Low Level Design (LLD)

Analyzing Swiggy : Bangalore delivery outlet data

Revision Number: 1.0

Last date of revision: 10/08/2024

Saurabh Kumar Kushwaha

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| **10/08/2024** | LLD-V1.0 | First Version of Complete LLD | Saurabh Kumar Kushwaha |
|  |  |  |  |

**Contents**

Document Version Control **2**

**Abstract 4**

1. **Introduction 5** 
   1. Why this Low-Level design document? 5
   2. Scope 5
   3. Constraints 5
2. **Technical Specifications 6** 
   1. Swiggy Bangalore Outlet Dataset 6
      1. Swiggy Bangalore Outlet Dataset Overview 6
      2. Overview of Variables 7
3. **Architecture 8** 
   1. Architecture Description 8
      1. Data Description 8
      2. Define the Use Cases 8
      3. Import the Dataset 8
      4. Exploratory Data Analysis (EDA) 9
      5. Data Pre-processing, Data Cleaning & Imputation (Handling the Categorical &

Numerical Variables) 9

* + 1. Analyse the Data 10
    2. Visualize & Share Meaningful Insights 10

1. **Technology Stack 11**

# Abstract

The online food ordering market encompasses a wide range of options, including meals prepared by restaurants, individuals, and groceries that can be ordered online and either picked up or delivered. The initial online food ordering service, initially named World Wide Waiter but now known as Waiter.com, was established in 1995. Online food ordering involves the act of requesting food through a website or a similar application. The items available can consist of ready-to-consume dishes or ingredients that have not been specifically prepared for immediate consumption.

In today's era marked by technological advancements and innovation, the food industry is progressing by integrating Data Science and Analytics. Leveraging data analysis can offer a fresh perspective on their operations and assist in enhancing service quality by pinpointing areas of weakness within the business. This research illustrates how diverse analyses contribute to informed business decisions and aid in examining customer patterns and contentment levels, ultimately paving the way for improved products and services. Various analyses, such as Extract, Transform, Load (ETL) Analysis and Descriptive Analysis, have been conducted across a range of scenarios to extract key insights from the data, which will serve as the foundation for making strategic business choices.

# 1 Introduction

## Why this Low-Level design document?

## The necessity of the Low-Level Document (LLD) is crucial in the context of the "Analyzing Swiggy" project due to its comprehensive depiction of the system's design and operational aspects. The LLD encompasses details about the system's structure, data movement, and the collaborative components responsible for generating the intended outcomes.

## Within this endeavor, the LLD encompasses the process of Extracting, Transforming, and Loading (ETL), which encompasses the retrieval of data from the source file, its conversion into a suitable format, and subsequent storage within a database. Moreover, the LLD incorporates an examination of the Swiggy dataset, unveiling insights pertaining to delivery patterns, popular culinary choices, and highly-rated eateries.

## The LLD serves as an instructive tool for developers and stakeholders engaged in the project, facilitating their comprehension of the system's mechanics, constituent elements, and the trajectory of data through the system. Additionally, the LLD aids in the identification of potential bottlenecks, performance challenges, and zones open for enhancement.

## In essence, the Low-Level Document is indispensable for the success of the "Analyzing Swiggy" initiative, guaranteeing a well-crafted, efficient system that aligns with project prerequisites.

