

Low Level Design (LLD)

(Swiggy Data Analysis)

Written By / Author	Prateek Magdum
Document Version	LLD-V1.0
Last Revised Date	05/04/2022

Document Version Control

Date Issued	Version	Description	Author
15/02/2022	LLD-V1.0	First Version of Complete LLD	Lokesh Attarde

Contents

Document Version Control	2
Abstract	4
1 Introduction	5
1.1 Why this Low-Level design document?	5
1.2 Scope	5
1.3 Constraints	5
2 Technical Specifications	6
2.1 Swiggy Dataset	6
2.1.1 Swiggy Dataset Overview	6
2.1.2 Overview of Variables	7
3 Architecture	8
3.1 Architecture Description	8
3.1.1 Data Description	8
3.1.2 Define the Use Cases	8
3.1.3 Import the Dataset	9
3.1.4 Exploratory Data Analysis (EDA)	9
3.1.5 Data Pre-processing, Data Cleaning & Imputation (Handling the Categorical & Numerical Variables)	9
3.1.6 Analyse the Data	10
3.1.7 Visualize & Share Meaningful Insights	10
4 Technology Stack	11

Abstract

The online food ordering market includes foods prepared by restaurants, prepared by independent people, and groceries being ordered online and then picked up or delivered. The first online food ordering service, World Wide Waiter (now known as Waiter.com), was founded in 1995. Online food ordering is the process of ordering food from a website or other application. The product can be either ready-to-eat food or food that has not been specially prepared for direction consumption.

In the world of rising new technology and innovation, Food 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 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

1.1 Why this Low-Level design document?

The purpose of this LLD or a Low-Level Design (LLD) document is to give the internal logical design of the actual program code for Swiggy Data Analysis project. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of this project and will be proposed to the higher management for its approval.

The main objective of the project is to analyse the various aspects with different use cases which covers many aspects of Swiggy Food Delivery Service. 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.

1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step-by step refinement process. This 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.

This study demonstrates the how different analysis help out to make better business decisions and help analyse customer trends and satisfaction, which can lead to new and better products and services.

1.3 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.

