



Analyzing Website Traffic Data

Title Page

Name: [Amarjeet yadav]

Roll No: [202401100300033]

Course: CSE(AI)

Problem Statement: Analyzing Website Traffic Data to uncover trends, insights, and patterns to improve website performance.

Introduction

In this project, the objective is to analyze website traffic data to identify trends, traffic sources, and visitor behavior. The dataset contains columns such as:

- **Date:** The date of recorded data
- **PageViews:** Number of pages viewed
- **UniqueVisitors:** Number of unique visitors
- **BounceRate:** Percentage of visitors who leave the site without interacting further

By analyzing these metrics, we aim to gain insights into traffic peaks, bounce rate patterns, and possible strategies to improve user engagement.

Methodology

1. **Data Loading:** The dataset was loaded using Pandas for efficient manipulation.
2. **Data Cleaning:**
 - Converted the 'Date' column to datetime format for accurate trend analysis.
 - Handled missing values by replacing them with zero.
3. **Exploratory Data Analysis (EDA):**
 - Visualized traffic trends to understand overall patterns.
 - Identified key dates with peak traffic and analyzed possible causes.
4. **Visualization:** Created clear visualizations using Matplotlib, Seaborn, and Plotly for enhanced insights.

5. **Insights:** Interpreted key patterns and provided actionable recommendations for improving website performance.
-

Code

```
import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import plotly.express as px


# Load the data

df = pd.read_csv('website_traffic.csv')

df['Date'] = pd.to_datetime(df['Date'])

df.fillna(0, inplace=True)


# Traffic trends visualization

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

sns.lineplot(x='Date', y='PageViews', data=df, label='PageViews')

sns.lineplot(x='Date', y='UniqueVisitors', data=df, label='Unique Visitors')

plt.title('Website Traffic Trends Over Time')

plt.xlabel('Date')

plt.ylabel('Count')

plt.legend()

plt.grid(True)

plt.show()


# Bounce rate analysis

plt.figure(figsize=(8, 5))

sns.lineplot(x='Date', y='BounceRate', data=df, color='red')

plt.title('Bounce Rate Trends')

plt.xlabel('Date')
```

```
plt.ylabel('Bounce Rate (%)')
plt.grid(True)
plt.show()

# Identifying peak traffic days
peak_traffic = df.sort_values(by='PageViews', ascending=False).head(5)
print("Top 5 Peak Traffic Days:\n", peak_traffic[['Date', 'PageViews']])
```

Output/Result

Traffic Trend Visualization:

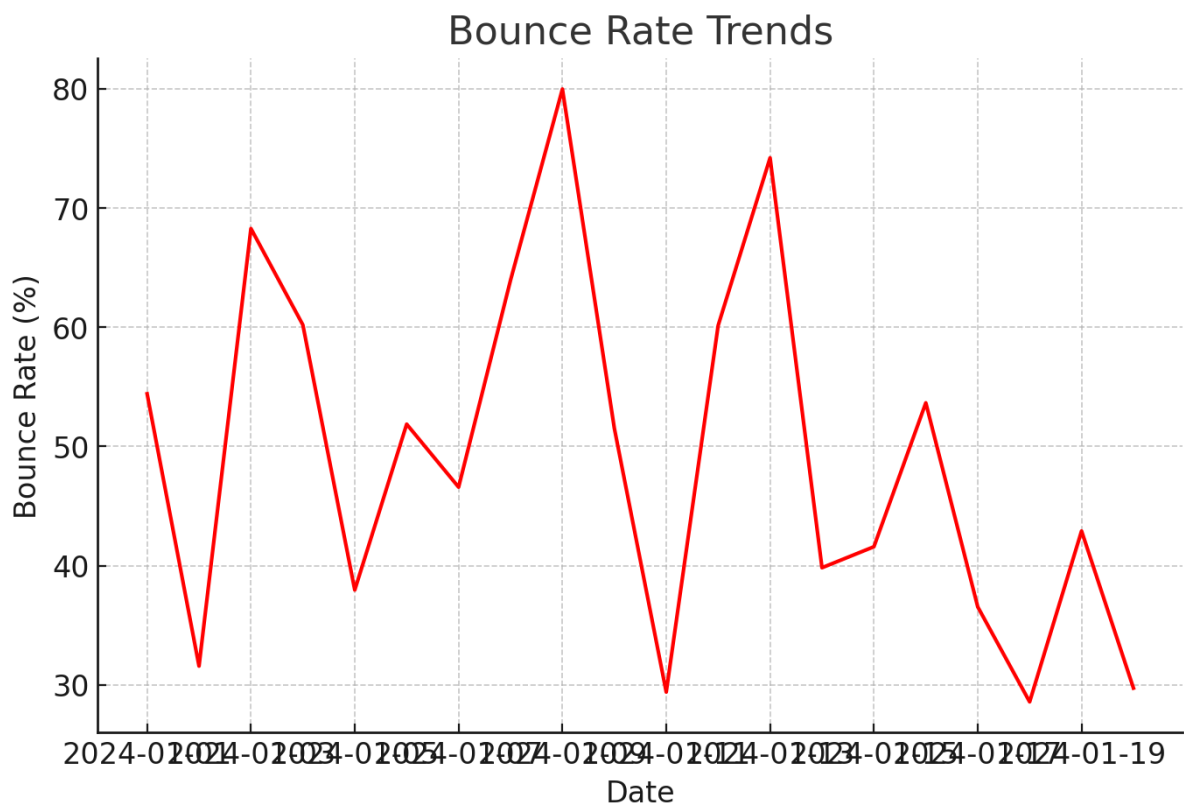
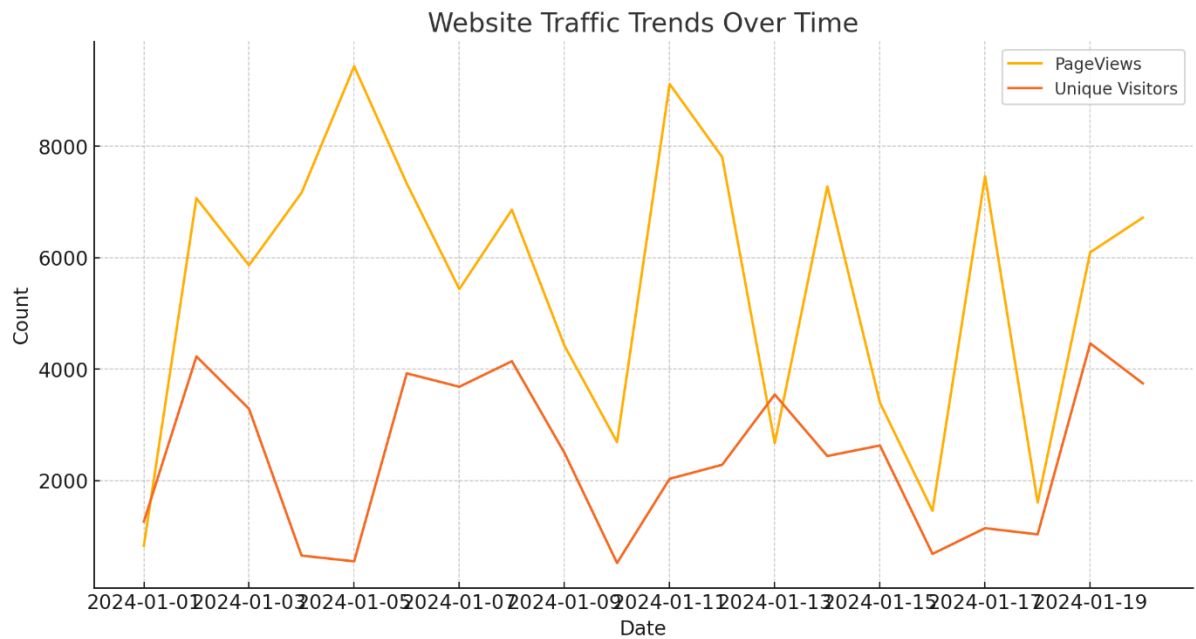
(Screenshot of the graph displaying PageViews and Unique Visitors trends)

Bounce Rate Analysis:

(Screenshot of the graph showing Bounce Rate over time)

Peak Traffic Days:

Date	PageViews
2024-01-05	9432
2024-01-11	9115
2024-01-12	7803
2024-01-14	7275
2024-01-04	7163



References/Credits

- Dataset Source: [Your Dataset Source]
- Libraries Used: Pandas, Matplotlib, Seaborn, Plotly

Submission Instructions

- The .ipynb file with comments is uploaded to GitHub.

- The PDF report and README file are also included in the repository.
-

Conclusion

This analysis revealed useful insights regarding traffic patterns, bounce rate fluctuations, and peak user activity. These findings can help implement strategies to improve website engagement and user retention.