INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

eazydiner

Title: Web Scraping and Exploratory Data Analysis of Eazy Dinner Restaurant Finder

INNOMATICS
RESEARCH LABS

About Us:

- We K.Chandu Goud and S.Rishikesh reddy B.Tech graduate passionate about Data Analytics and Visualization.
- Exploring data-driven insights through real-world datasets.
- Interested in web data extraction and business trend analysis.
- **LINKEDIN(Chandu)**: https://www.linkedin.com/in/k-chandu-goud-32482a29a/
- **GITHUB:** https://github.com/chandu63012
- **LINKEDIN(Rishikesh)**: https://www.linkedin.com/in/rishikesh-reddy-3013a5233
- **GITHUB:** https://github.com/rishikeshreddy1405/Workco.git

INNOMATICS RESEARCH LABS

Introduction:

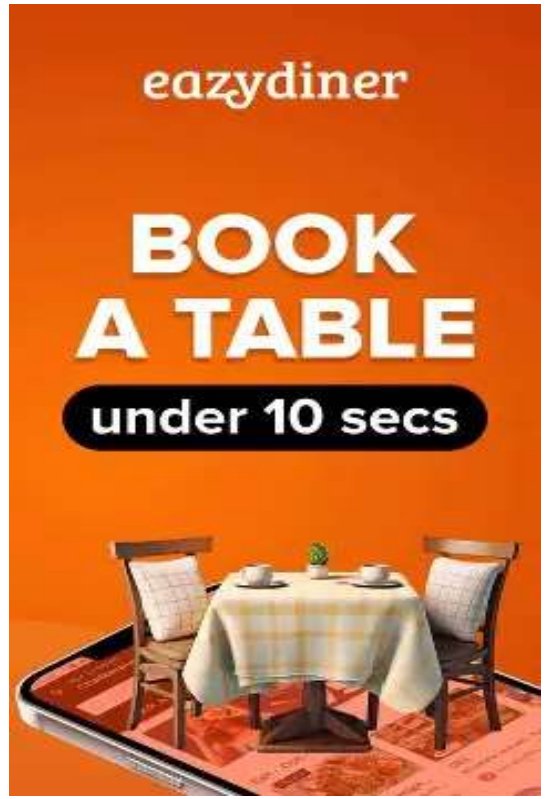
- Eazy Dinner is an online restaurant discovery platform that provides details such **as Restaurant names**, **Cuisines**, **Ratings,Price,Offers** and **Locations**.
- This project focuses on **web scraping data** from the **EazyDiner website (Hyderabad)** using **Python (BeautifulSoup, Requests, Pandas)** to collect, clean, and analyze restaurant information.
- The main goal is to extract structured data to identify **popular cuisines, top-rated restaurants, and dining trends** across Hyderabad.
- This end-to-end project highlights the use of **data automation and analysis** for real-world insights in the food industry.
- Additionally, it demonstrates how **web scraping simplifies data collection** from large dynamic websites.
- The project serves as a foundation for **future visualization and trend analysis using Power BI or Tableau**.

INNOMATICS
RESEARCH LABS

Business Problem:

- Restaurant information on EazyDiner is unstructured and spread across multiple web pages, making manual data collection time-consuming.
- No direct access to a consolidated dataset for analyzing restaurant ratings, cuisines, or location-based trends.
- Difficulty in comparing restaurants across areas due to inconsistent or incomplete online data formats.
- Lack of automated tools to extract and update restaurant insights regularly for analytics or visualization purposes.

eazydiner



INNOMATICS
RESEARCH LABS

Objectives:

- Extract structured restaurant data (Name, Cuisine, Rating,Offers,Prices and location) from the EazyDiner website using Python-based web scraping.
- Clean and preprocess the collected data to ensure consistency, accuracy, and usability for analysis.
- Analyze restaurant trends to identify popular cuisines, highly rated areas, and dining patterns in Hyderabad.
- Build a foundation for future visualization and interactive dashboards using tools like Power BI or Tableau.
- Automate data extraction and updates to enable continuous monitoring of restaurant performance and emerging food trends.

Web Scraping:

