



INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON
Eazydiner Analysis using EDA and
Webscraping

Presented By :- Rohan Santosh Chauthe

About me

- **Name :-** Rohan Santosh Chauthe .
- **Education :-** G. H. Raisoni college of engineering and business management ,Jalgaon.
- **Why I Want To Learn Data Science :-**
 - I want to learn data science because ,I've always been curious about how companies makes the decision about what customers likes or dislikes ,that urge of knowing how this decision making works in companies created my interest in data science .
- **Connect With Me :-**
LinkedIn - www.linkedin.com/in/rohanchauthe
GitHub - <https://github.com/Rohanchauthe>

Business Objectives

- Analyse how restaurants are distributed across **cities, locations, cuisines, and meal types**.
- Identify **price ranges**, affordability, and whether **premium locations charge higher prices**.
- Understand how ratings vary with **price, cuisine, meal type, and location**.
- Examine how **offers/discount percentages** vary by **price range, restaurant type, and rating**.
- Detect restaurants that offer **high ratings at relatively lower prices**, helping customers make better choices.

Web-Scraping : Details

- Data was collected from the **EazyDiner website**.
- Restaurants information was extracted using **web scraping**.
- The dataset represents **real-time publicly available restaurant listings**.
- Scrapped attributes include:
 - Restaurant Name
 - City & Location
 - Cuisine
 - Meal Type
 - Price
 - Ratings
 - Offers

Data Summary

Database Schema

- **Data Source:** EazyDiner website (web scraped)
- **Data Format:** Structured tabular data (CSV/DataFrame)

Total Records	5692
Total Features	9(Columns)
Cities Covered	10
Unique Cuisines	500
Price Range	₹150 - ₹8000

Data Quality & Readiness

- Missing values and inconsistent formats due to web scraping
- Duplicates identified and removed
- Data cleaned and standardised for EDA

Data Cleaning

Missing Value Handling

- Identified missing values in price, rating, and offers
- Applied appropriate removal or imputation techniques.
- City has missing values because some restaurants don't have a city in data

Missing Values

```
hotel          0
address        0
ratings        3
prices          0
cuisines       21
offers          0
Meal_Type      0
city          1590
dtype: int64
```

After handling missing values

```
Restaurants      0
address          0
city             0
ratings          0
prices          0
cuisines         0
Meal_Type        0
Restaurant_offer_percentage 0
Online_payment_offer_percentage 0
dtype: int64
```

Data Type Standardisation

- Converted offers and prices from text to numeric format
- Removed currency symbols and special characters

