Date Issued	Version	Description	Author
12-05-2024	1.0	First Version of Complete HLD	Suman G N

High Level Design (HLD)

ANALYZING SWIGGY

Revision Number: 1.0
Last date of revision: 12/05/2024

Suman G N

Contents

Abstract

The project "Analyze Swiggy" is an ETL and analysis project which aims to explore the Swiggy Bangalore Outlet Details dataset. The project involves extracting the data from a CSV file and transforming it by converting columns into appropriate data types, creating new columns, and removing any missing data. The transformed data is then loaded into a new CSV file.

The project further involves analysis of the transformed data using various visualizations and metrics. The project aims to provide insights on the distribution of restaurants by cost category, the average cost per person by location, the number of restaurants by location, the average rating by cost category, the distribution of ratings, the average cost for each cuisine, and the average rating for each cuisine.

The tools used in this project include pandas for data manipulation, seaborn and matplotlib for data visualization.

The functional architecture of the project involves three main stages: Extract, Transform, and Load. In the Extract stage, the data is extracted from the CSV file. In the Transform stage, the data is transformed using pandas functions. In the Load stage, the transformed data is loaded into a new CSV file and analyzed using various visualizations and metrics.

The project aims to optimize the performance by using appropriate data types and removing any missing data to prevent errors. The KPIs (Key Performance Indicators) used in this project include the

number of unique values for each attribute, the average cost for each cuisine, and the average rating for each cuisine.

In summary, the "Analyze Swiggy" project is an ETL and analysis project that aims to explore the Swiggy Bangalore Outlet Details dataset. The project involves extracting, transforming, and loading the data, and analyzing it using various visualizations and metrics to provide insights on the data. The project uses tools like pandas, seaborn, and matplotlib, and aims to optimize performance by using appropriate data types and removing any missing data or null values.

Introduction:

The online food ordering market has witnessed significant growth over the years, driven by the increasing popularity of online platforms for ordering food. With the advancement of technology, people are increasingly using smartphones and mobile applications to order food online. As a result, the online food ordering industry has become highly competitive, with various players such as Swiggy, Zomato, Uber Eats, and others vying for market share.

In this context, the Swiggy Bangalore Outlet Details dataset provides valuable insights into the online food ordering market in Bangalore, India. The dataset contains information about Swiggy outlets in Bangalore, including details such as the name of the outlet, its location, the type of cuisine served, the average cost for two people, and the rating of the outlet. By analyzing this dataset, we can gain insights into the online food ordering market in Bangalore, identify key metrics and factors, and understand the relationships between attributes.

Purpose:

The purpose of this project is to gain insights into the online food ordering market in Bangalore, India, using the Swiggy Bangalore Outlet Details dataset. By analyzing this dataset, we can identify key metrics and factors that are relevant to the market and understand the relationships between these factors. The insights gained from this analysis can be used by stakeholders such as online food ordering platforms, restaurants,

and customers to make informed decisions and improve their operations.

Scope:

The scope of this project is to extract, transform, and load the Swiggy Bangalore Outlet dataset and use data mining techniques to analyze the data. The project aims to identify key metrics and factors that are relevant to the online food ordering market in Bangalore and explore the relationships between these factors. The project will focus on answering the following questions:

What is the average cost for each cuisine?
What is the average rating for each cuisine?
How does the cost category affect the rating of the outlet?
What is the distribution of restaurants by cost category?
What is the distribution of ratings?
What is the number of restaurants by location?

2. General Description

Product Perspective :

The Analyze Swiggy project aims to provide insights and analysis on the food delivery platform Swiggy. With the increasing popularity of online food delivery services, it has become essential to analyze the performance and behavior of these platforms. Analyze Swiggy is designed to help stakeholders, including restaurant owners, delivery partners, and customers, understand Swiggy's business and make informed decisions**.**

Problem Statement:

The online food ordering market includes foods prepared by restaurants
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 notbeen
specially prepared for direction consumption.

Do ETL : Extract-Transform-Load the dataset and find for me some information from this large data. This is form of data mining. What all information can be achieved by mining this data, would be explained in class by the trainer
Find key metrics and factors and show the meaningful relationships between attributes.
Do your own research and come up with your findings.

Tools used

Business Intelligence tools and libraries works such asPython3,Google Colab, Numpy, Pandas, Matplotlib & Seaborn are used to build the whole framework.