- EazyDiner’s Official website was selected as the main data source for restaurants information in Hyderabad.
- Used browser developer tools(inspect Element) to identify relevant HTML elements containing restaurant names,cuisine,ratings,price,offers and location
- Used BeautifulSoup and Requests libraries in python to extract structured restaurants data from the web pages.
- Sent automated HTTP requests to retrieve multiple restaurant listings across different pages.
- Cleaned and consolidated the scraped data into a structured dataset for further analysis and visualization.

```
url="https://www.eazydiner.com/restaurants?city=hyderabad&location=hyderabad"
```

```
response=requests.get(url)
```

```
response
```

```
<Response [200]>
```

```
from bs4 import BeautifulSoup
import requests
import numpy as np
import time

Hotel_name = []
Price = []
Rating = []
Location = []
Timings= []
Offers= []
Resterunt_types= []
pagenum = []

total_time = time.time()

for i in range(1, 98):
    start_time = time.time()
    URL = f"https://www.eazydiner.com/restaurants?location=hyderabad"
    page = requests.get(URL)
    soup = BeautifulSoup(page.text, "html.parser")

    for x in soup.find_all('div', attrs={'class':'relative pointer listing_touch_box__00o7f'}):
        Hname = x.find('a', attrs={'class':'ellipsis listing_res_name__uVIN8'})
        price = x.find('div', attrs={'class':'flex align-v-center'})   # discounted price
        rating = x.find('svg',attrs={'xmlns':'http://www.w3.org/2000/svg'})
        location = x.find('div', attrs={'class':'font-12 ellipsis flex listing_res_address__LuVOm'})
        timings= x.find('div', attrs={'class':'flex margin-r-10'})
        offers = x.find('div',attrs={'class':'flex align-v-center ellipsis'})
        res_type= x.find('div', attrs={'class': 'flex align-v-center ellipsis listing_multi_cuisines__zLLXK'})
```

```
Hotel_df.to_csv("hotel.csv",index=False)
```

INNOMATICS RESEARCH LABS

Tools Used:

Data Cleaning Steps:

- **Removed Unwanted Characters & Formatting:** Cleaned restaurant names, cuisines, and location fields by removing special symbols, extra spaces, and inconsistent text formats to ensure uniformity.
- **Handled Missing & Incomplete Values:** Replaced null or missing entries in rating and cuisine columns with suitable defaults or “Not Available” to maintain data integrity and avoid errors during analysis.
- **Data Type Conversion & Standardization:** Converted rating columns from string/object type to numeric (float) for accurate statistical analysis and visualization. Ensured consistent text casing for categorical fields like cuisine and city.
- **Verified Location & Structured Dataset:** Cross-checked location and address information for accuracy, formatted city names properly, and stored the cleaned output into a structured DataFrame and CSV for further analysis.

```
df['Price'] = df['Price'].replace(r'[^0-9.]', '', regex=True)
```

```
df['Offers'] = df['Offers'].str.extract(r'(\d+%)')
```

df

	Hotel_name	Price	Rating	Location	Offers	Restaurant_types
0	63 Degree Modern Regional Buffet	2400.	4.1	Gachibowli, Hyderabad	25%	Regional Indian
1	63 Degree Modern Regional Buffet	2400.	4.1	Gachibowli, Hyderabad	25%	Regional Indian
2	Golconda Pavilion	2500.	4.4	ITC Kohenur, Hyderabad	10%	North Indian, Multicuisine, Continental, Desse...
3	Golconda Pavilion	2500.	4.4	ITC Kohenur, Hyderabad	10%	North Indian, Multicuisine, Continental, Desse...
4	Exotica	1500.	4.3	Banjara Hills, Hyderabad | View All(2) Outlets	15%	Multicuisine
...	...	...	...	...	...	...
1741	one8 Commune	2500.	4.6	Hitech City, Hyderabad	25%	Multicuisine
1742	Pizza Zone	1000.	4.0	Habsiguda, Hyderabad | View All(2) Outlets	25%	Italian
1743	Pizza Zone	1000.	4.0	Habsiguda, Hyderabad | View All(2) Outlets	25%	Italian
1744	Fuel Cafe	800.	4.0	Banjara Hills, Hyderabad	15%	Cafe
1745	Fuel Cafe	800.	4.0	Banjara Hills, Hyderabad	15%	Cafe

