COGNIFYZ DATA ANALYSIS INTERNSHIP PROJECT REPORT

Title: Restaurant Data Analysis

Student Name: Mohammed Afreen

Internship Organization: Cognifyz Technologies

Duration: 01/01/2026 – 01/02/2026

Tools & Technologies: Python, Pandas, Matplotlib, Power BI

---

1. Introduction

Data analysis plays a vital role in converting raw data into meaningful insights that support effective decision-making. During my internship at Cognifyz Technologies, I worked on a comprehensive data analysis project involving restaurant data. The project included data cleaning, visualization, and the development of an interactive dashboard using Power BI. The primary objective was to extract actionable insights and present them in a clear and user-friendly manner.

---

2. Objectives of the Project

- To understand and analyze the restaurant dataset
- To clean and preprocess raw data using Python
- To create meaningful visualizations using Matplotlib
- To design an interactive dashboard using Power BI
- To derive insights and support data-driven decision-making

---

3. Dataset Description

The dataset used in this project consists of restaurant-related information collected from Cognifyz Technologies.

Key attributes include:

- Restaurant ID, Restaurant Name

- Country Code, City, Address, Locality, Locality Verbose
- Longitude, Latitude
- Cuisines
- Average Cost for Two, Currency
- Has Table Booking, Has Online Delivery, Is Delivering Now
- Switch to Order Menu
- Price Range
- Aggregate Rating, Rating Color, Rating Text
- Votes

The dataset was provided in CSV format and contains **9,551 rows** and **21 columns**.

---

4. Methodology

4.1 Data Collection

The dataset ("Dataset.csv") was provided by Cognifyz Technologies as part of the Data Analysis Internship Program.

4.2 Data Cleaning and Preprocessing

Using Python and Pandas, the following preprocessing steps were carried out:

- Handling missing and null values
- Removing duplicate records
- Renaming columns for better clarity
- Converting data types where necessary

4.3 Data Visualization Using Matplotlib

Matplotlib was used to generate visual insights through:

- Bar charts
- Line graphs
- Pie charts

- Histograms

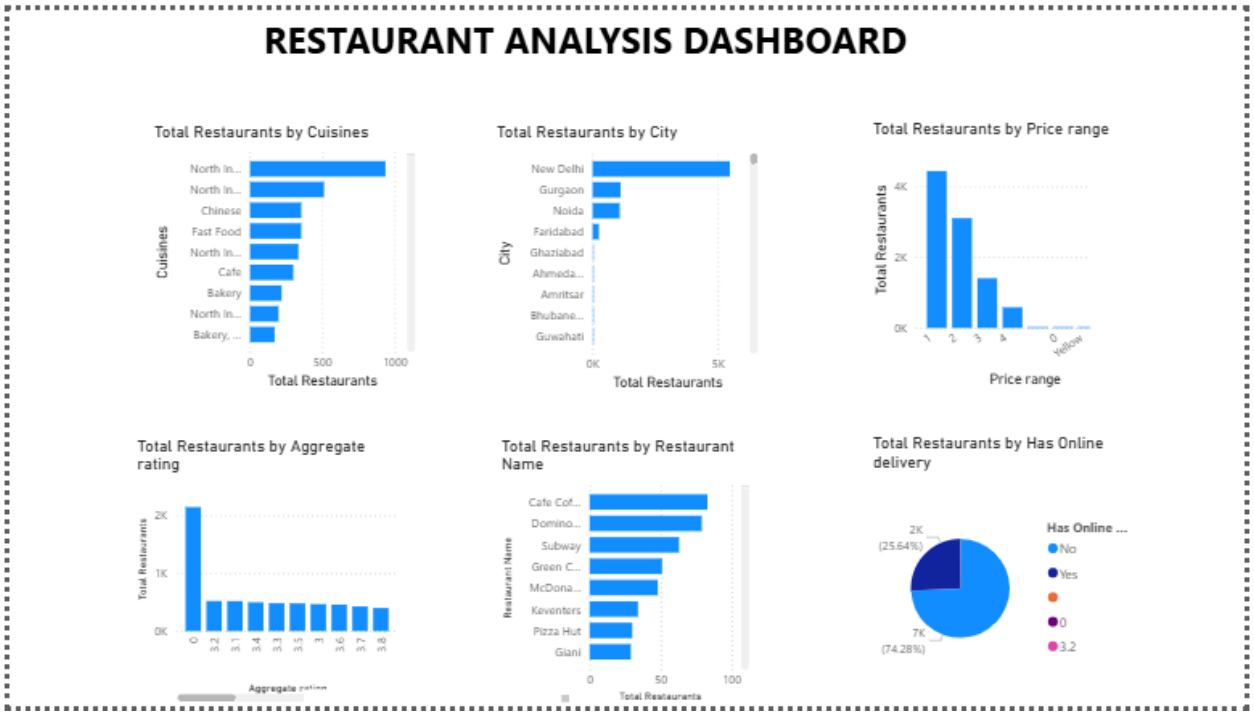
These visualizations helped identify patterns, trends, and comparisons within the dataset.