## 1.2 Scope

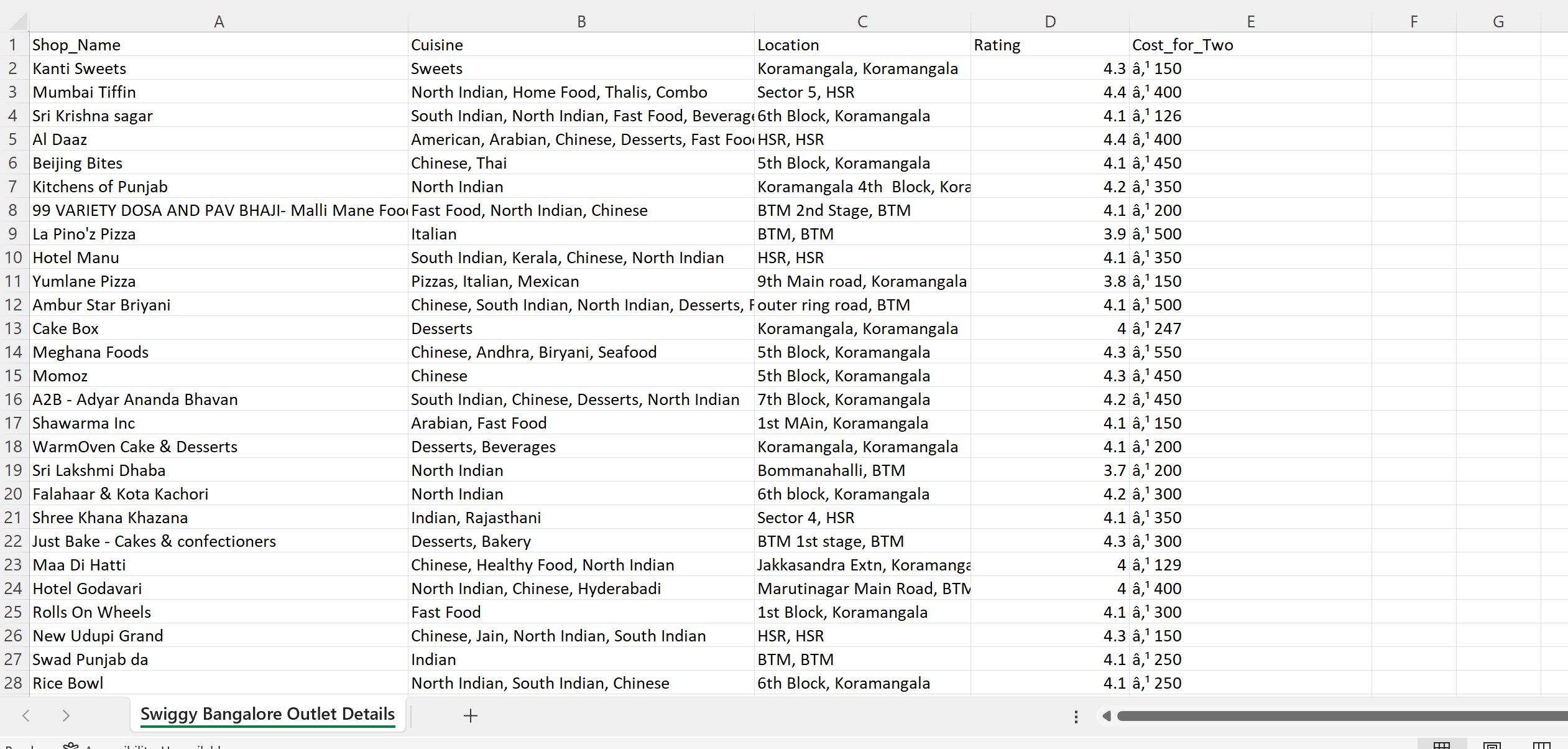
Low-level design (LLD) refers to the process of designing individual components, characterized by a systematic and iterative refinement approach. It finds its application in crafting data structures, essential software architecture, source code, and culminates in the development of efficient algorithms. The arrangement of data might be initially outlined during the requirement analysis phase and subsequently perfected during the data design phase. This investigation illustrates the ways in which diverse analyses contribute to enhancing business choices and aid in assessing customer patterns and contentment. Such insights can pave the way for improved offerings, ushering in novel and superior products and services.

## 1.3 Constraints

The examination needs to have a user-friendly approach, with well-organized and tidy code. The process of Extract, Transform, Load (ETL) should be highly automated to maximize time savings. Additionally, it's essential that users do not need any coding expertise, as the detailed insights they seek are provided alongside relevant visuals.

