***How to explain this project to a stake holder start to end***

Explaining a data analysis project to stakeholders involves outlining the project's objectives, methodology, findings, and implications in a clear and concise manner. Here's a structured approach to presenting this project from start to end:

**1. Introduction**

**Project Objective:** The goal of this project is to revamp the restaurant consolidator's B-to-C portal using intelligent automation technology. Specifically, we aim to identify and recommend restaurants based on various data-driven insights.

**Key Questions:**

* What is the geographical distribution of the restaurants?
* How are the ratings distributed?
* Which franchise has the most national presence?
* What is the ratio of restaurants allowing table booking vs. those that do not?
* What percentage of restaurants provide online delivery?
* Is there a difference in the number of votes for restaurants that deliver vs. those that don't?
* What are the top 10 cuisines served across cities?
* What is the maximum and minimum number of cuisines a restaurant serves?
* What is the relationship between the number of cuisines served and ratings?
* How do cost and other variables affect ratings?

**2. Data Collection and Preparation**

**Data Sources:** We collected data on various restaurants, including details such as location, cuisines offered, average cost for two people, table booking options, online delivery options, ratings, votes, and franchises.

**Data Cleaning:**

* Removed duplicates and handled missing values.
* Standardized data formats for consistency.

**3. Exploratory Data Analysis (EDA)**

**Geographical Distribution:**

* Analyzed the distribution of restaurants across different cities.
* Identified cities with the maximum and minimum number of restaurants.

**Visualization:**

python

Copy code

import matplotlib.pyplot as plt

import seaborn as sns

# Example plot

city\_counts = data['City'].value\_counts()

plt.figure(figsize=(15, 8))

sns.barplot(x=city\_counts.index, y=city\_counts.values, palette='viridis')

plt.xticks(rotation=90)

plt.title('Number of Restaurants by City')

plt.xlabel('City')

plt.ylabel('Number of Restaurants')

plt.show()

**Findings:**

* Cities with the highest concentration of restaurants.
* Cities with fewer restaurant options.

**Ratings Distribution:**

* Analyzed the overall distribution of ratings.

**Visualization:**

python

Copy code

plt.figure(figsize=(10, 6))

sns.histplot(data['Aggregate rating'], bins=20, kde=True)

plt.title('Distribution of Restaurant Ratings')

plt.xlabel('Aggregate Rating')

plt.ylabel('Frequency')

plt.show()

**Franchise Analysis:**

* Identified the franchise with the most national presence.

**4. Specific Insights**

**Table Booking and Online Delivery:**

* Calculated the ratio of restaurants that allow table bookings vs. those that don't.
* Calculated the percentage of restaurants providing online delivery.

**Visualization:**

python

Copy code

# Table Booking Ratio

table\_booking\_counts = data['Has Table booking'].value\_counts()

plt.figure(figsize=(6, 6))

plt.pie(table\_booking\_counts, labels=table\_booking\_counts.index, autopct='%1.1f%%', colors=['skyblue', 'lightgreen'])

plt.title('Table Booking Availability')

plt.show()

# Online Delivery Percentage

online\_delivery\_counts = data['Has Online delivery'].value\_counts()

plt.figure(figsize=(6, 6))

plt.pie(online\_delivery\_counts, labels=online\_delivery\_counts.index, autopct='%1.1f%%', colors=['lightcoral', 'lightskyblue'])

plt.title('Online Delivery Availability')

plt.show()

**Votes Analysis:**

* Compared the number of votes for restaurants that deliver vs. those that don't.

**Visualization:**

python

Copy code

plt.figure(figsize=(10, 5))

data.boxplot(column='Votes', by='Has Online delivery')

plt.title('Votes for Delivery vs Non-Delivery Restaurants')

plt.suptitle('')

plt.xlabel('Online Delivery')

plt.ylabel('Number of Votes')

plt.show()

**Cuisine Analysis:**

* Identified the top 10 cuisines served across cities.
* Calculated the maximum and minimum number of cuisines a restaurant serves.

**Visualization:**

python

Copy code

# Top 10 Cuisines

top\_cuisines = data['Cuisines'].str.split(',').explode().value\_counts().head(10)

plt.figure(figsize=(12, 6))

sns.barplot(x=top\_cuisines.values, y=top\_cuisines.index, palette='magma')

plt.title('Top 10 Cuisines Served')

plt.xlabel('Number of Restaurants')

plt.ylabel('Cuisine')

plt.show()

**Cost and Other Variables:**

* Explored the relationship between cost, number of cuisines, delivery option, and ratings.

**Visualization:**

python

Copy code

# Cost vs Ratings

plt.figure(figsize=(10, 6))

sns.scatterplot(x='Average Cost for two', y='Aggregate rating', data=data, hue='Has Online delivery', palette='coolwarm')

plt.title('Cost vs Ratings')

plt.xlabel('Average Cost for Two')

plt.ylabel('Aggregate Rating')

plt.show()

# Number of Cuisines vs Ratings

data['Num\_Cuisines'] = data['Cuisines'].str.split(',').apply(len)

plt.figure(figsize=(10, 6))

sns.scatterplot(x='Num\_Cuisines', y='Aggregate rating', data=data)

plt.title('Number of Cuisines vs Ratings')

plt.xlabel('Number of Cuisines')

plt.ylabel('Aggregate Rating')

plt.show()

**5. Conclusions and Recommendations**

**Key Insights:**

* Geographical distribution highlights key markets and potential areas for expansion.
* Rating distribution helps understand overall customer satisfaction.
* Popular franchises can be benchmarked for best practices.
* Table booking and online delivery significantly impact customer engagement and satisfaction.
* Cuisine diversity influences restaurant ratings.

**Recommendations:**

* Focus marketing efforts on cities with fewer restaurants to increase brand presence.
* Encourage restaurants to offer online delivery and table booking to enhance customer convenience and satisfaction.
* Promote popular cuisines and explore introducing diverse cuisine options.

**Next Steps:**

* Implement intelligent automation for personalized restaurant recommendations based on user preferences and behavior.
* Regularly update and monitor data to refine recommendations and improve the user experience.

**6. Q&A Session**

Encourage stakeholders to ask questions and provide feedback. This helps ensure they understand the analysis and can provide valuable insights for further refinement.

**Visual Aids**

* Use charts and graphs to support your explanations.
* Highlight key points and insights in your visual aids to make them easily understandable.