**WEBSITE TRAFFIC ANALYSIS**

**PROJECT OBJECTIVES:**

Analyze user navigation patterns, including pages visited, time spent on each page, and common paths taken through the website.Purpose To understand how users interact with the website, identify popular content, and pinpoint any bottlenecks in the user journey. To gain insights into the stages of the customer journey, optimize content and interactions at each stage, and improve overall user satisfaction and loyalty.Compare website traffic metrics with competitors to identify strengths, weaknesses, opportunities, and threats.To benchmark the website's performance against industry competitors and identify areas where the website can gain a competitive edge.

**CODING:**

import pandas as pd

import matplotlib.pyplot as plt

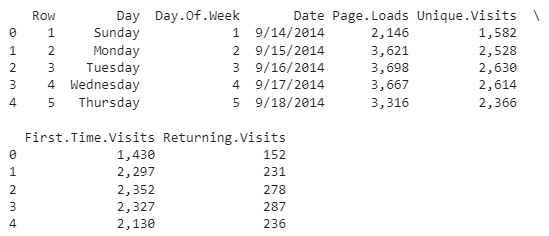
import seaborn as sns

# You can read the data into a pandas DataFrame

data = pd.read\_csv('daily-website-visitors.csv')

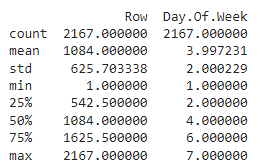
# Display the first few rows of the dataset to verify the import

print(data.head())

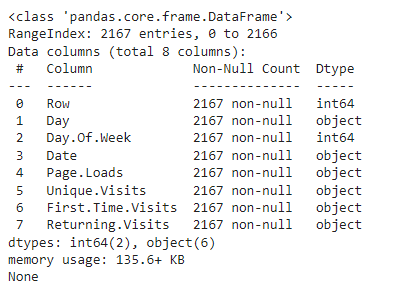


# Basic statistics and info about the dataset

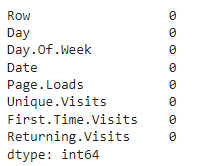
print(data.describe())



print(data.info())



data.isna().sum()



# Example 1: Bar chart for Day of the Week and Unique Visits

day\_of\_week\_counts = data['Day.Of.Week'].value\_counts().sort\_index()

plt.figure(figsize=(8, 6))

day\_of\_week\_counts.plot(kind='bar')

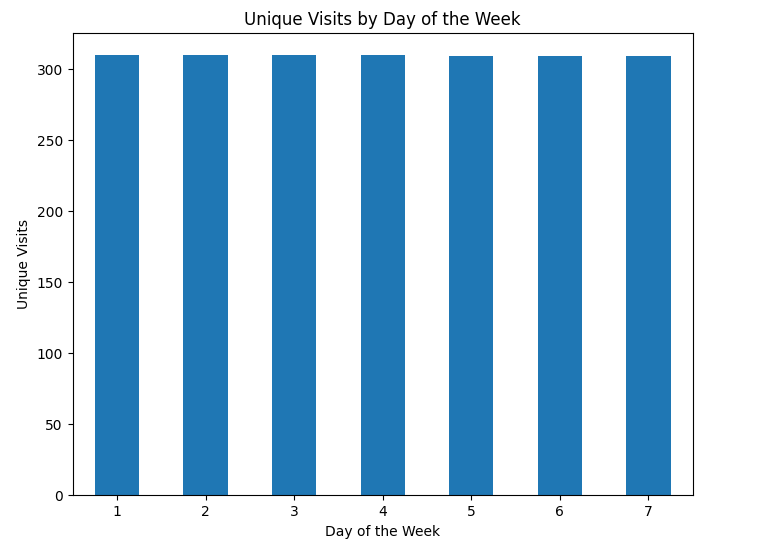
plt.title('Unique Visits by Day of the Week')

plt.xlabel('Day of the Week')

plt.ylabel('Unique Visits')

plt.xticks(rotation=0)

plt.show()



