**Project Documentation**  
**Internship Domain**: Data Analysis  
**Company**: Cognifyz Technologies  
**Intern Name**: Priya Patil  
**Dataset Used**: Restaurant Dataset (Dataset.csv)

**About Cognifyz Technologies**

Cognifyz Technologies is a growing tech company that provides cutting-edge solutions in artificial intelligence, machine learning, and data analytics. They conduct internship programs designed to equip students and graduates with practical, industry-relevant skills.

**Internship Objective**

The objective of this internship was to perform an in-depth analysis of a restaurant dataset, apply real-world data science skills, and extract meaningful insights. The internship was divided into 3 levels covering various aspects of data analysis from basic exploration to advanced visualization and correlation analysis.

**Development Tools Used**

* Python Programming Language
* Libraries: Pandas, Matplotlib, Seaborn, Folium
* IDE: PyCharm Community Edition
* Jupyter Notebook (for quick visualization)
* CSV Dataset: Restaurant.csv

**Level 1: Exploratory Data Analysis (EDA)**

**Task 1: Top Cuisines**

* Identified the top 3 most common cuisines.
* Calculated the percentage of restaurants offering each of them.

**Task 2: City Analysis**

* Identified the city with the highest number of restaurants.
* Calculated average rating by city.
* Determined the city with the highest average rating.

**Task 3: Price Range Distribution**

* Plotted price range distribution using bar charts.
* Calculated the percentage of restaurants in each range.

**Task 4: Online Delivery Analysis**

* Found the percentage of restaurants offering online delivery.
* Compared average ratings of restaurants with and without online delivery.

**Level 2: Data Grouping & Geographic Visualization**

**Task 1: Restaurant Ratings**

* Analyzed distribution of ratings.
* Found most frequent rating.
* Calculated average number of votes per restaurant.

**Task 2: Cuisine Combination**

* Identified frequent cuisine combinations.
* Compared average rating for each combination.

**Task 3: Geographic Analysis**

* Used latitude and longitude to map restaurants.
* Used Folium and Marker Clustering to visualize clusters.

**Task 4: Restaurant Chains**

* Detected restaurants with the same name across multiple locations.
* Compared average rating and vote count.

**Level 3: Text & Correlation Analysis**

**Task 1: Restaurant Reviews**

* Tokenized reviews.
* Identified most common positive and negative words.
* Analyzed relationship between review length and rating.

**Task 2: Votes Analysis**

* Identified restaurants with highest and lowest votes.
* Explored correlation between votes and ratings.

**Task 3: Price Range vs Online Delivery and Table Booking**

* Analyzed service availability based on price range.
* Found that higher-priced restaurants more frequently offered table booking.

**Key Learnings**

* Real-world data handling and analysis
* Data visualization best practices
* Geospatial mapping with Python
* Relationship analysis between different restaurant attributes
* Development using PyCharm and Jupyter Notebook

**Note**: This documentation is based on the analysis of a restaurant dataset and showcases practical application of data analytics concepts in a real internship environment.