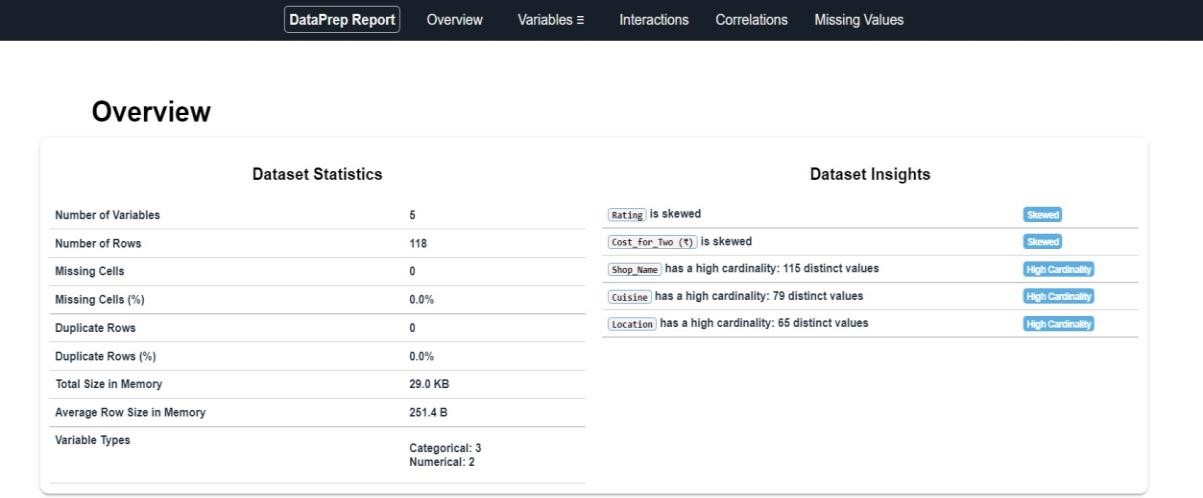
# 2 Technical Specifications

## 2.1 Swiggy Bangalore Outlet Dataset

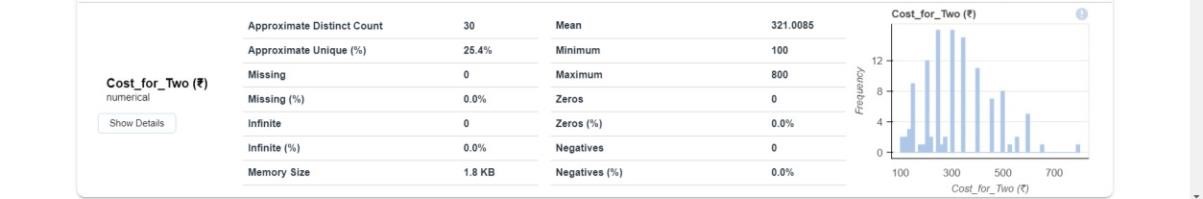
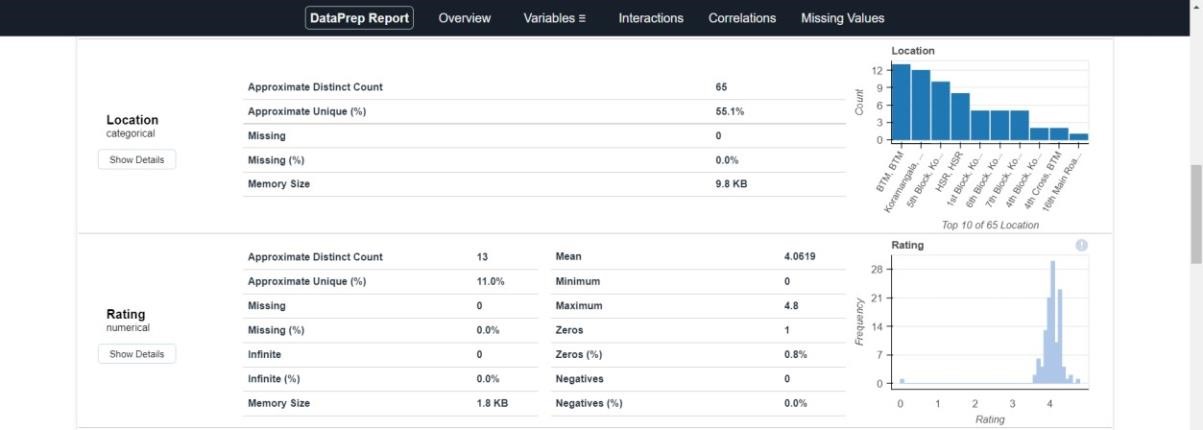
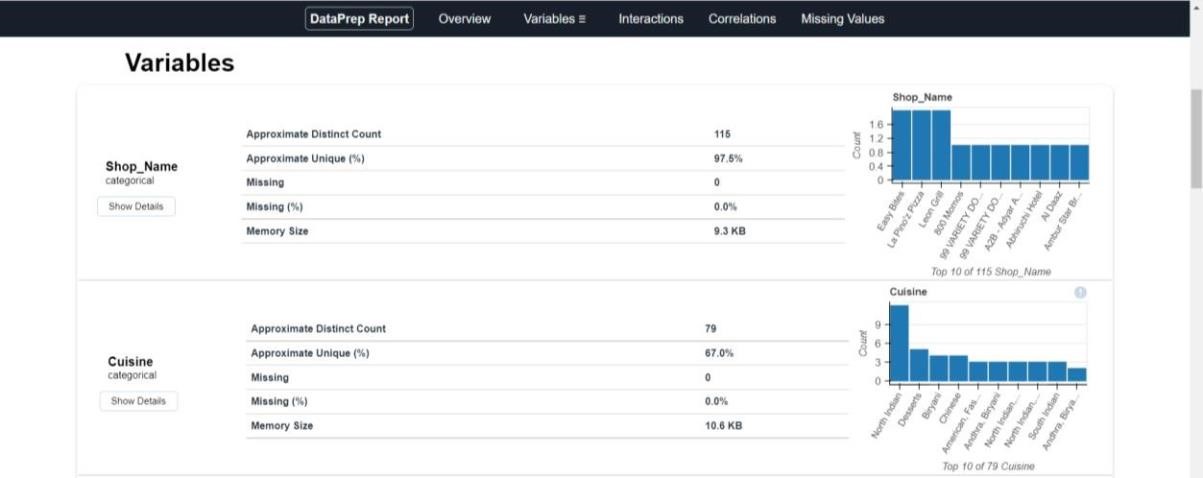


### 2.1.1 Swiggy Bangalore Outlet Dataset Overview –

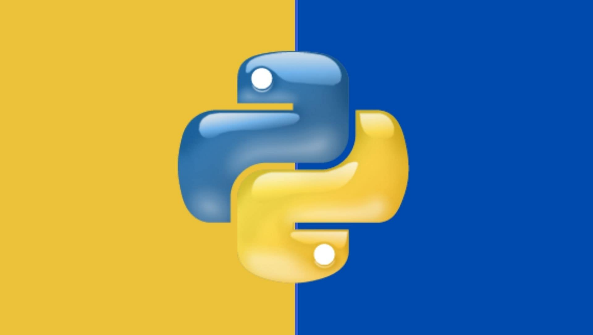
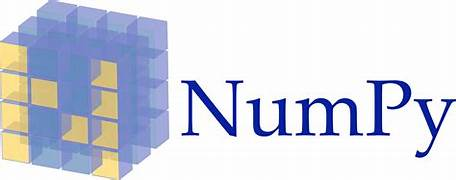
The Listings dataset comprises a tabular structure containing 118 entries and 5 attributes. These attributes are divided into 2 Continuous characteristics and 3 Categorical characteristics. Notably, the dataset does not contain any Missing Values, accounting for 0% of the total records. In summary, the dataset is devoid of any instances with Missing Values.



### 2.1.2 Overview of Variables:



# 3 Architecture



## 3.1 Architecture Description –

### 3.1.1 Data Description –

### In the previous sections, we observed that within our dataset pertaining to the Swiggy Bangalore Outlet, there are approximately 118 entries containing 5 distinct attributes. These attributes are divided into 2 continuous variables and 3 categorical variables. The data is presented in the format of Comma Separated Values (.csv).

