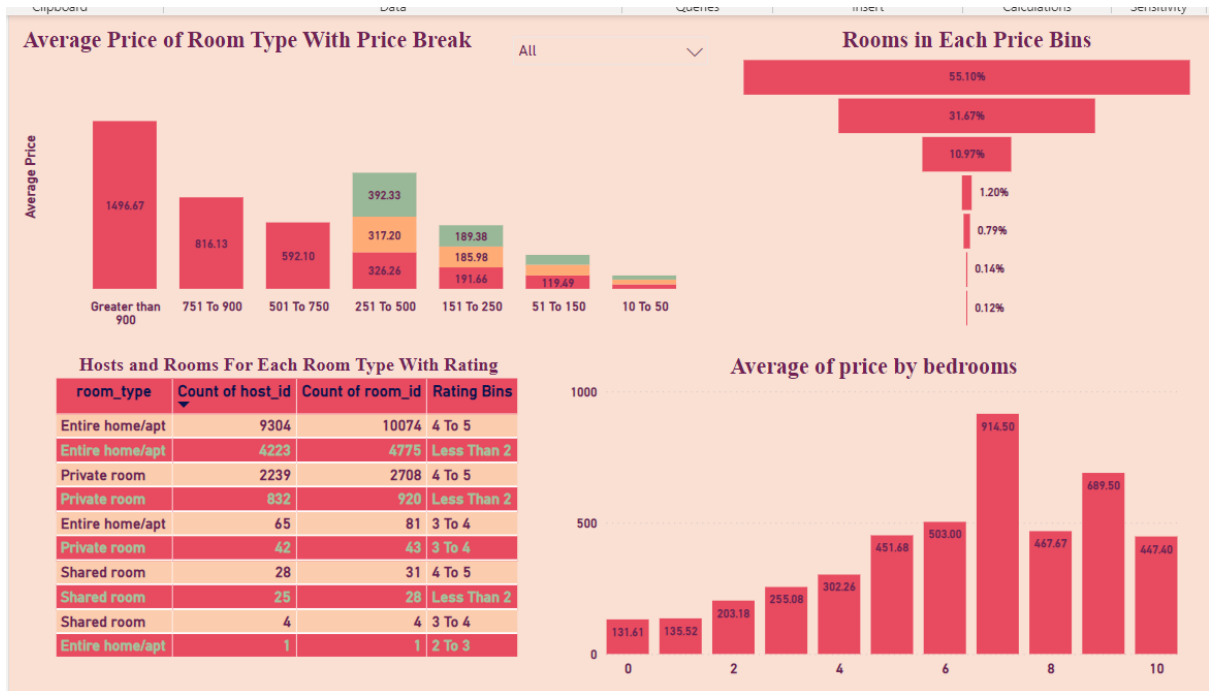
	room_id	survey_id	host_id	room_type	city	neighborhood	Status
1	10176931		1476	Shared room	Amsterdam	De Pijp / Rivierenbuurt	Valid
2	89358771		1476	Shared room	Amsterdam	Centrum West	Valid
3	14011697		1476	Shared room	Amsterdam	Watergraafsmeer	Valid
4	6137978		1476	Shared room	Amsterdam	Centrum West	Valid
5	18630616		1476	Shared room	Amsterdam	De Baarsjes / Oud West	Valid
6	5790170		1476	Shared room	Amsterdam	De Pijp / Rivierenbuurt	Valid
7	934060		1476	Shared room	Amsterdam	Oostelijk Havengebied / Indische Buurt	Valid
8	19590049		1476	Shared room	Amsterdam	Westerpark	Valid
9	5020280		1476	Shared room	Amsterdam	Oud Oost	Valid
10	15810783		1476	Shared room	Amsterdam	Centrum West	Valid
11	9188521		1476	Shared room	Amsterdam	De Pijp / Rivierenbuurt	Valid
12	10162121		1476	Shared room	Amsterdam	Centrum Oost	Valid
13	14217287		1476	Shared room	Amsterdam	Centrum Oost	Valid
14	19304330		1476	Shared room	Amsterdam	De Baarsjes / Oud West	Valid
15	9060570		1476	Shared room	Amsterdam	Oostelijk Havengebied / Indische Buurt	Valid
16	13797285		1476	Shared room	Amsterdam	Oostelijk Havengebied / Indische Buurt	Valid
17	8692643		1476	Shared room	Amsterdam	Geuzenveld / Sloterveer	Valid
18	12314602		1476	Shared room	Amsterdam	De Pijp / Rivierenbuurt	Valid
19	18985831		1476	Shared room	Amsterdam	Buitenveldert / Zuidas	Valid
20	11122903		1476	Shared room	Amsterdam	Buitenveldert / Zuidas	Valid
21	9086830		1476	Shared room	Amsterdam	Oud Noord	Valid
22	9537235		1476	Shared room	Amsterdam	Westerpark	Valid
23	14309046		1476	Shared room	Amsterdam	De Baarsjes / Oud West	Valid
24	19243789		1476	Shared room	Amsterdam	De Pijp / Rivierenbuurt	Valid
25	18117741		1476	Shared room	Amsterdam	Rekenen / Centrum	Valid