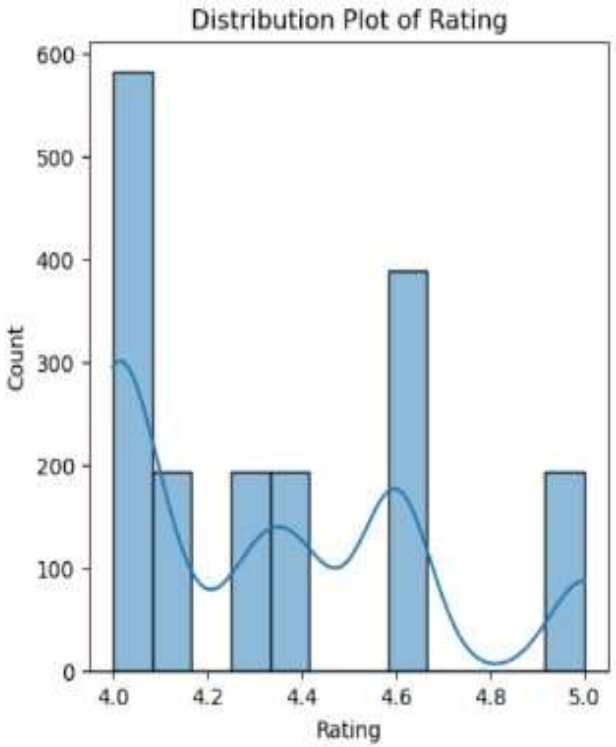
1746 rows × 6 columns

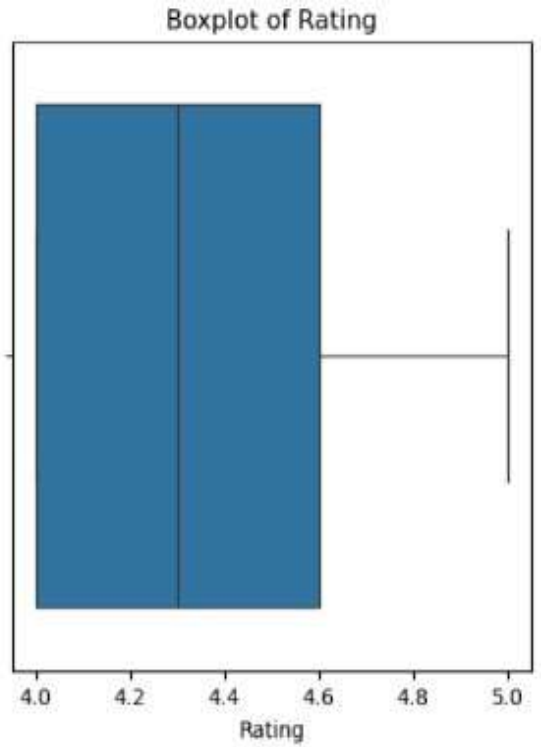
INNOMATICS RESEARCH LABS

Data Visualization:

Ui-Variate Analysis

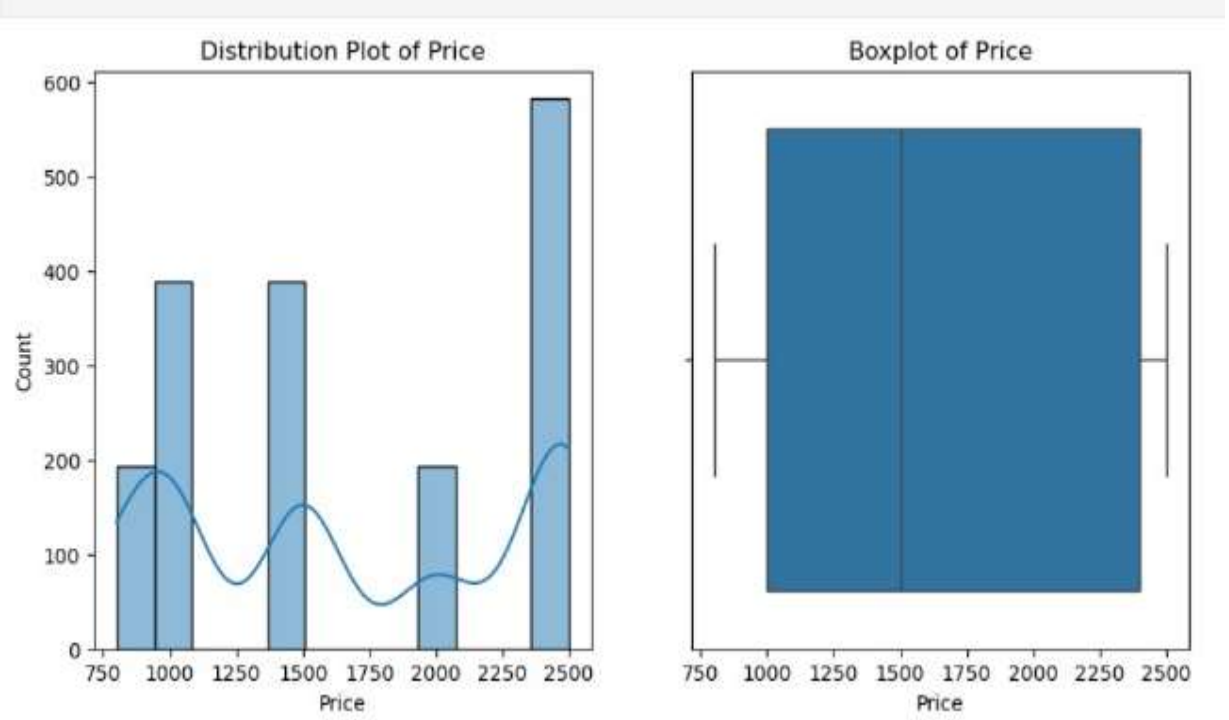
- The distribution plot shows most restaurants have ratings between 4.0 and 4.7, indicating generally high customer satisfaction on the platform.
- The KDE curve shows multiple small peaks, suggesting slight variation in rating popularity, but still concentrated around the higher range.
- The boxplot confirms a high median rating, with a small interquartile range, meaning most restaurants consistently receive good ratings.
- There are few outliers near the upper end (~5.0), indicating only a limited number of exceptionally high-rated restaurants.





