



Data Collection and Preprocessing Phase

| Date | 27 January 2025 |
|---------------|----------------------------------|
| Team ID | SWUID20240011509 |
| Project Title | Restaurant Recommendation System |
| Maximum Marks | 6 Marks |

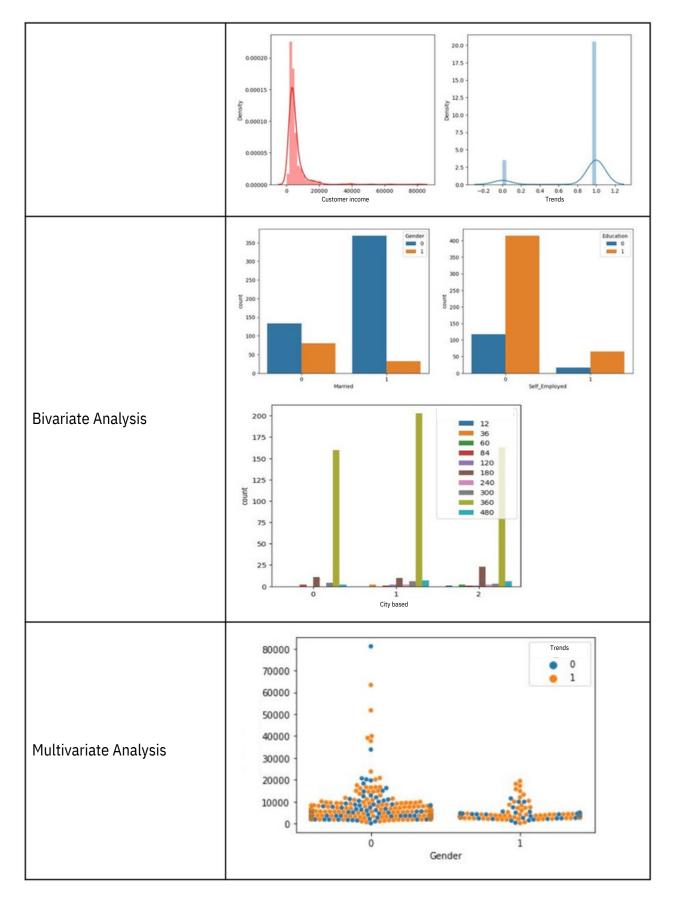
Data Exploration and Preprocessing Report

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

| Section | Description |
|---------------------|--|
| Data Overview | Dimension: 614 rows × 13 columns Descriptive statistics: |
| | Sink Stde Antagery address specificates properly filters under the phone of the contraction of the contracti |
| | Nov. |
| | OSCI-VING OSCI |
| | OSR-1995 |
| | COUSTAL SIGNA SIGNAN SIGNA SIGNAN SIGNA SIGNAN SIGNA SIGNAN SIGNA SIGNAN |
| Univariate Analysis | |











Outliers and Anomalies Data Preprocessing Code Screenshots **Loading Data** def get_latitude_longitude(restaurant_name): # Find the row where the restaurant name matches row = restaurant_data.loc[restaurant_data['title'] == restaurant_name] # Extract latitude and longitude from the row latitude = row['latitude'].values[0]
longitude = row['longitude'].values[0] return latitude, longitude except IndexError: Handling Missing Data # Handle the case where the restaurant name is not found in the DataFrame print(f"Restaurant '{restaurant_name}' not found in the dataset.") return None, None def get_recommendations_with_latlon(restaurant_name, cosine_sim=cosine_sim): # Get the index of the restaurant that matches the input name idx = indices[restaurant_name] sim_scores = list(enumerate(cosine_sim[idx])) def extract_link(json_str): if pd.isna(json_str):
 return "" **Data Transformation** data = json.loads(json_str) if data and isinstance(data, list):
 link = data[0].get("link", "") return link except (json.JSONDecodeError, IndexError, AttributeError): Feature Engineering Attached the codes in final submission.