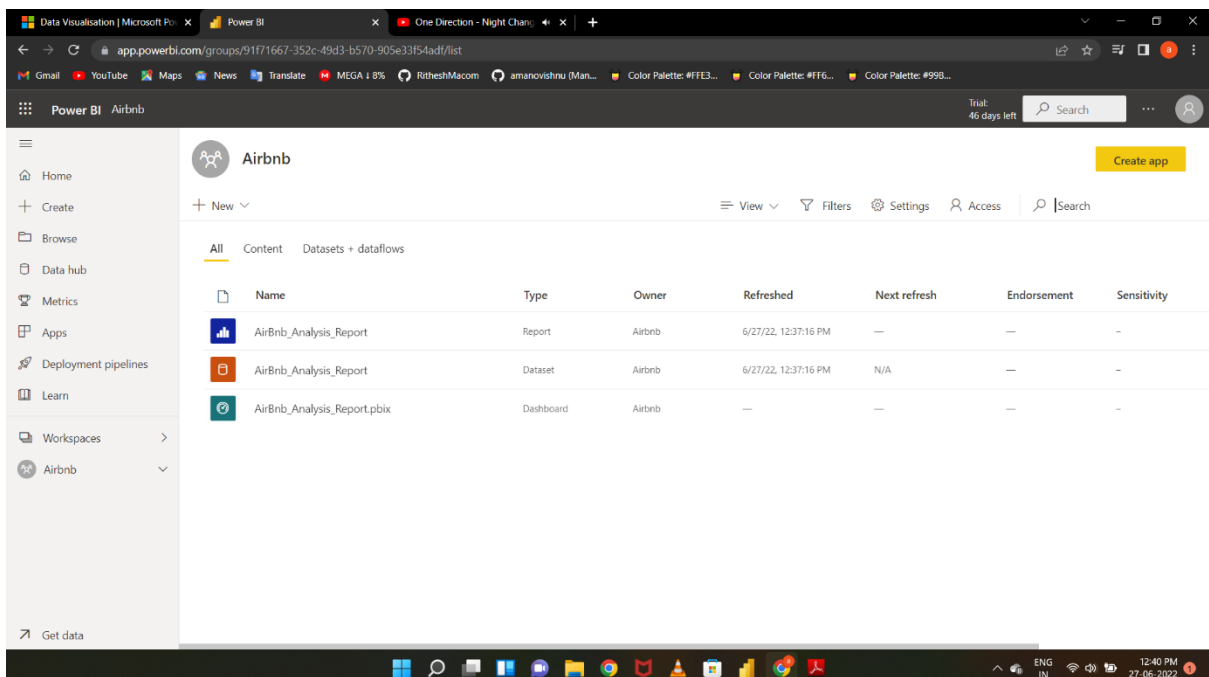
### 3.5 Data to Insights through Visualizations and Excel Data Analysis



## Low Level Design (LLD)



## 3. Deployment to Power BI Service



## Low Level Design (LLD)