### 3.1.2 Define the Use Cases –

### Currently, considering the provided dataset and identified business issues, we have established various scenarios for analysis. This approach will undoubtedly assist in extracting crucial insights from the data, which will serve as a foundation for making business choices. Moreover, it aids not only in comprehending significant connections among attributes but also empowers us to conduct independent investigations and generate our own discoveries.

### 3.1.3 Import the Dataset –

Since the dataset has been provided in the Comma Separated Value (.csv) format, we can utilize the Pandas read\_csv() function to import it.



#### 3.1.4 Extract, Transform, Load (ETL) –

• "Extract Transform Load" (ETL) constitutes a crucial phase known as "Data Exploration" within the Data Analysis Process. During this stage, various methodologies are applied to gain a deeper comprehension of the employed dataset.

• Grasping the nature of the dataset encompasses a range of aspects, which include but are not confined to the following:

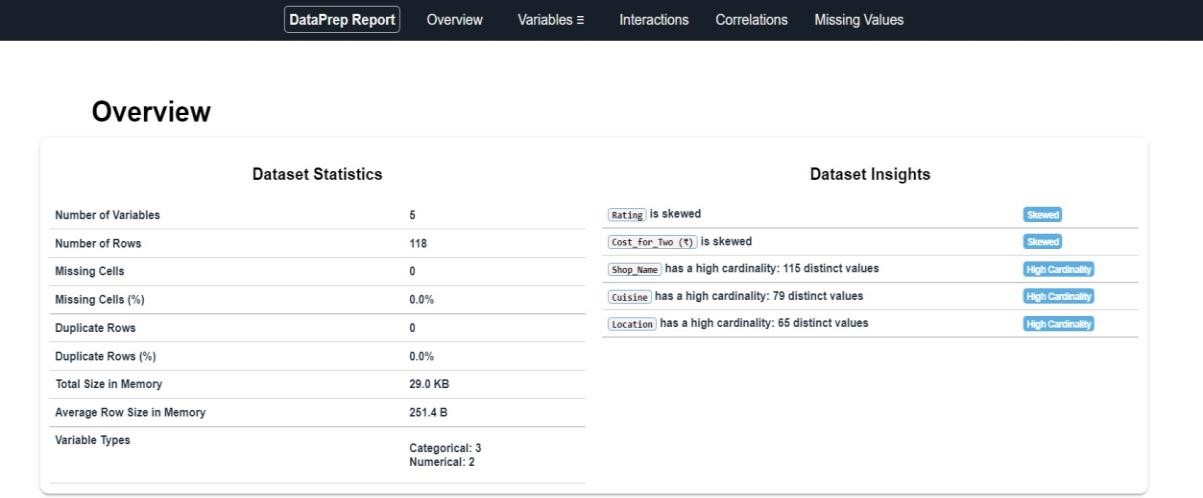
* Extracting significant "Variables."
* Detecting "Outliers," "Missing Values," or instances of "Human Error."
* Uncovering the interplay between different variables.
* Ultimately, enhancing insights into the dataset while minimizing the potential for subsequent errors.

• In simpler terms, this process facilitates a more profound grasp of both the "Variables" and the "Relationships" they share.

• To automate our Exploratory Data Analysis (EDA), we leverage the dataprep module.

• The module provides the ensuing information:

* Overview: Identifying column types within a Data Frame.
* Variables: Detailing variable attributes, unique values, distinct count, and instances of missing values.
* Quartile statistics, encompassing minimum and maximum values, Q1, median, Q3, range, and interquartile range.
* Descriptive statistics, such as mean, mode, standard deviation, sum, median absolute deviation, coefficient of variation, kurtosis, and skewness.
* Correlations: Signaling highly correlated variables, along with Spearman, Pearson, and Kendall matrices.
* Handling Missing Values: Visual representations through Bar Charts, Heatmaps, and a spectrum of missing value occurrences.