# Example 2: Pie chart for Unique Visits by Day of the Week

unique\_visits\_by\_day = data.groupby('Day.Of.Week')['Unique.Visits'].sum()

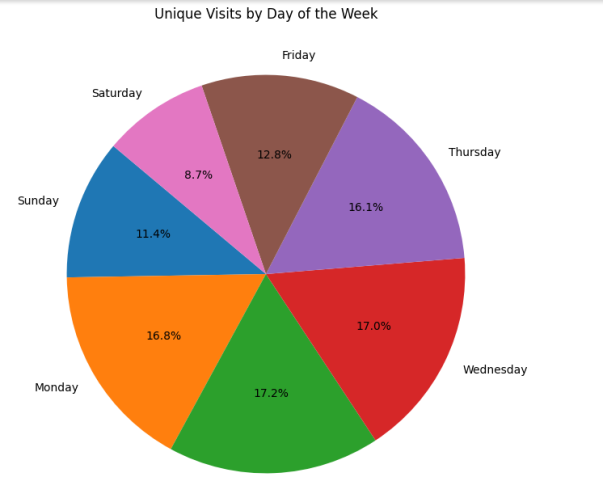
day\_labels = ['Sunday', 'Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday']

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

plt.pie(unique\_visits\_by\_day, labels=day\_labels, autopct='%1.1f%%', startangle=140)

plt.title('Unique Visits by Day of the Week')

plt.show()



# Example 3: Creating a pairplot for numeric columns

columns\_to\_plot = ['Page.Loads', 'Unique.Visits', 'First.Time.Visits', 'Returning.Visits']

# Check that the specified columns exist in your dataset

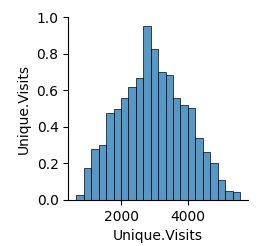
if all(col in data.columns for col in columns\_to\_plot):

    sns.pairplot(data[columns\_to\_plot])

    plt.show()

else:

    print("One or more specified columns do not exist in the dataset.")



# Example 4: Line chart for Page Loads over time

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

plt.plot(data['Date'], data['Page.Loads'], marker='o', linestyle='-')

plt.title('Page Loads Over Time')

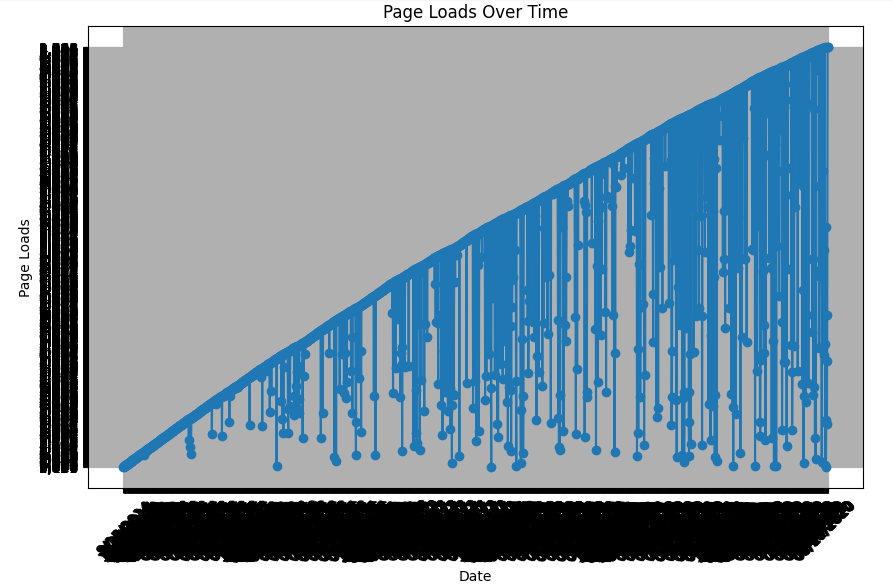
plt.xlabel('Date')

plt.ylabel('Page Loads')

plt.xticks(rotation=45)

plt.grid(True)

plt.show()