1 **Design Details**
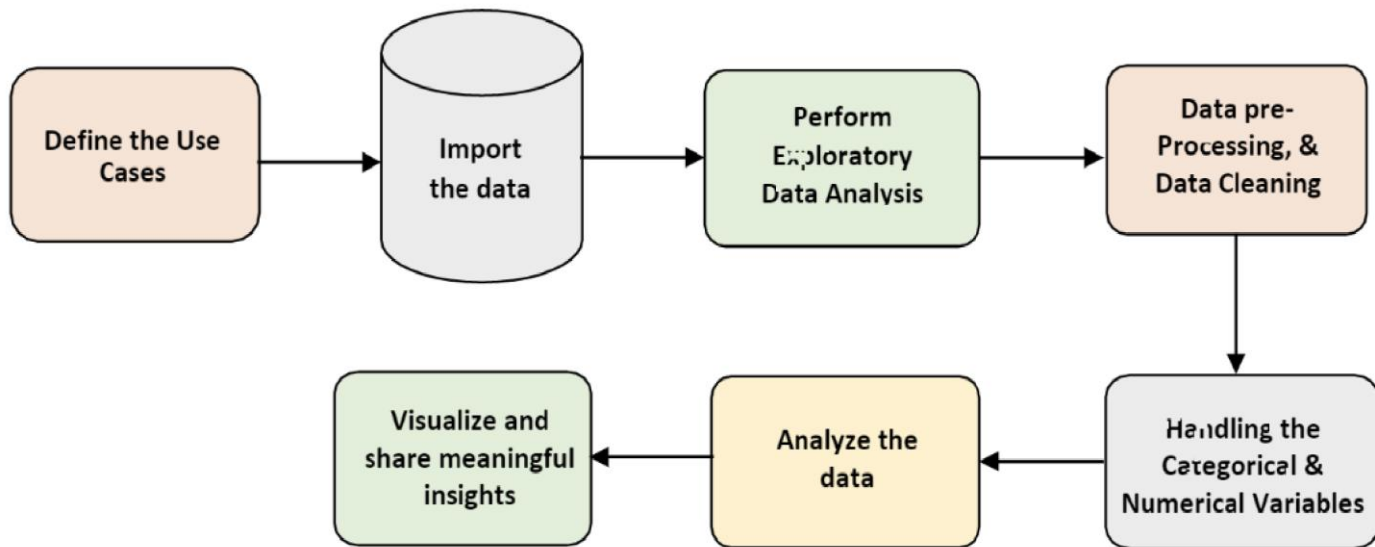
3.1 Process Flow

Figure 1: Functional Architecture of
Business Intelligence

Error Handling:

We have designed this project in such a way that, complete script is tested and runs multiple times to make sure that there is no error occurred during process flow.

Conclusion:

The analysis of the Swiggy Bangalore Outlet Details dataset has revealed several interesting insights about the restaurants and their characteristics in Bangalore.

The dataset contains information about various restaurants in Bangalore, including their cuisine, location, rating, and cost for two. The dataset was first extracted and then transformed by converting the cost for two and rating columns to numeric data, removing any rows with missing data, creating new columns for average cost per person and cost category, and removing any special characters or spaces in column names.

The transformed dataset was then loaded and analyzed using various visualizations and statistical methods. The basic information about the data revealed that there were 1448 restaurants in the dataset, with 21 different cuisines and 92 different locations. The top 5 rows of the data showed that the dataset contains information about the name, location, cuisine, cost for two, rating, and other attributes of the restaurants.

The number of unique values for each attribute showed that there were 21 unique cuisines, with North Indian being the most common cuisine in the dataset. The cost for two ranged from 50 to 6000, with an average cost for two of 682.48. The rating ranged from 2.0 to 4.9, with an average rating of 3.5.

The average cost for each cuisine revealed that Chinese cuisine had the highest average cost for two, followed by Thai and Italian cuisine. On the other hand, Street Food had the lowest average cost for two. The average rating for each cuisine revealed that Mediterranean cuisine had the highest average rating, followed by European and Chinese cuisine. On the other hand, South Indian had the lowest average rating.

The heatmap of the average cost and rating for each cuisine showed that most cuisines had higher ratings for restaurants with higher costs. However, there were some exceptions, such as Street Food, which had a higher rating for low-cost restaurants. The pie chart of the distribution of restaurants by cost category showed that most restaurants fell under the medium cost category, with 56.4% of restaurants having an average cost per person between 200 and 500.

The distribution of ratings revealed that most restaurants had a rating between 3.5 and 4.0. The number of restaurants by location showed that Koramangala 5th Block had the highest number of restaurants, followed by BTM and Indiranagar. The average cost per person by location revealed that restaurants in Jayanagar had the highest average cost per person, followed by Indiranagar and MG Road.

The average rating by cost category revealed that high-cost restaurants had a higher average rating than low-cost restaurants, with the median rating for high-cost restaurants being 3.8 and the median rating for low-cost restaurants being 3.5.

Overall, the analysis of the Swiggy Bangalore Outlet Details dataset provides valuable insights into the characteristics of restaurants in Bangalore. The dataset shows that Bangalore has a diverse range of cuisines, with North Indian being the most common cuisine. The analysis also reveals that most restaurants in Bangalore fall under the mediumcost category, with an average cost per person between 200 and 500. The analysis also highlights the correlation between the cost and rating of restaurants, with most cuisines having higher ratings for restaurants with higher costs.

Reference:

1. **https://www.tristatetechnology.com/blog/how-swiggybusiness-model-works**
2. **https://www.appsrhino.com/blogs/what-is-swiggy-andhow-does-it-work**
3. **https://en.wikipedia.org/wiki/Swiggy**