2 Technical Specifications

2.1 Swiggy Dataset –

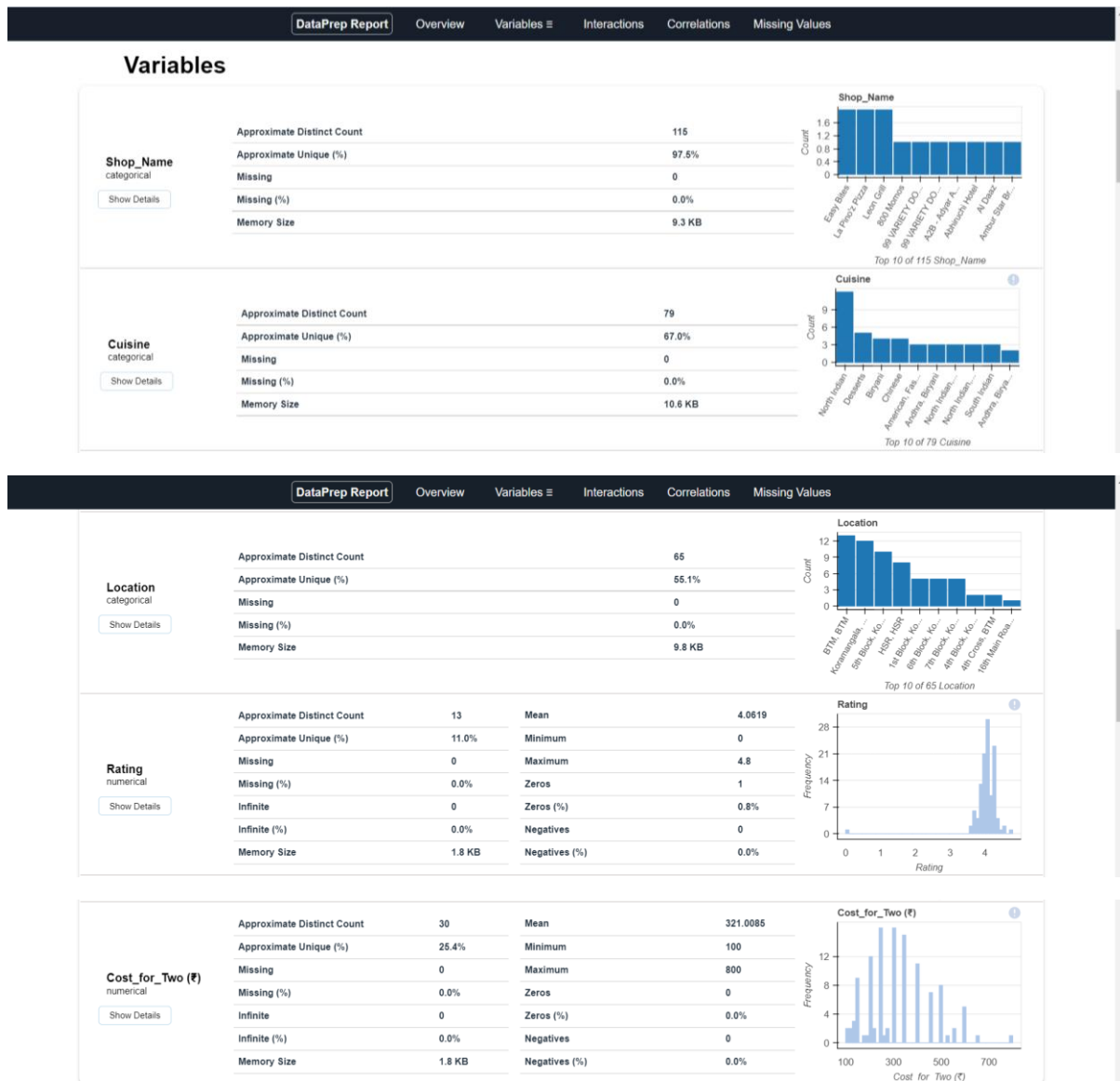
	A	B	C	D	E
1	Shop_Name	Cuisine	Location	Rating	Cost_for_Two
2	Kanti Sweets	Sweets	Koramangala, Koramangala	4.3	₹, 150
3	Mumbai Tiffin	North Indian, Home Food, Thalys, Combo	Sector 5, HSR	4.4	₹, 400
4	Sri Krishna sagar	South Indian, North Indian, Fast Food, Beverages, Jain	6th Block, Koramangala	4.1	₹, 126
5	Al Daaz	American, Arabian, Chinese, Desserts, Fast Food, Mughlai, North Indian	HSR, HSR	4.4	₹, 400
6	Beijing Bites	Chinese, Thai	5th Block, Koramangala	4.1	₹, 450
7	Kitchens of Punjab	North Indian	Koramangala 4th Block, Koramangala	4.2	₹, 350
8	99 VARIETY DOSA AND PAV BHAJI	Fast Food, North Indian, Chinese	BTM 2nd Stage, BTM	4.1	₹, 200
9	La Pino'z Pizza	Italian	BTM, BTM	3.9	₹, 500
10	Hotel Manu	South Indian, Kerala, Chinese, North Indian	HSR, HSR	4.1	₹, 350
11	Yumlane Pizza	Pizzas, Italian, Mexican	9th Main road, Koramangala	3.8	₹, 150
12	Ambur Star Briyani	Chinese, South Indian, North Indian, Desserts, Fast Food, Kerala, Andhra	outer ring road, BTM	4.1	₹, 500
13	Cake Box	Desserts	Koramangala, Koramangala	4	₹, 247
14	Meghana Foods	Chinese, Andhra, Briyani, Seafood	5th Block, Koramangala	4.3	₹, 550
15	Momoz	Chinese	5th Block, Koramangala	4.3	₹, 450
16	A2B - Adyar Ananda Bhavan	South Indian, Chinese, Desserts, North Indian	7th Block, Koramangala	4.2	₹, 450
17	Shawarma Inc	Arabian, Fast Food	1st MAin, Koramangala	4.1	₹, 150
18	WarmOven Cake & Desserts	Desserts, Beverages	Koramangala, Koramangala	4.1	₹, 200
19	Sri Lakshmi Dhaba	North Indian	Bommanahalli, BTM	3.7	₹, 200
20	Falahaar & Kota Kachori	North Indian	6th block, Koramangala	4.2	₹, 300
21	Shree Khana Khazana	Indian, Rajasthani	Sector 4, HSR	4.1	₹, 350
22	Just Bake - Cakes & confectioners	Desserts, Bakery	BTM 1st stage, BTM	4.3	₹, 300
23	Maa Di Hatti	Chinese, Healthy Food, North Indian	Jakkasandra Extn, Koramangala	4	₹, 129
24	Hotel Godavari	North Indian, Chinese, Hyderabad	Marutnagar Main Road, BTM	4	₹, 400
25	Rolls On Wheels	Fast Food	1st Block, Koramangala	4.1	₹, 300
26	New Udupi Grand	Chinese, Jain, North Indian, South Indian	HSR, HSR	4.3	₹, 150
27	Swad Punjab da	Indian	BTM, BTM	4.1	₹, 250

2.1.1 Swiggy Dataset Overview –

The Listings dataset consists of a table with 118 records and 5 features. Features are distributed as 2 Continuous features and 3 Categorical features. There are a total 0% of records having Missing Values. In short, there are no Missing Values present in the dataset.