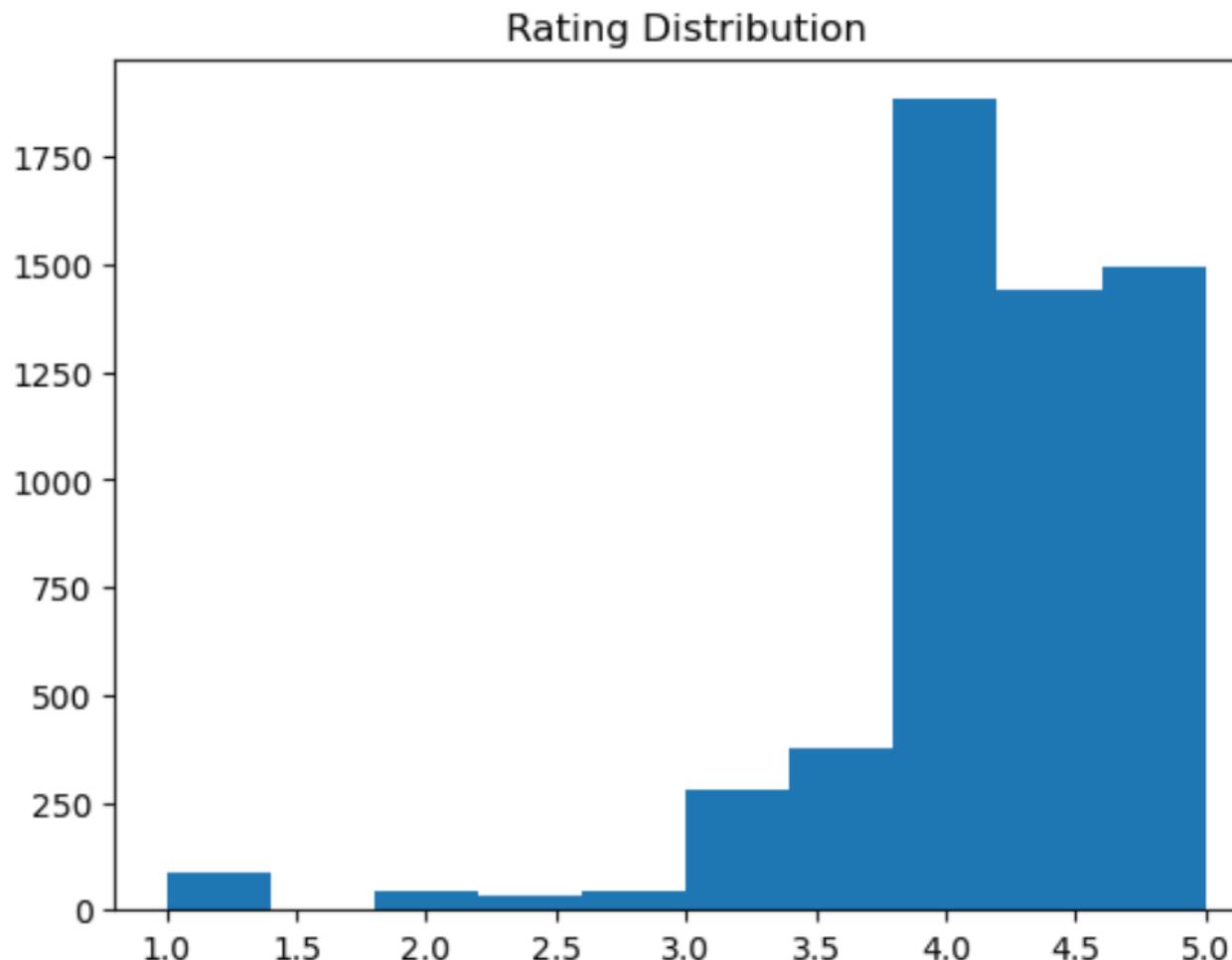
```
hotel          object
address        object
ratings       float64
prices          object
cuisines        object
offers          object
Meal_Type      object
dtype: object
```

```
Restaurants      object
address          object
city             object
ratings       float64
prices          int64
cuisines         object
Meal_Type        object
Restaurant_offer_percentage  int64
Online_payment_offer_percentage  int64
dtype: object
```