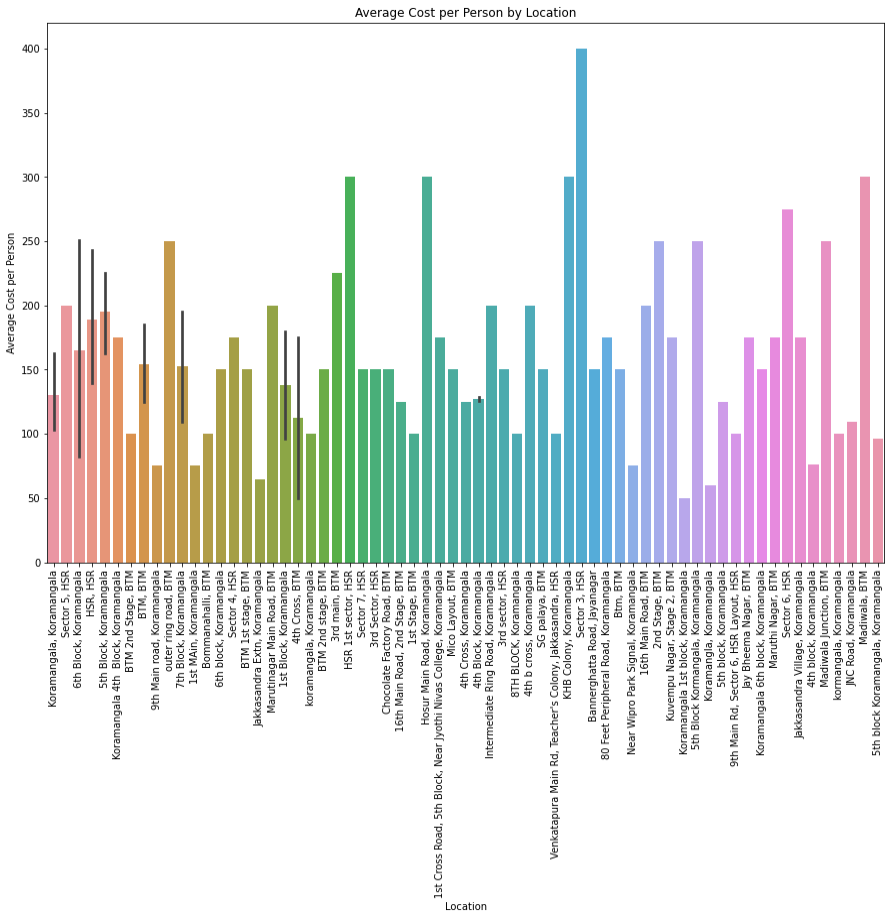
### 3.1.5 Data Pre-processing, Data Cleaning & Imputation (Handling the Categorical & Numerical Variables) –

Data pre-processing involves the preparation of raw data to make it suitable for analysis. This includes extensive data cleaning, addressing missing values through suitable imputation methods, and handling both categorical and numerical variables appropriately. In the context of this project, we performed substitution or imputation for missing values using mean, median, or mode based on the variable's nature. Additionally, we eliminated columns that were not relevant to our analysis.

### 3.1.6 Analyse the Data –

### After completing the pre-processing phase, we are ready to proceed with our real analysis. This involves crafting lines of code and logical instructions to arrange our data according to the specified intended purposes.

### 3.1.7 Visualize & Share Meaningful Insights –

At last, the moment has come to transform our data into a visual depiction of information. In essence, data visualization involves converting extensive data sets and measurements into various visual formats like Bar Plots, Pie Charts, Heat maps, Box Plots, Scatter Plots, and numerous others. These visuals effectively present the data, facilitating the recognition and dissemination of insights derived from the information. Below is a captivating preview of one of our visual creations:

# 4 Technology Stack

|  |  |
| --- | --- |
| **Data Manipulation Library** | Pandas |
| **Visualization Library** | Matplotlib, Seaborn, Plotly, etc |
| **ETL** | dataprep |
| **Dataset** | .CSV Format |
| **IDE** | Jupyter Notebook |