DataPrep Report	Overview	Variables	Interactions	Correlations	Missing Values
Overview					
Dataset Statistics			Dataset Insights		
Number of Variables	5		Rating is skewed	Skewed	
Number of Rows	118		Cost_for_Two (₹) is skewed	Skewed	
Missing Cells	0		Shop_Name has a high cardinality: 115 distinct values	High Cardinality	
Missing Cells (%)	0.0%		Cuisine has a high cardinality: 79 distinct values	High Cardinality	
Duplicate Rows	0		Location has a high cardinality: 65 distinct values	High Cardinality	
Duplicate Rows (%)	0.0%				
Total Size in Memory	29.0 KB				
Average Row Size in Memory	251.4 B				
Variable Types	Categorical: 3 Numerical: 2				

2.1.2 Overview of Variables:



3 Architecture



3.1 Architecture Description –

3.1.1 Data Description –

As we have seen earlier, in our Swiggy dataset, we have around 118 records with 5 different features. Features are distributed as 2 Continuous features and 3 Categorical features. These datasets are given in the form of Comma Separated Value (.csv) format.

3.1.2 Define the Use Cases –

At this stage, based on the given dataset and business problems we have defined the several Use Cases to perform the analysis on and this will definitely help out get the key insights from this data based on which business decisions will be taken. Furthermore, 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.

3.1.3 Import the Dataset –

As we have received the dataset in the form of Comma Separated Value (.csv) format, therefore we can import the same using Pandas `read_csv()` function.

Reading Data

```
In [2]: df_Swiggy = pd.read_csv('Swiggy Bangalore Outlet Details.csv', sep = ',')
df_Swiggy
```

Out[2]:

	Shop_Name	Cuisine	Location	Rating	Cost_for_Two
0	Kanti Sweets	Sweets	Koramangala, Koramangala	4.3	₹ 150
1	Mumbai Tiffin	North Indian, Home Food, Thalis, Combo	Sector 5, HSR	4.4	₹ 400
2	Sri Krishna sagar	South Indian, North Indian, Fast Food, Beverag...	6th Block, Koramangala	4.1	₹ 126
3	Al Daaz	American, Arabian, Chinese, Desserts, Fast Foo...	HSR, HSR	4.4	₹ 400
4	Beijing Bites	Chinese, Thai	5th Block, Koramangala	4.1	₹ 450
...
113	Wok Paper Scissors	Pan-Asian, Chinese, Asian	JNC Road, Koramangala	3.9	₹ 219
114	Savoury Restaurant	Arabian, Middle Eastern, North Indian, Grill, ...	Madiwala, BTM	4.1	₹ 600
115	Royal Treat	North Indian, Chinese, Seafood, Biryani	5th block Koramangala, Koramangala	4.2	₹ 193
116	Thali 99	North Indian	Koramangala, Koramangala	4.3	₹ 200
117	Mani's Dum Biryani	Andhra, Biryani	1st Block, Koramangala	4.2	₹ 400

118 rows × 5 columns

3.1.4 Exploratory Data Analysis (EDA) –

- "Exploratory Data Analysis" (EDA) is a "Data Exploration" step in the Data Analysis Process, where a number of techniques are used to better understand the dataset being used.
- Understanding the Dataset can refer to a number of things including but not limited to...
 - Extracting Important "Variables".
 - Identifying "Outliers", "Missing Values", or "Human Error".
 - Understanding the Relationships between variables.
 - Ultimately, maximizing our insights of a dataset and minimizing potential "Error" that may occur later in the process.
- In other words, it will give you a better Understanding of the "Variables" and the "Relationships" between them.
- Here, we make use of dataprep module to automate our EDA process.
- It provides the following information:
 - Overview: detect the types of columns in a DataFrame.
 - Variables: variable type, unique values, distinct count, missing values
 - Quartile statistics like minimum value, Q1, median, Q3, maximum, range, interquartile range
 - Descriptive statistics like mean, mode, standard deviation, sum, median absolute deviation, coefficient of variation, kurtosis, skewness.
 - Correlations: highlighting of highly correlated variables, Spearman, Pearson and Kendall matrices
 - Missing Values: Bar Chart, Heatmap and spectrum of missing values.

