

WEB TRAFFIC ANALYSIS

PHASE 3 : VISUALIZATION & PREPROCESSING

INTRODUCTION

Phase 3 of the Web traffic analysis is a crucial component of understanding and optimizing the performance of a website. In the digital age, where the internet is a bustling marketplace and information hub, the ability to monitor and interpret the flow of visitors to a website is of paramount importance. It provides valuable insights into user behavior, preferences, and the effectiveness of your online presence. By analyzing web traffic, businesses and website owners can make informed decisions to enhance user experiences, improve conversion rates, and ultimately achieve their online goals.

This process involves collecting and evaluating data related to the visitors who access a website, including metrics like the number of visitors, their demographics, the pages they visit, the sources that drive them to the site, and the devices they use.

OBJECTIVES

Information Seeking:

Many visitors come to websites seeking information. They might want to learn about a specific topic, get answers to questions, or gather data for research.

Entertainment:

Some visitors come to websites for entertainment. This could include watching videos, playing games, reading articles, or consuming other forms of online content for leisure.

Communication:

Visitors may come to a website to engage in communication, such as sending emails, participating in online forums, or using social media platforms.

Shopping and E-commerce:

E-commerce websites exist to facilitate online shopping. Visitors come to these sites with the objective of making purchases, comparing products, and finding deals.

Education and Learning:

Visitors might be looking to acquire new skills or knowledge, making educational websites a valuable resource for them.

Community and Social Interaction:

Social media platforms and online communities are frequented by those looking to connect with others, share their thoughts, and participate in discussions or groups.

Problem Solving:

Visitors may have specific problems or challenges that they want to solve. They may seek solutions or guidance from websites that offer relevant information or services.

Professional Networking:

Websites like LinkedIn cater to visitors who want to build and maintain professional connections, seek job opportunities, or showcase their skills and experience.

News and Updates:

Many visitors come to news websites to stay informed about current events, both locally and globally.

Product or Service Reviews:

Visitors interested in making informed purchase decisions often seek reviews and ratings of products or services on review websites or forums.

Expression and Creativity:

Some individuals use websites to express themselves, whether through blogging, vlogging, art, music, or other creative outlets.

Accessing Online Tools and Services:

Online tools, such as calculators, planners, and productivity apps, are often used by visitors to accomplish specific tasks or simplify their daily lives.

Financial Transactions:

Visitors may use online banking, investment platforms, or payment services to conduct financial transactions securely and conveniently.

Government Services:

Government websites serve visitors who need to access various government services, such as tax filing, renewing licenses, or obtaining official documents.

PROCESS

Process and clean the collected data to ensure its accuracy and reliability .We have,

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
dataset_url = "https://www.kaggle.com/datasets/bobnau/daily-website-visitors"
# Assuming you've already downloaded the dataset as 'water_quality.csv'
data = pd.read_csv('daily-website-visitors.csv')