4.4 Dashboard Creation Using Power BI

After preprocessing, the dataset was imported into Power BI. An interactive dashboard was designed featuring:

- Graphical representations of key insights
- Visual comparison across different categories
- Analysis of trends
- Interactive filtering for deeper insights

4.5 Dashboard Overview



The dashboard enables stakeholders to easily explore restaurant performance and customer preferences.

---

5. Task-Based Analysis

Level 1 Analysis

Task 1: Top Cuisines

- Identified the top three cuisines
- Calculated their percentage contribution

Task 2: City Analysis

- Identified the city with the highest number of restaurants
- Calculated the average rating for each city

Task 3: Price Range Distribution

- Analyzed the distribution of restaurants across different price ranges

Task 4: Online Delivery

- Compared ratings of restaurants with and without online delivery

Level 2 Analysis

Task 1: Restaurant Ratings

- Analyzed rating distribution
- Identified the most common rating range
- Calculated the average number of votes

Task 2: Cuisine Combination

- Identified common cuisine combinations
- Compared ratings across combinations

Task 3: Geographic Analysis

- Plotted restaurant locations using latitude and longitude
- Identified geographic clusters

Task 4: Restaurant Chains

- Identified restaurant chains
- Analyzed ratings and popularity of different chains

---

6. Results and Insights

The key insights obtained from the analysis are as follows:

- A small number of cuisines dominate the restaurant market
- A higher number of restaurants in a city does not always imply higher ratings
- Most restaurants fall within the mid-price range
- Restaurants offering online delivery tend to have slightly higher ratings
- The majority of restaurants are rated between 3.5 and 4.5
- Certain cuisine combinations receive better customer ratings
- Restaurants are concentrated in urban and commercial areas
- Popular restaurant chains maintain consistent ratings across locations

---

7. Challenges Faced

- Handling inconsistent and missing data
- Understanding complex relationships among variables
- Designing a simple yet informative dashboard layout

These challenges were addressed through systematic preprocessing and iterative dashboard refinement.

---

8. Conclusion

This project successfully demonstrated the application of Python and Power BI in real-world data analysis. By transforming raw restaurant data into meaningful insights, the project emphasizes the value of data-driven decision-making. The internship significantly enhanced my analytical, technical, and visualization skills.

---

9. Learning Outcomes

- Practical experience with Python, Pandas, and Matplotlib
- Hands-on exposure to Power BI dashboard development
- Enhanced analytical and problem-solving abilities

---

10. Future Scope

- Automating data refresh in Power BI
- Incorporating advanced and interactive visualizations
- Applying machine learning techniques for predictive analysis

---

11. References

- Python Official Documentation
- Power BI Dashboard
- Github link: https://github.com/Afreen-1517/COGNIFYZ_DATA-ANALYST-PROJECT

Student Name: Mohammed Afreen

Btech-CSE-lV year

Date: 20/01/2026