DataPrep Report	Overview	Variables	Interactions	Correlations	Missing Values
Overview					
Dataset Statistics			Dataset Insights		
Number of Variables	5		<u>Rating</u> is skewed	Skewed	
Number of Rows	118		<u>Cost_for_Two (*)</u> is skewed	Skewed	
Missing Cells	0		<u>Shop_Name</u> has a high cardinality: 115 distinct values	High Cardinality	
Missing Cells (%)	0.0%		<u>Cuisine</u> has a high cardinality: 79 distinct values	High Cardinality	
Duplicate Rows	0		<u>Location</u> has a high cardinality: 65 distinct values	High Cardinality	
Duplicate Rows (%)	0.0%				
Total Size in Memory	29.0 KB				
Average Row Size in Memory	251.4 B				
Variable Types	Categorical: 3 Numerical: 2				

3.1.5 Data Pre-processing, Data Cleaning & Imputation (Handling the Categorical & Numerical Variables) –

Data pre-processing is a process of preparing the raw data and making it suitable for our analysis purpose, where we have to do lot of Data Cleaning, handle the missing values by using appropriate imputation techniques and based on that variable nature i.e. either of Categorical & Numerical variable. Here, in this project, we have done the substitution/imputation of missing values using either mean, median or mode according to the nature of those variables. Moreover, we also removed the columns which are does not participate in our analysis.

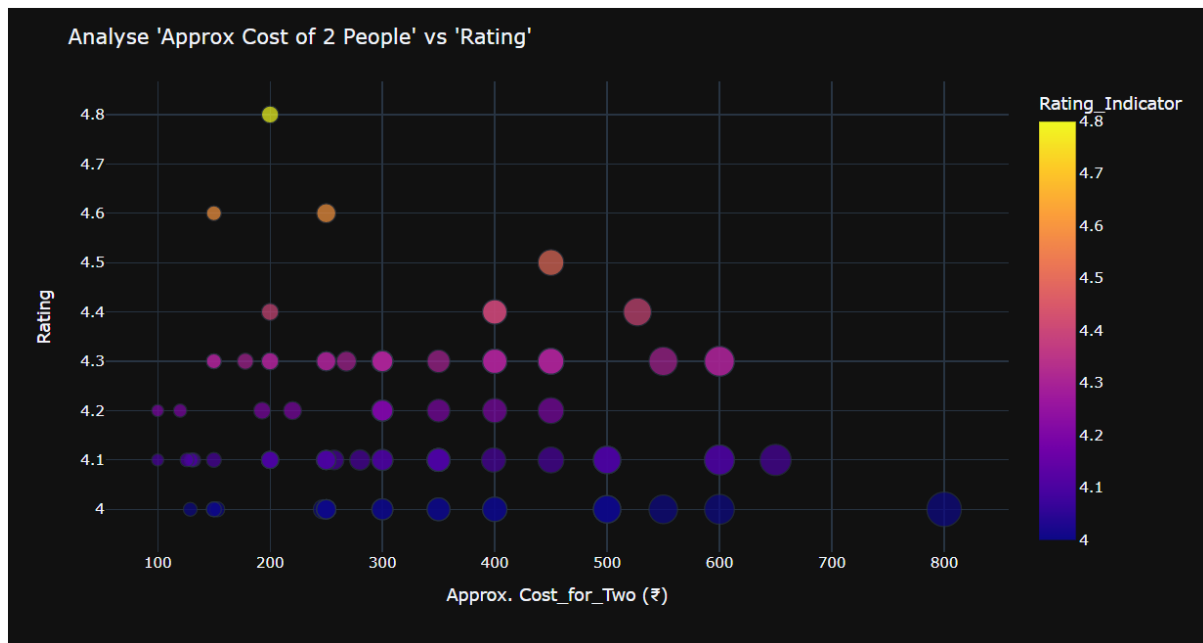
3.1.6 Analyse the Data –

Once the pre-processing is done, we are good to go with our actual analysis where we write lines of codes and logics to prepare our data as per the defined use cases.

3.1.7 Visualize & Share Meaningful Insights –

Finally, it's time to turn our data into some sort of visual representation. In short, Data visualization is the process of translating large data sets and metrics into charts, graphs and other visuals such as Bar Plot, Pie Chart, Heat map, Box Plot, Scatter Plot, and many more. The resulting visual representation of data makes it easier to identify and share insights about the information represented in the data.

Here is the beautiful glimpse of one of our visuals are –



All those different analysis help out to make better business decisions and help analyse customer trends and satisfaction, which can lead to new and better products and services.

4 Technology Stack

Data Manipulation Library	Pandas
Visualization Library	Matplotlib, Seaborn, Plotly, WordCloud, etc
EDA	dataprep
Dataset	.CSV Format
IDE	Jupyter Notebook