After Converting data types

Before Converting data types

Ratings Distribution

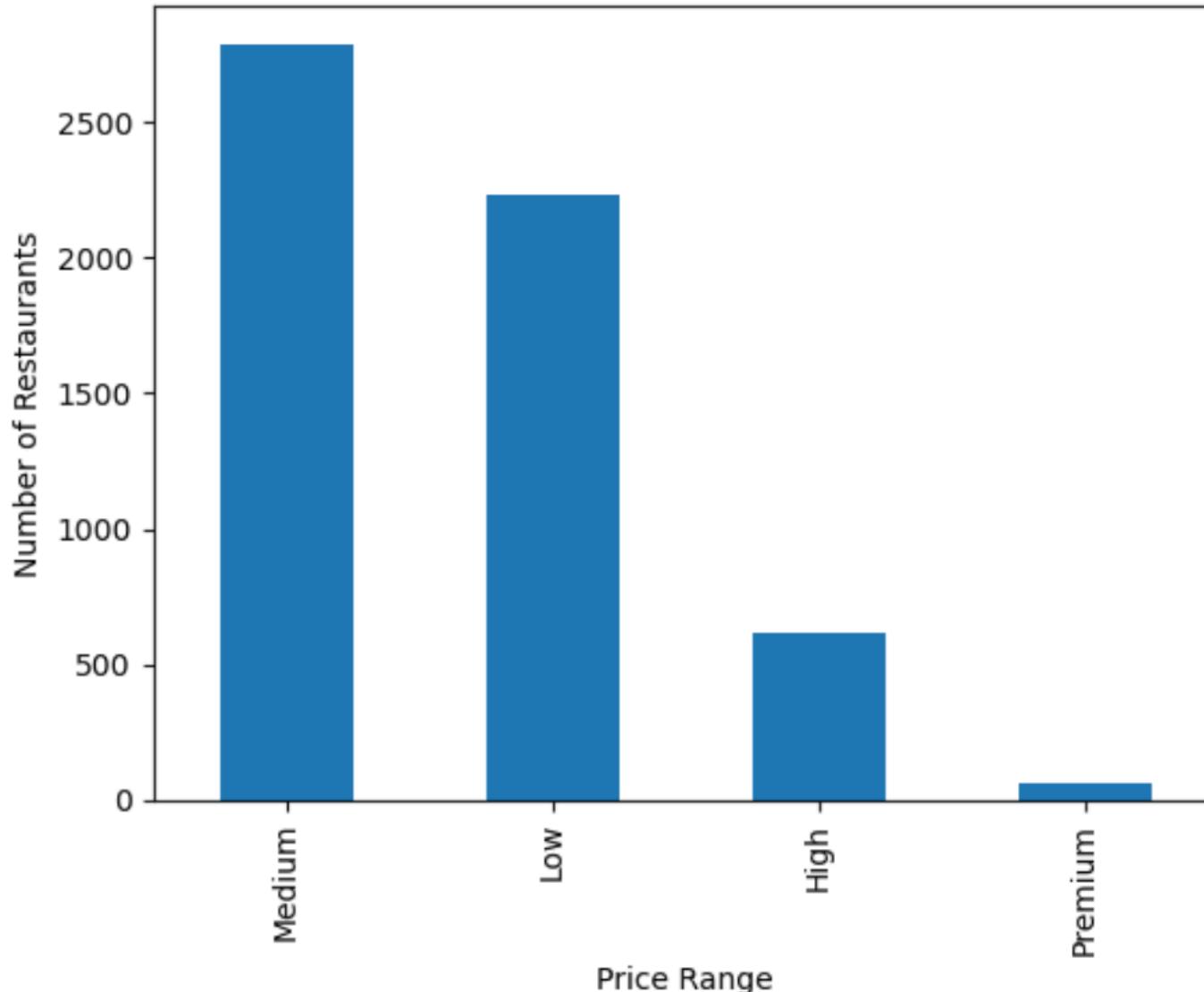


Insights :-

- The highest bars are in the **4.0–4.5** range
- This shows **most restaurants are well-rated**
- Indicates generally **positive customer satisfaction.**