Data Visualization:

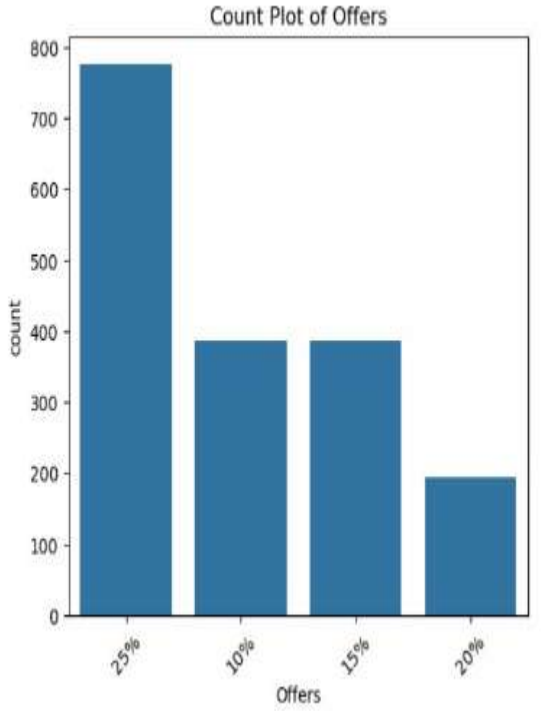
- The distribution plot shows that restaurant prices mostly fall between ₹1,000 and ₹2,500, indicating a mix of mid-range to premium dining options.
- Multiple peaks in the KDE curve suggest varied pricing clusters, showing presence of both budget and premium restaurants.
- The boxplot confirms a wide price range, with the median price leaned toward the higher side, meaning many restaurants are moderately expensive.
- The whiskers show few extreme values, indicating limited outliers and that most prices lie within a consistent and expected range.

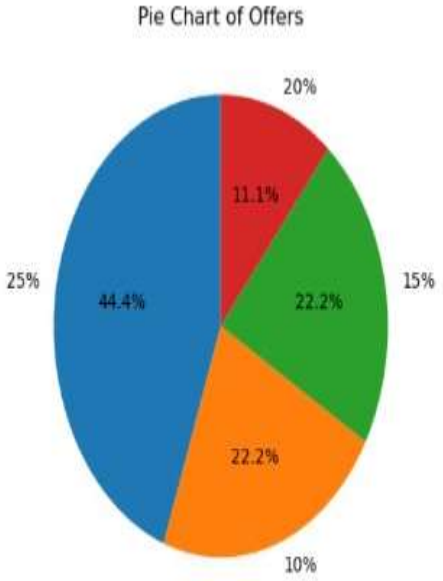


INNOMATICS RESEARCH LABS

Data Visualization:

- The count plot shows that 25% offers are the most common, followed by 10% and 15%, while 20% offers are the least frequent.rating trend on EazyDiner.
- The pie chart confirms this trend visually, with 25% offers making up the largest share and 20% offers contributing the smallest portion.
- The distribution suggests that restaurants tend to provide higher discount rates (25%) more frequently to attract customers.
- The variety of offer percentages indicates a competitive pricing and discount strategy among restaurants to drive customer traffic.





INNOMATICS RESEARCH LABS

Data Visualization:

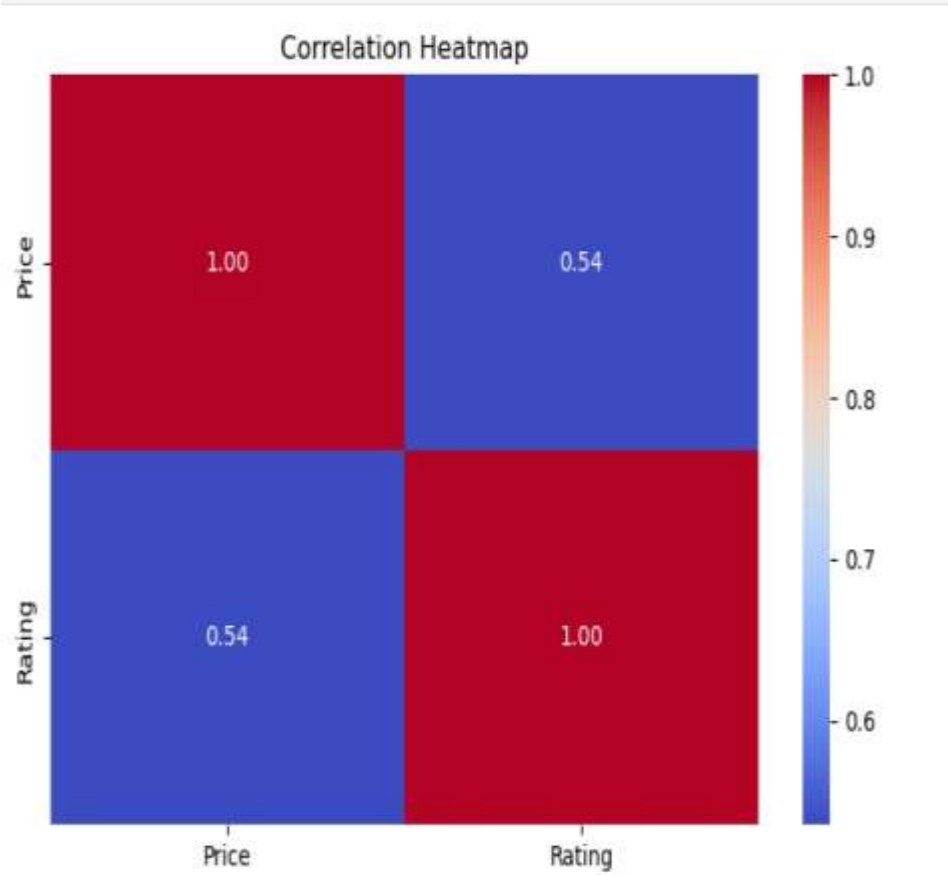
Bi-Variate Analysis

- These visualizations compare restaurant price ranges across different offer percentages (25%, 10%, 15%, and 20%).
- The boxplot highlights the median, inter-quartile range, and outliers, showing price variation within each offer category.
- The violin plot provides a deeper view of data distribution and density, showing how frequently certain price points appear within each offer level.
- Restaurants with 25% and 10% offers have higher price ranges, indicating premium restaurants tend to offer attractive discounts.
- The 15% and 20% offer groups show lower and narrower price ranges, suggesting more budget-friendly restaurants offer smaller, consistent discounts.

INNOMATICS RESEARCH LABS

Data Visualization:

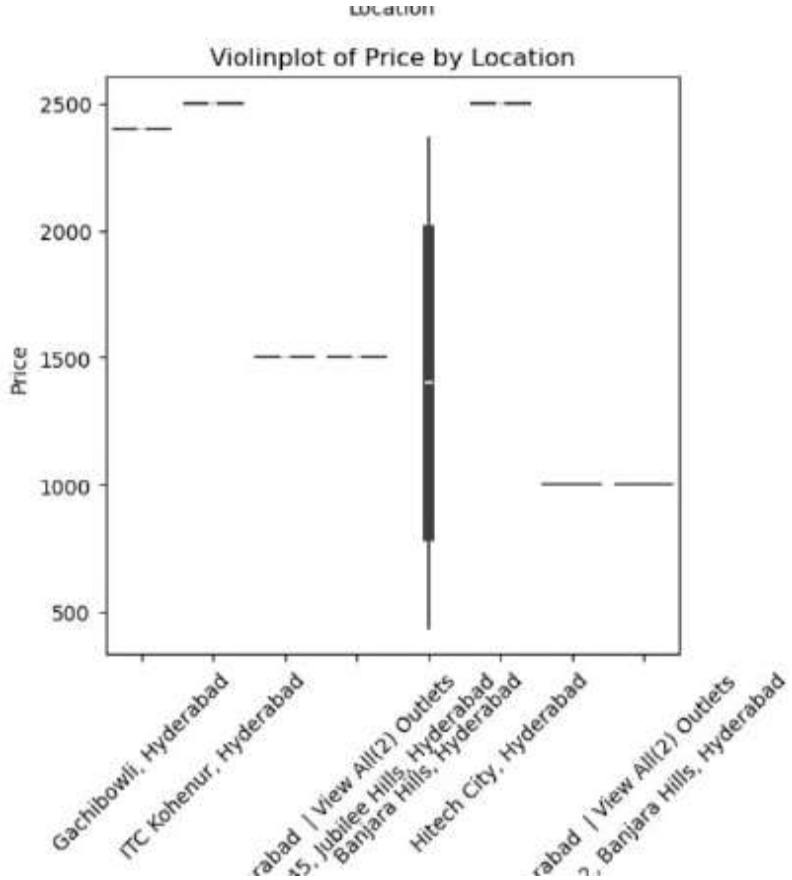
- This heatmap shows the correlation between restaurant price level and customer ratings..
- The diagonal values are 1.00, representing perfect self-correlation of each variable.
- The correlation value of 0.54 between price and rating indicates a moderate positive relationship.
- This suggests that higher-priced restaurants tend to receive better ratings, although the relationship is not very strong.
- The heatmap uses color intensity to clearly show the strength of correlation — red for strong positive correlation and blue for weaker relationships.