# Example 5: Heatmap

import seaborn as sns

# Assuming you have a CSV file with the following columns: Row, Day, Day.Of.Week, Date, Page.Loads, Unique.Visits, First.Time.Visits, Returning.Visits

# You can read the data into a pandas DataFrame

data = pd.read\_csv('daily-website-visitors.csv')

# Remove commas from 'Page.Loads' and convert it to a numeric format

data['Page.Loads'] = data['Page.Loads'].str.replace(',', '').astype(int)

# Extract the hour from the 'Date' column

data['Hour'] = pd.to\_datetime(data['Date']).dt.hour

# Create a pivot table to prepare the data for the heatmap

traffic\_pivot = data.pivot\_table(index='Hour', columns='Day.Of.Week', values='Page.Loads', aggfunc='sum')

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

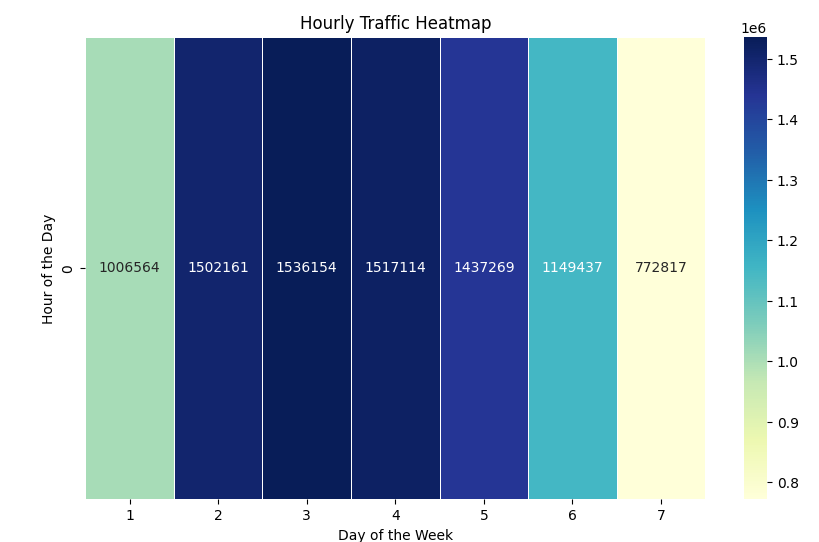
sns.heatmap(traffic\_pivot, cmap='YlGnBu', linewidths=.5, annot=True, fmt='d')

plt.title('Hourly Traffic Heatmap')

plt.xlabel('Day of the Week')

plt.ylabel('Hour of the Day')

plt.show()



# Example 6:Histogram

plt.figure(figsize=(8, 6))

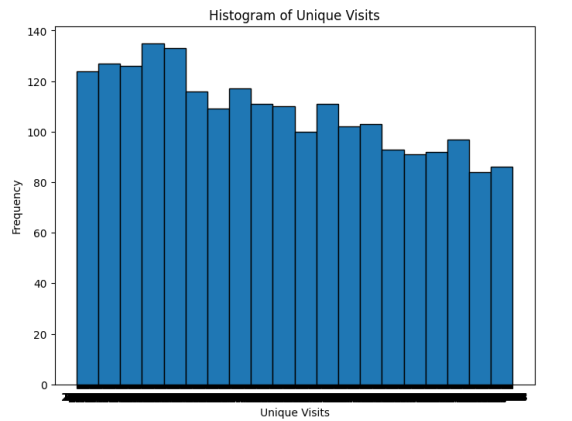
plt.hist(data['Unique.Visits'], bins=20, edgecolor='k')

plt.title('Histogram of Unique Visits')