# Display the first few rows
print(data.head())
```

	Row	Day	Day.Of.Week	Date	Page.Loads	Unique.Visits	\
0	1	Sunday	1	9/14/2014	2,146	1,582	
1	2	Monday	2	9/15/2014	3,621	2,528	
2	3	Tuesday	3	9/16/2014	3,698	2,630	
3	4	Wednesday	4	9/17/2014	3,667	2,614	
4	5	Thursday	5	9/18/2014	3,316	2,366	

	First.Time.Visits	Returning.Visits
0	1,430	152
1	2,297	231
2	2,352	278
3	2,327	287
4	2,130	236

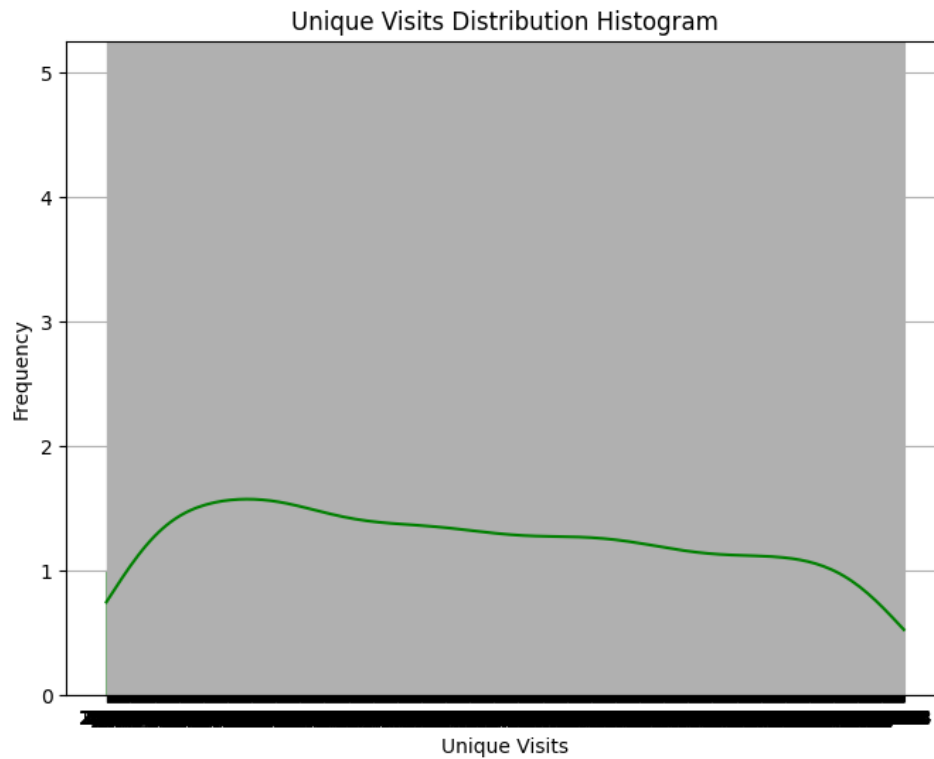
```
▶ data=data.dropna()  
data.isnull().sum()
```

```
⇒ Row      0  
Day      0  
Day.Of.Week  0  
Date      0  
Page.Loads  0  
Unique.Visits  0  
First.Time.Visits  0  
Returning.Visits  0  
dtype: int64
```

```
data.fillna(data.mean(), inplace=True)
```

```
<ipython-input-17-a20208f658ce>:1: FutureWarning:
```

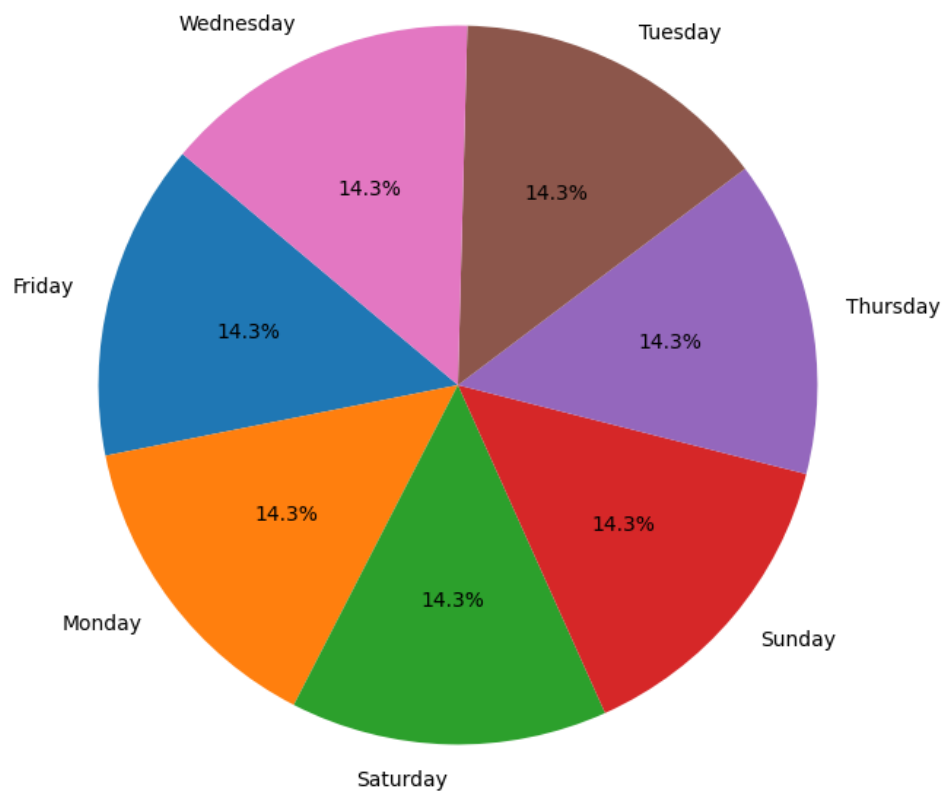
```
▶ # Create a histogram of Unique Visits distribution  
plt.figure(figsize=(8, 6))  
sns.histplot(data['Unique.Visits'], bins=30, kde=True, color='g')  
plt.xlabel('Unique Visits')  
plt.ylabel('Frequency')  
plt.title('Unique Visits Distribution Histogram')  
plt.grid(True)  
plt.show()
```