Distribution of Restaurants Across Price Ranges

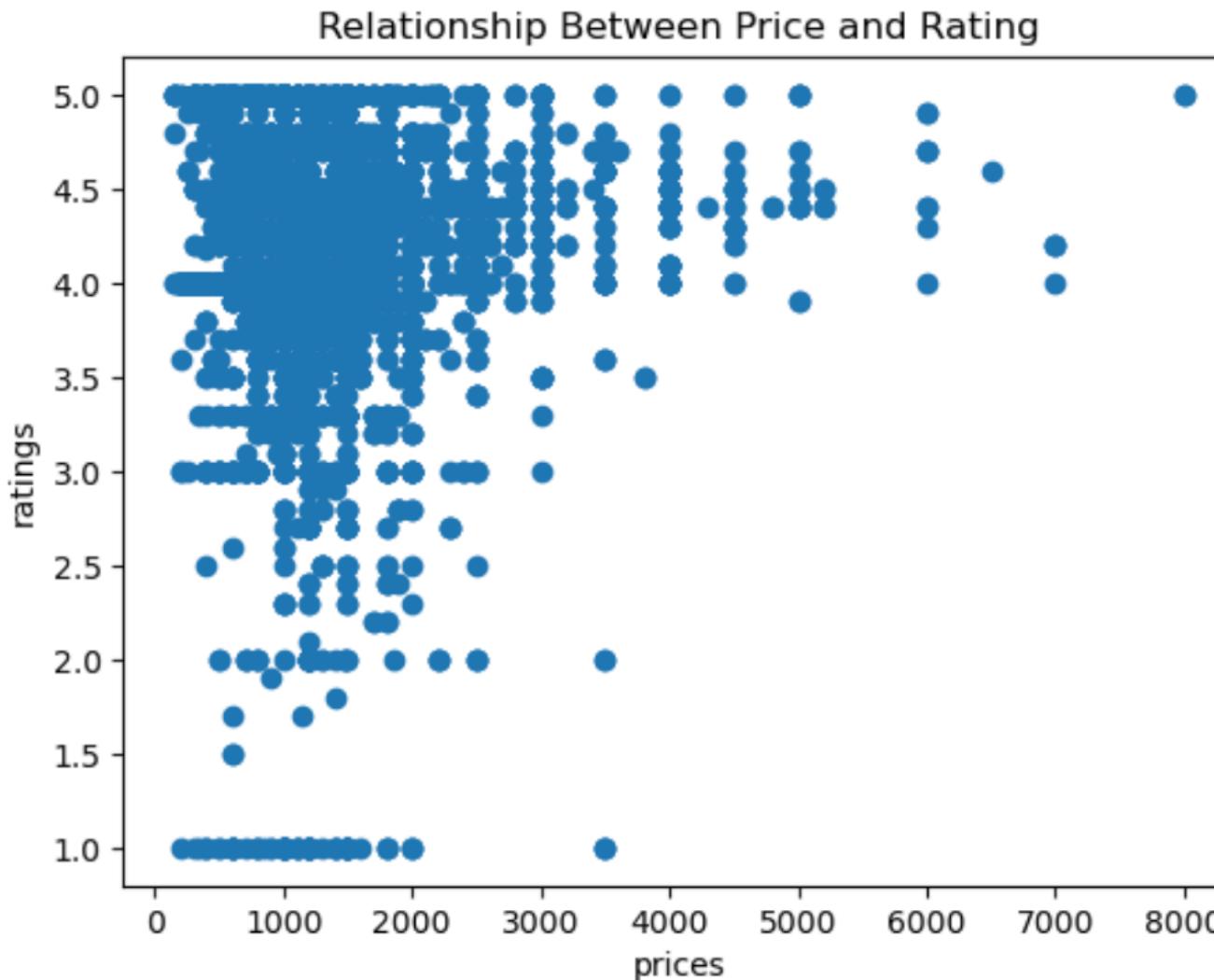
Distribution of Restaurants Across Price Ranges



Insights :-

- The platform is largely dominated by affordable to mid-range dining options, targeting mass-market customers.
- Fewer restaurants operate in the high-price segment, possibly due to limited demand or higher operational costs.

Relationship Between price and Rating



Insights :-

- Restaurants across **all price ranges** receive **both high and low ratings**
- High ratings (4 - 5) are not exclusive to expensive restaurants.
- Affordable restaurants dominate the dataset and show **high variability in customer satisfaction**.

Relationship Between Restaurant Price and Offer Percentage



Insights :-

- Discount strategies are mainly used to attract customers in budget and mid-range restaurants.
- Premium restaurants rely more on brand value and experience rather than heavy discounts.

Challenges

- EazyDiner uses dynamically loaded content, making scraping more complex and time-consuming
- Columns like ratings and offers contained both text and numeric values, complicating preprocessing.
- Some restaurants had missing, ratings, city names, requiring extensive cleaning.

Conclusions and Recommendations

Conclusions :-

- Market is dominated by mid-range restaurants.
- High ratings are common, but mostly clustered Ratings are heavily concentrated between 3.5 and 4.5.
- Discounts are mainly used by budget and mid-range restaurants.
- Customers prioritise quality and experience over discounts
- Offers are more of a marketing tool than a quality indicator

Recommendations :-

- Restaurants should **focus on food quality and service**, as higher prices do not guarantee higher ratings.
- Mid-range restaurants should use **moderate discounts strategically** to attract price-sensitive customers.
- Premium restaurants should **avoid heavy discounting** and instead highlight experience, ambience, and exclusivity.
- Budget restaurants can leverage offers as a customer acquisition tool in competitive markets.

THANK YOU