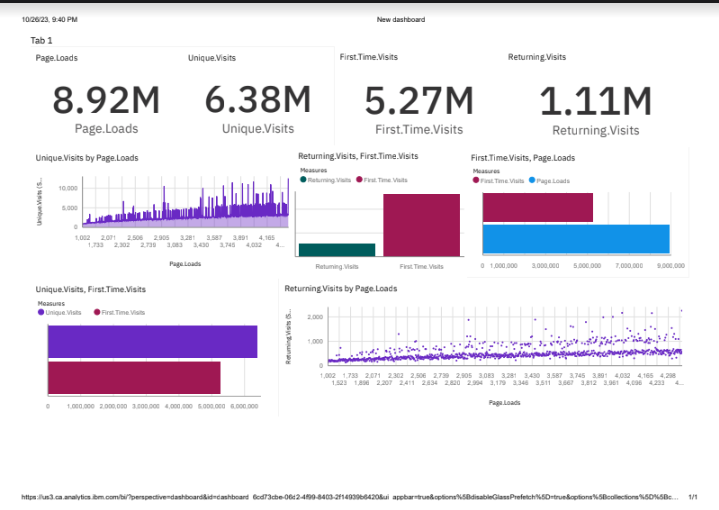
plt.xlabel('Unique Visits')

plt.ylabel('Frequency')

plt.show()



**IBM COGNOS ANALYSIS:**

****

**Insights from the Analysis:**

The analysis has unveiled valuable insights into the dynamics of website traffic and user engagement. Notably, it has revealed that certain website performance metrics, such as 'Page Loads,' 'Unique Visits,' 'First-Time Visits,' and 'Returning Visits,' play pivotal roles in understanding user behavior and optimizing web content.

* Page Loads: This metric provides insights into the overall user engagement with the website. Understanding the trends in page loads helps in assessing the popularity of different web pages and the overall traffic volume.
* Unique Visits: Tracking unique visits helps in identifying the number of distinct users accessing the website. It offers insights into the website's reach and the size of the audience.
* First-Time Visits: This metric highlights the number of users visiting the website for the first time. It provides essential insights into user acquisition and the effectiveness of marketing strategies or campaigns in attracting new visitors.
* Returning Visits: Monitoring returning visits helps gauge user retention and loyalty. It indicates how successful the website is at keeping users engaged and coming back for more.

These insights are instrumental in making data-driven decisions about website optimization, content strategies, and user experience improvements. They enable website owners, marketing teams, and developers to enhance website performance and user engagement.

The project's outcomes are particularly valuable for website owners and digital marketers seeking to improve their online presence and user interaction. The insights serve as a foundation for ongoing website traffic monitoring and management, aiding in the continuous enhancement of the website's effectiveness and the satisfaction of its users.

**Project Phases Summary**

**1. Data Collection and Understanding:**

* Gathered website traffic data, which may include information such as page views, unique visitors, time spent on site, and referral sources.
* Gained a clear understanding of the dataset's structure and the meaning of each attribute, including how data is collected and recorded.

**2. Data Preprocessing:**

* Handled missing or incomplete data to ensure data quality.
* Normalized or standardized the data as needed to maintain consistent scales for analysis.
* Efficiently split the data into relevant time periods or segments for analysis, such as daily, weekly, or monthly data.

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

* Conducted a thorough statistical analysis, calculating key statistics like average page views, unique visitor counts, and bounce rates.
* Explored correlations between variables, such as the relationship between time of day and website traffic.
* Addressed any outliers or anomalies in the data that could impact the analysis.

**4. Data Visualization:**

* Created informative visualizations, including line charts, bar graphs, and heatmaps, to better understand and communicate insights and trends within the website traffic data.
* Visualized patterns in traffic, including peak traffic hours, days, or seasons.

**5. Predictive Modeling:**

* Chose an appropriate model or algorithm for website traffic analysis, such as time series forecasting or anomaly detection.
* Built models to predict future traffic patterns or identify abnormal behavior, which may indicate issues with the website.

**6. Insights and Interpretation:**

* Based on the model's results, gained insights into the factors influencing website traffic, such as marketing campaigns, content changes, or external events.
* Offered recommendations and guidelines for improving website performance, user engagement, and traffic growth.

**7. Conclusion and Reporting:**

* Summarized findings and their implications for website performance and user engagement.
* Created a comprehensive report or presentation to share insights and recommendations with stakeholders, such as website owners, marketing teams, or web developers, to make informed decisions about website optimization and content strategies.