

**Project Report**

* **Project Title**: Analysing website traffic data
* **Student Name** Adarsh Srivastava
* **Roll Number**: 202401100300013
* **Course**: Introduction to artificial intelligence (Mr Bikki Kumar)
* **Date**: 11-03-2024

**1. Introduction**

This project aims to analyse website traffic data using Python to extract meaningful insights, such as visitor trends, page views, and bounce rates. By leveraging data analysis techniques, we can enhance website performance and user experience.

**2. Problem Statement**

-Analysing the website traffic data by the given dataset.

The primary challenge in managing a website is understanding user interaction and behavior. Without proper analysis, businesses cannot optimize their website for better engagement or identify potential issues affecting user retention.

This project focuses on analyzing website traffic data to:

* Identify trends in user visits and page views.
* Determine peak traffic hours.
* Assess bounce rates to understand user engagement.
* Provide insights for improving website performance.

**3. Methodology**

To analyze the website traffic data, the following approach is used:

1. **Data Collection**:
   * The dataset contains information such as timestamps, page views, unique visitors, and bounce rates.
   * Data is loaded from a CSV file for processing.
2. **Data Preprocessing**:
   * Cleaning and formatting data (e.g., converting timestamps, handling missing values).
   * Removing outliers that may affect analysis.
3. **Data Analysis**:
   * Identifying trends in traffic using statistical measures.
   * Analyzing peak traffic hours and visitor distribution.
   * Calculating key performance metrics like bounce rate.
4. **Visualization**:
   * Generating graphs to illustrate trends in page views and visitor patterns.
   * Visualizing search space reduction and user behavior insights.
5. **Interpretation & Insights**:
   * Extracting meaningful conclusions from the analysis.
   * Suggesting improvements based on observed trends.

**4. Code Implementation**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

# Load the website traffic data

df = pd.read\_csv("traffic\_data.csv")

#using desribe to show the basic layout of the dataset

print(df.describe())

# Plot website page views over time

plt.bar(df["Date"], df["PageViews"], label="Page Views", color='b')

plt.xlabel("Date")

plt.ylabel("Number of Page Views")

plt.title("Website Page Views Over Time")

plt.show()

# Calculate average bounce rate

average\_bounce\_rate = df["BounceRate"].mean()

print(f"Average Bounce Rate: {average\_bounce\_rate:.2f}%")

# Find peak traffic days

peak\_days = df.sort\_values(by="PageViews", ascending=False).head(5)

print("Top 5 Peak Traffic Days:")

print(peak\_days[["Date", "PageViews"]])

# Plot unique visitors over time

plt.plot(df["Date"], df["UniqueVisitors"], label="Unique Visitors", color='g')

#using line graph to visualize the data

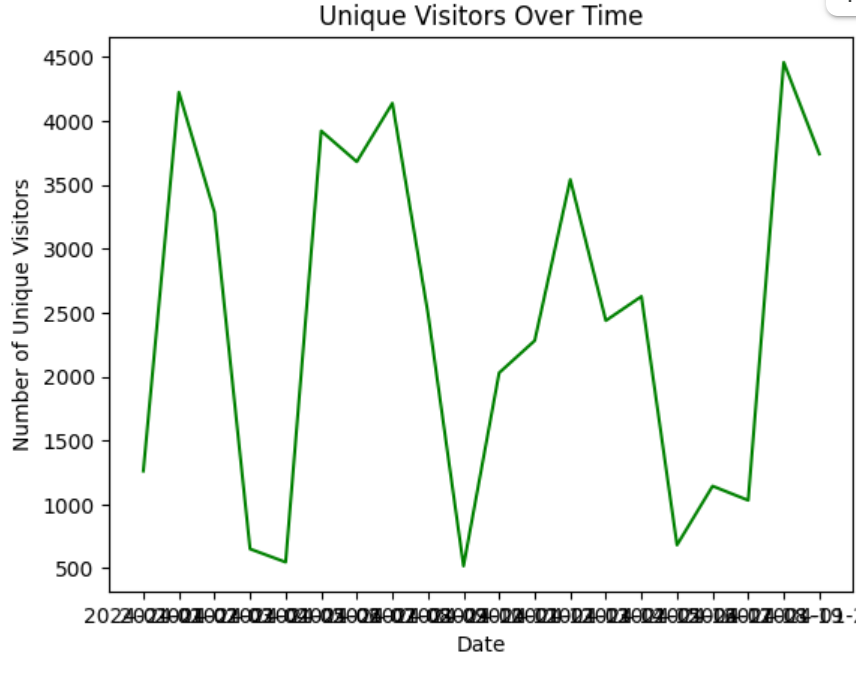
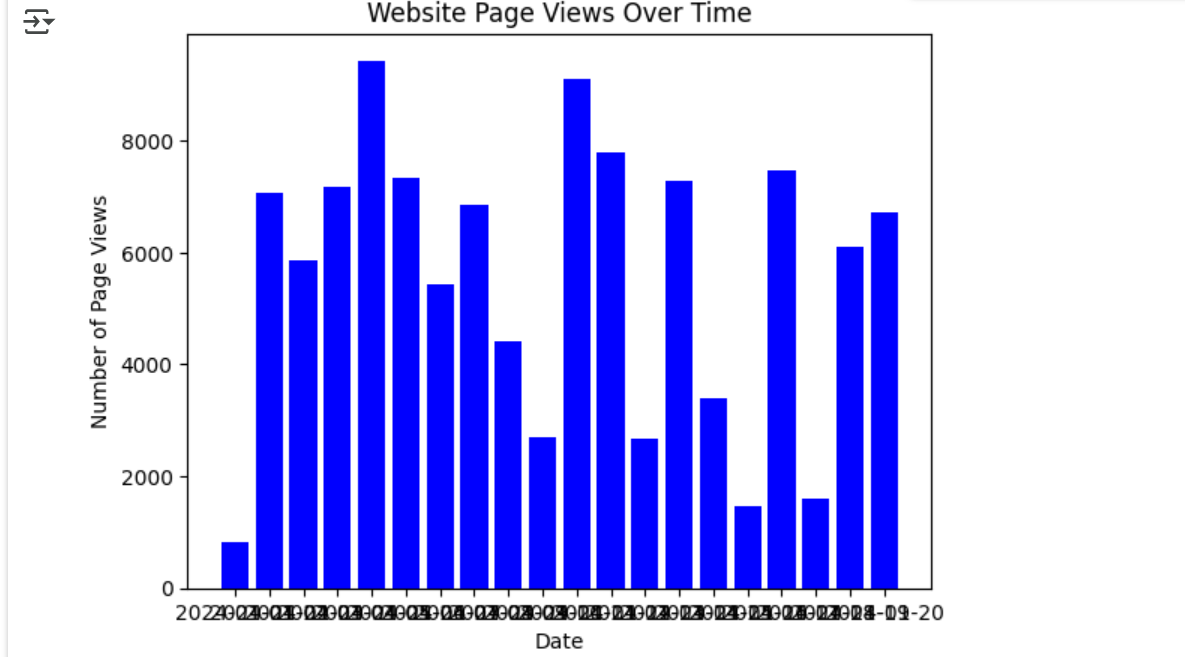
plt.xlabel("Date")

plt.ylabel("Number of Unique Visitors")

plt.title("Unique Visitors Over Time")

plt.show()

**5. Results and Output**

The output is visualized by graphs after cleaning and analysing the data

**8. References/Credits**

* Excel Datasheet to analyse the data
* Python libraries like pandas, numpy, matplotlib, sklearn modules.
* Python project by Adarsh Srivastava