Capstone Project Report

Project Title: Identifying Star Restaurants & Customer Preferences Using Data Analysis

Situation

During my internship as a Data Analyst, I worked on a project for a restaurant aggregator aiming to improve its B2C platform. The existing recommendation approach relied on limited metrics and did not accurately reflect customer behavior. The business required a more data-driven framework to identify high-performing ("star") restaurants and improve customer recommendations.

Task

My responsibilities included:

- Cleaning and preparing raw restaurant data for analysis
- Performing exploratory data analysis (EDA) to understand customer preferences and restaurant performance
- Defining measurable KPIs to identify star restaurants
- Communicating insights clearly to business stakeholders through visualizations

Action

I followed a structured data analysis lifecycle:

1. Data Cleaning & Preparation

- Imported two datasets (restaurant data and country code mapping) and merged them using a common key
- Inspected data structure, identified missing values, and removed exact duplicates to ensure data quality
- Created derived features such as cuisine count to enable deeper analysis

2. Exploratory Data Analysis (EDA)

Analyzed geographic distribution to identify cities with the highest and lowest restaurant concentration

Evaluated operational features such as online delivery and table booking

Compared ratings and votes across key factors including:

- Number of cuisines
- Average cost for two
- Delivery availability

Used visualizations (histograms and box plots) to validate trends and support insights

3. Metric Definition & Insight Generation

Identified key drivers of restaurant performance:

Online delivery strongly correlated with higher customer engagement and votes

Restaurants with moderate pricing and higher cuisine variety performed better

Defined star restaurants using explainable, business-aligned criteria:

- High aggregate rating
- Strong customer engagement (votes)
- Availability of online delivery

4. Data Visualization & Communication

- Exported the cleaned dataset and built an interactive Tableau dashboard
- Included KPIs, filters (city, cuisine, cost range), and comparative visualizations
- Designed the dashboard to enable non-technical stakeholders to independently explore insights

Result

Delivered a clean, analysis-ready dataset and a structured framework to evaluate restaurant performance. Built an interactive dashboard that allowed stakeholders to identify star restaurants using data-backed logic. Improved visibility into how pricing, cuisines, and delivery options influence ratings and customer engagement

Impact

Demonstrated end-to-end data analysis skills: data cleaning → EDA → metric definition → visualization. Enabled a shift from intuition-based decisions to data-driven recommendations. Created reusable KPIs that can support future automation or recommendation models. Strengthened stakeholder confidence by translating complex analysis into clear, actionable insights

This project strengthened my ability to transform raw data into meaningful insights, define clear performance metrics, and communicate results effectively through dashboards—skills directly aligned with a Data Analyst role.