Cognifyz Data Analysis Internship Report

Introduction:-

This report documents the work completed during the Data Analysis Internship at Cognifyz Technologies. The goal of this internship was to analyze a restaurant dataset and derive meaningful insights using Python and data analysis libraries. The tasks were divided into multiple levels, each focusing on different analytical objectives.

Tools Used:-

- Python for data analysis
- Pandas and NumPy for data manipulation
- Matplotlib for data visualization
- Jupyter Notebook as the development environment

Dataset Overview:-

The dataset contains detailed information about restaurants, including location, cuisine types, price range, ratings, votes, and service availability such as online delivery and table booking. Each record in the dataset represents a single restaurant.

Level 1: Exploratory Data Analysis:-

In Level 1, the dataset was loaded and cleaned by handling missing values and checking for duplicates. Basic exploratory analysis was performed to understand the structure of the data.

Key insights included:

- Identification of the most commonly served cuisines
- City-wise analysis of restaurant count and average ratings
- Distribution of restaurants across different price ranges
- Percentage of restaurants offering online delivery
- This level helped build a strong foundation for further analysis.

Level 2: Advanced Analysis:-

Restaurant Ratings

The distribution of aggregate ratings showed that many restaurants have a rating of 0.0, indicating unrated restaurants. Most rated restaurants fall between 3.0 and 4.0, suggesting generally good customer satisfaction. The average number of votes per restaurant was around 157, indicating moderate customer engagement.

Cuisine Combinations

North Indian cuisine, both alone and combined with Chinese, appeared most frequently in the dataset. However, some less common cuisine combinations showed very high average ratings, suggesting that niche cuisines tend to receive better feedback.

Geographic Analysis

Using latitude and longitude, restaurant locations were plotted on a map. The visualization showed dense clusters of restaurants in specific areas, likely representing major cities and urban regions.

Restaurant Chains

Several restaurant chains were identified based on repeated restaurant names. Well-known chains such as Cafe Coffee Day and Domino's Pizza had multiple outlets. Some chains showed high average ratings and vote counts, indicating strong customer trust and popularity.

Level 3: Additional Analysis:-

In Level 3, further analysis was performed to study:

- Restaurants with the highest and lowest number of votes
- Relationship between ratings and votes
- Comparison of price range with online delivery and table booking availability

The analysis suggested that higher-priced restaurants are more likely to offer additional services such as online delivery and table booking.

Conclusion:-

This internship provided hands-on experience with real-world data analysis. By working through different levels of tasks, meaningful insights were derived related to customer preferences, restaurant performance, pricing, and service availability. Overall, the project enhanced practical data analysis and problem-solving skills.

Intern Details

Name: Divya Lakhtariya
Internship: Data Analysis Intern
Organization: Cognifyz Technologies