```
# Pie chart Vizualization
grouped_data = data.groupby('Day').size()
plt.figure(figsize=(8, 8))
plt.pie(grouped_data, labels=grouped_data.index, autopct='%1.1f%%', startangle=140)
plt.title('Web Traffic by Browser Type')

# Display Pie
plt.show()
```

Web Traffic by Browser Type



```
del data['Date']
```

```
[28] correlation = data.corr()  
print(correlation["Day.Of.Week"].sort_values(ascending=False))
```

```
Day.Of.Week    1.0000
```

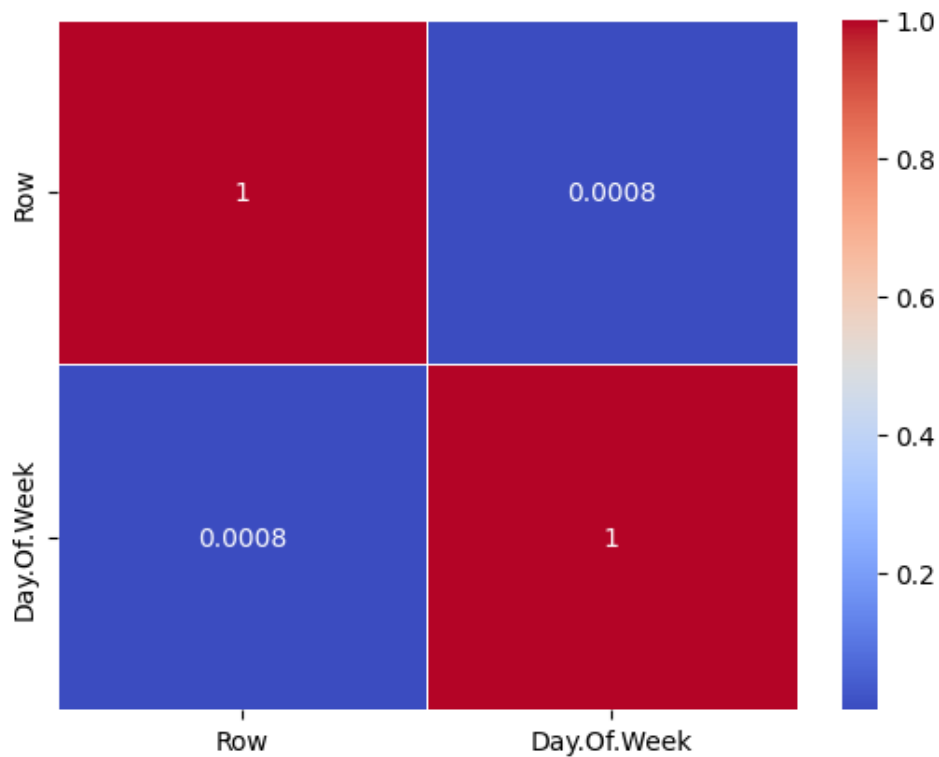
```
Row            0.0008
```