INNOMATICS RESEARCH LABS

Data Visualization:

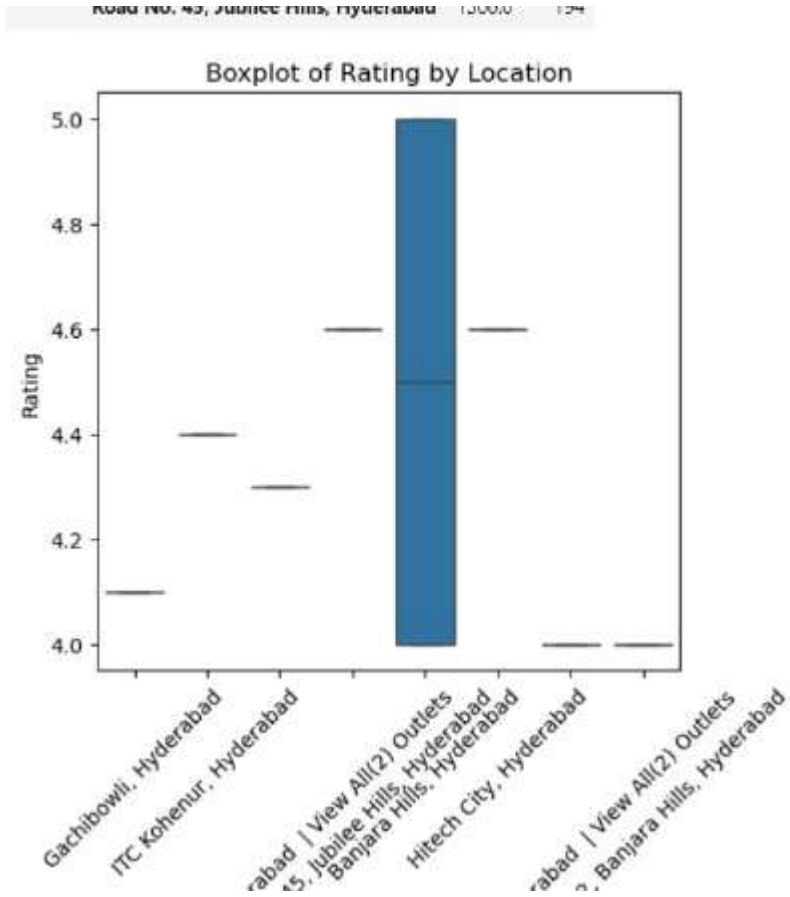
- This violin plot compares restaurant price ranges across different locations in Hyderabad.
- Locations like Gachibowli, ITC Kohenur, and Jubilee Hills show higher price ranges, indicating premium dining zones.
- Areas such as Hitech City and Banjara Hills show a broader spread, meaning a mix of budget and high-end restaurants.
- The distribution reveals that restaurant pricing varies significantly by locality, showing how location influences dining cost and target customer group.



INNOMATICS RESEARCH LABS

Data Visualization:

- This boxplot compares restaurant ratings across different locations in Hyderabad.
- Locations such as Banjara Hills and Jubilee Hills show higher median ratings, indicating strong customer satisfaction in premium areas.
- Some locations like Gachibowli and Hitech City display a wider spread in ratings, meaning a mix of both average-rated and high-rated restaurants.
- The presence of consistent higher ratings in top areas suggests these locations offer better service quality and dining experience.
- Lower variation in ratings for certain areas reflects consistent performance, while higher variation indicates diverse restaurant quality levels within that locality.

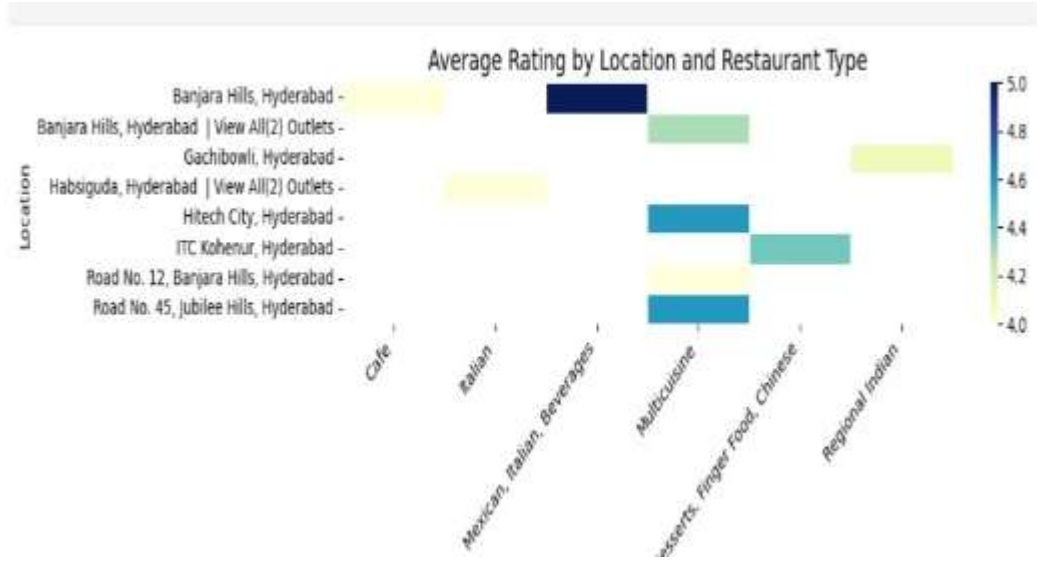


INNOMATICS RESEARCH LABS

Data Visualization:

Multi-Variate Analysis

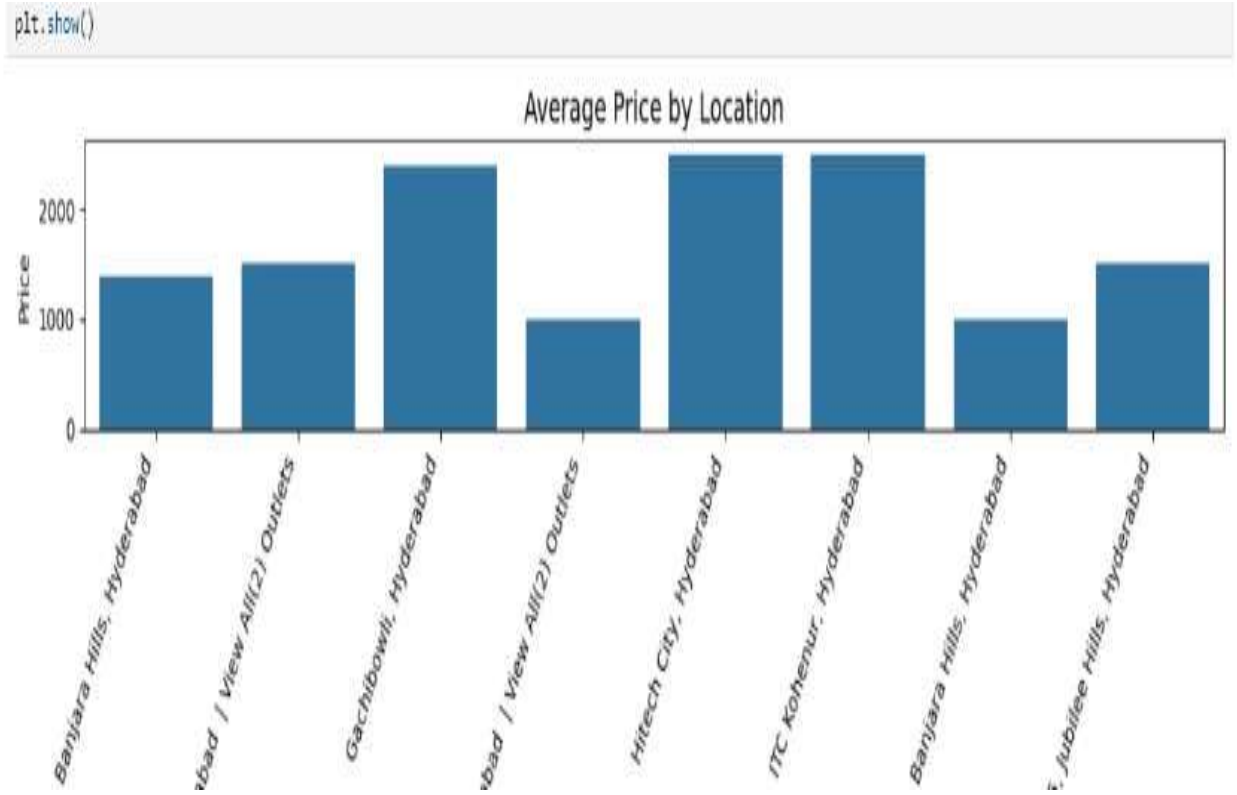
- The boxplot shows that restaurant pricing varies significantly by restaurant type — with Regional Indian, Chinese, and Desserts/Finger Food restaurants having higher price ranges, while Cafes and Italian outlets tend to be more budget-friendly.
- Multicuisine restaurants show a wider spread in price, indicating the presence of both mid-range and premium dining options in this category.
- The heatmap highlights that premium areas like Banjara Hills, Jubilee Hills, and Gachibowli consistently achieve higher average ratings, especially for Multicuisine and International-style restaurants.
- The combined view suggests that location and restaurant type together influence pricing and ratings, with upscale areas showing higher-priced and better-rated restaurants.



INNOMATICS RESEARCH LABS

Data Visualization:

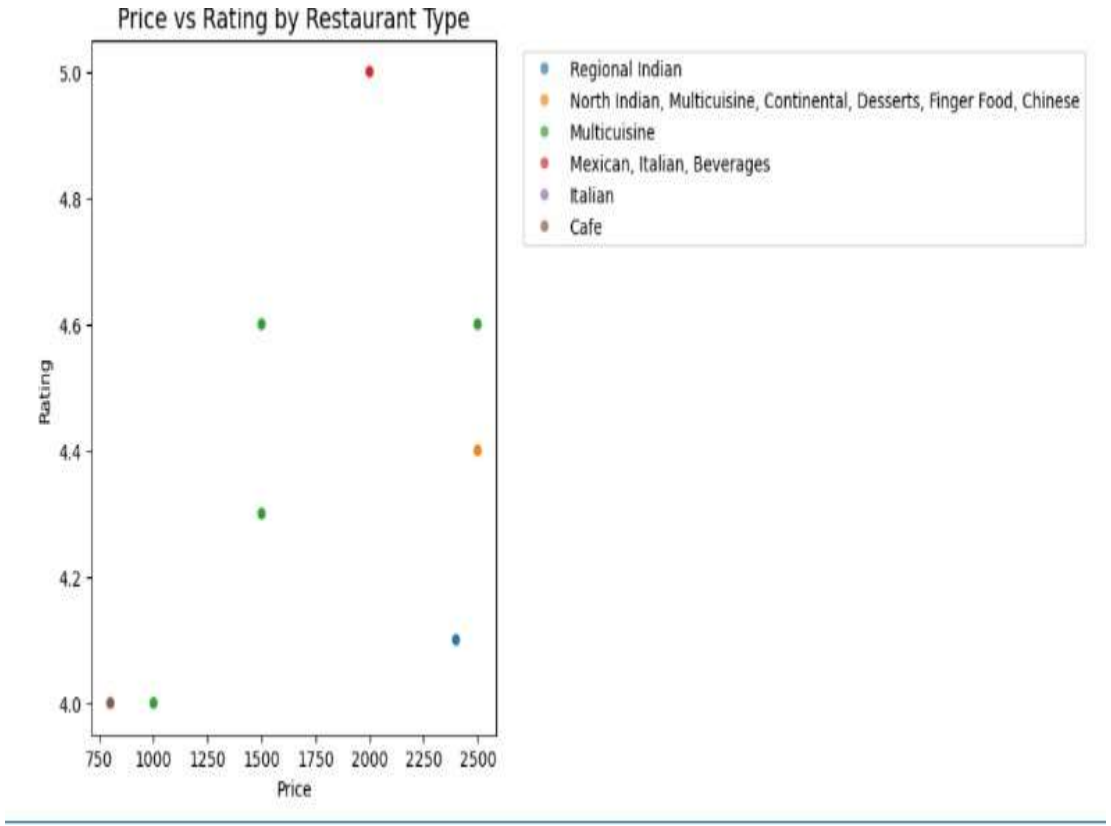
- The chart shows how average restaurant pricing varies across key locations in Hyderabad, highlighting differences in dining affordability.
- Premium areas like Gachibowli, Hitech City, and ITC Kohenur have the highest average prices, indicating a strong presence of upscale restaurants.
- Locations like Banjara Hills and Jubilee Hills show moderately high prices, reflecting a mix of luxury and mid-range dining options.
- Areas with lower average prices suggest more budget-friendly restaurants, catering to a wider audience and casual dining crowd.



INNOMATICS RESEARCH LABS

Data Visualization:

- This scatter plot compares restaurant pricing with customer ratings across different cuisine/restaurant types.
- Higher-priced categories like Multicuisine and International (Mexican/Italian/Beverages) tend to show higher ratings, indicating premium dining often aligns with better customer satisfaction.
- Some budget-friendly segments like Cafes and Regional Indian outlets have lower price points with moderate ratings, showing affordability but mixed experiences.
- The spread reflects that while higher price generally correlates with better ratings, quality varies by cuisine type, and customer preferences differ across restaurant segments.



Key Business Questions

- Which areas in Hyderabad have the most restaurants?
- What are the most popular cuisines?
- Which locations have the highest-rated restaurants?
- How are restaurant ratings distributed across the city?
- Do certain cuisines receive higher ratings consistently?

Conclusion:

- Successfully scraped restaurant data from the EazyDiner platform using Python & BeautifulSoup.
- Performed data cleaning and formatting to convert raw HTML data into a structured dataset.
- Conducted analysis on top locations, cuisines, and restaurant ratings across Hyderabad.
- Observed that **Banjara Hills, Jubilee Hills, and Gachibowli** have the highest restaurant concentration and top ratings.
- Identified **North Indian, Chinese, and Multicuisine** as the most preferred cuisines among users.
- This project demonstrates a complete data workflow: **Web Scraping → Data Cleaning → Exploratory Analysis → Visualization** and highlights how web data can support real-world business insights.

Q&A

Experience– Web Scraping & Data Analysis:

- Scraped restaurant data from EazyDiner.com using Python, BeautifulSoup, and Requests.
- Extracted key information including restaurant names, cuisines, locations, and ratings from multiple pages.
- Cleaned and standardized data by removing noisy characters, handling missing values, and formatting columns for consistency.
- Performed exploratory data analysis to identify top restaurant zones, popular cuisines, and rating patterns in Hyderabad.
- Completed end-to-end workflow: Scraping → Cleaning → Analysis → Visualization → Insights Presentation.

Challenges – Web Scraping & Data Analysis:

- Dynamic and inconsistent HTML structure across restaurant pages
- Extracting complete details for each restaurant (name, cuisine, location, rating)
- Handling missing, duplicated, and unclean text data
- Managing pagination and avoiding request restrictions during scraping
- Selecting clear and meaningful visualizations to highlight insights

INNOMATICS
RESEARCH LABS