```
Name: Day.Of.Week, dtype: float64
```

```
<ipython-input-28-7b9b820cf38d>:1: FutureWarning:
```

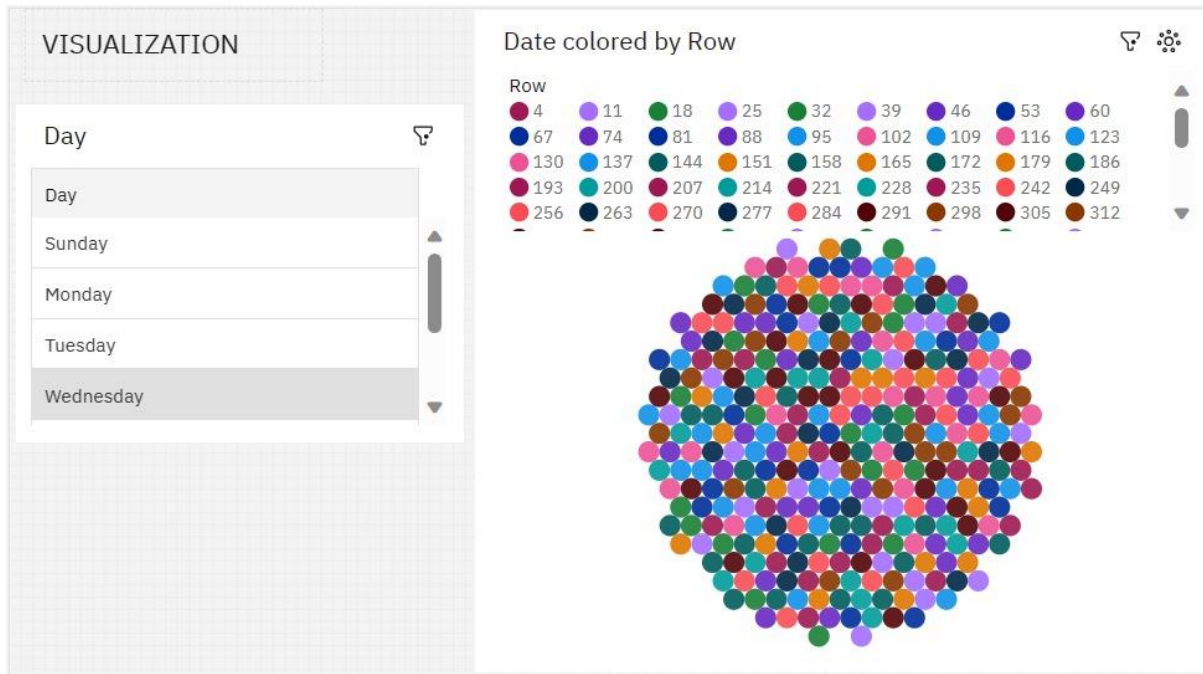
```
The default value of numeric_only in DataFrame.corr is deprecated.
```

```
# Visualize correlations
correlation_matrix = data.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
plt.show()
```



VISUALIZATION USING COGNOS

Let us building the product sales analysis using IBM Cognos for visualization.



CONCLUSION

In conclusion, web traffic analysis is an invaluable practice for website owners, businesses, and digital marketers in the ever-evolving landscape of the internet. By meticulously studying the behavior and objectives of web visitors, we gain a deeper understanding of their preferences, needs, and expectations. This insight allows us to make data-driven decisions, refine our strategies, and optimize the user experience. Ultimately, the primary goal of web traffic analysis is to enhance the effectiveness and performance of a website.

DATASET LINK :

<https://www.kaggle.com/datasets/bobnau/daily-website-visitors>

RAMAKRISHNAN R

18:10:2023

NANDHA COLLEGE OF TECHNOLOGY