PHASE – 4

AIM:

To make advanced analysis for web traffic by using python with integrated IBM cognos.

DESCRIPTION:

- Building the analysis by creating visualizations using IBM Cognos .
- integrating Python code for advanced analysis.
- Use IBM Cognos to create interactive dashboards and reports that display insights such as popular pages, traffic sources, and user engagement metrics.
- Use Python libraries like Pandas and Matplotlib to perform more complex analyses on the data, such as time series analysis, user segmentation, or machine learning-based predictions.

INTEGRATION OF PYTHON CODE:

Time series:

A time series is a data set that tracks a sample over time. In particular, a time series allows one to see what factors influence certain variables from period to period.

We have used the basic plotting function to achieve the time series to plot the graph(Date vs PageLoads)

Here we assume first 100 data among the given dataset(daily-website-visitors)

Program:

```
import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

data=pd.read_csv("daily-website-visitors.csv")

data["Date"] = pd.to_datetime(data["Date"],format="%m/%d/%Y")

plt.figure(figsize=(20,80))

plt.plot(data['Date'].head(100),data['Page.Loads'].head(100))

plt.title("Time Series Data")

plt.xlabel("Date")

plt.ylabel("Page.Loads")

plt.legend()

plt.show()
```
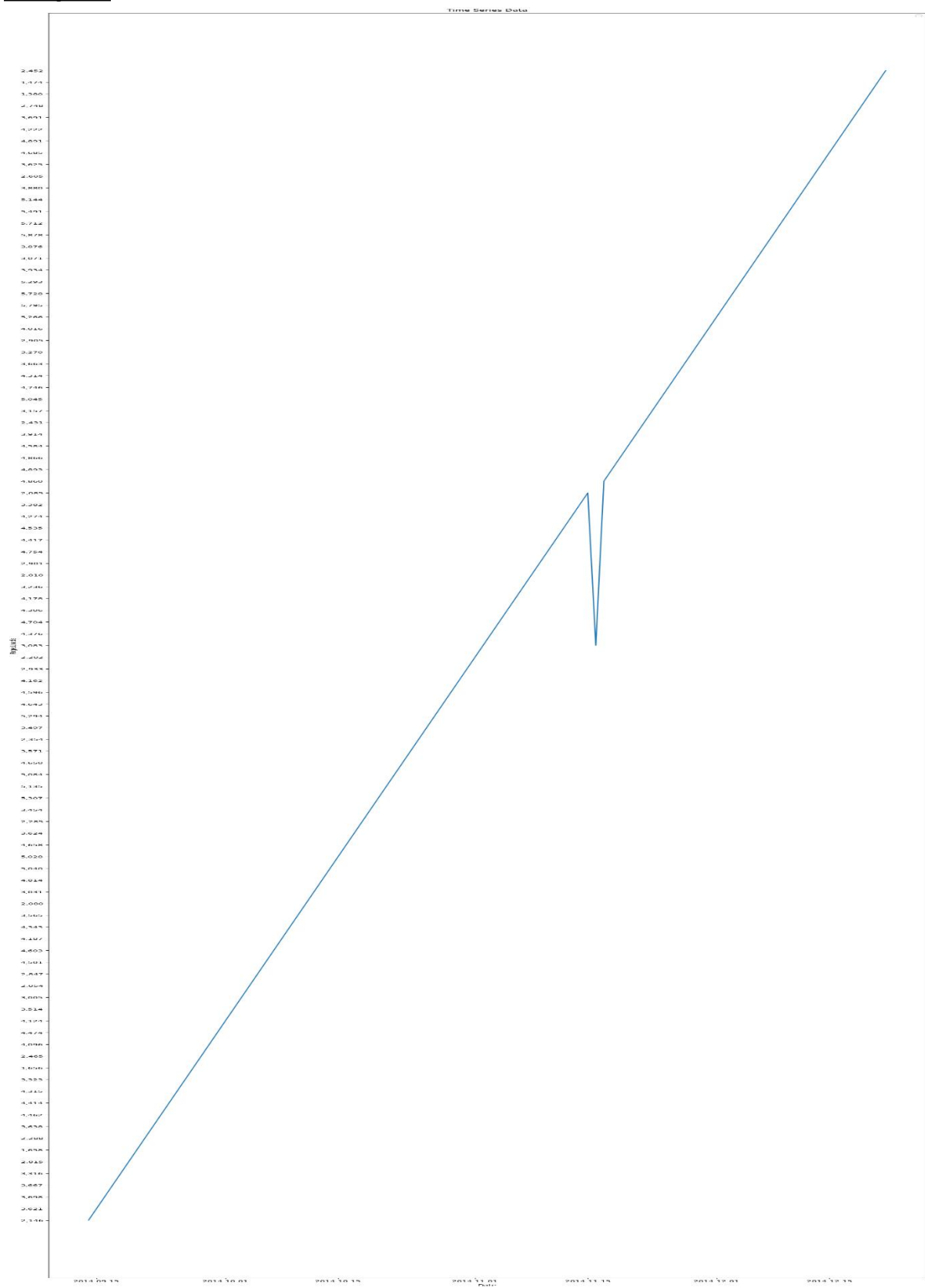
Time Series Data

Conclusion based on time series:

The time series analysis of our website traffic data has been instrumental in unraveling patterns, trends, and fluctuations in user activity over a specific period. By examining this data, we've gained valuable insights that can inform our decision-making and strategic planning.

Understanding the temporal dynamics of website traffic enables us to identify peak activity periods, and assess the impact of marketing campaigns or website changes. It equips us with the knowledge needed to allocate resources efficiently, optimize content release schedules, and ensure that our website is performing at its best during critical moments.

As we move forward, it's crucial to continue monitoring and analyzing website traffic data over time. This ongoing time series analysis will help us adapt to changing user behaviors and market conditions, ensuring that we can remain responsive and proactive in meeting our goals and objectives.

Segmentation:

  User segmentation in data analytics involves dividing a target audience or customer base into distinct groups based on specific characteristics or behaviors. These segments are created to better understand, target, and tailor marketing strategies, product offerings, or user experiences.This way of grouping data may include demographic information, behavioral data, purchase history, website interactions, app usage, and more.

  Here we use the basic bar plot to achieve the user segmentation

  We divided the given segments of pageloads based on the day of week. Each number represents the days of the week as

1. Sunday
2. Monday
3. Tuesday
4. Wednesday
5. Thursday
6. Friday
7. Saturday

The below program inucludes basic bar plotting using python

Program:

```python
import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

data=pd.read_csv("daily-website-visitors.csv")

ds=data['Page.Loads'].head()

xy=data['Day.Of.Week'].head()

df=data.describe()

print(df)

#segmentation

data["Date"] = pd.to_datetime(data["Date"],format="%m/%d/%Y")

plt.figure(figsize=(40, 80))

plt.bar(data['Day.Of.Week'].head(100), data['Page.Loads'].head(100));

plt.show()
```

Output:
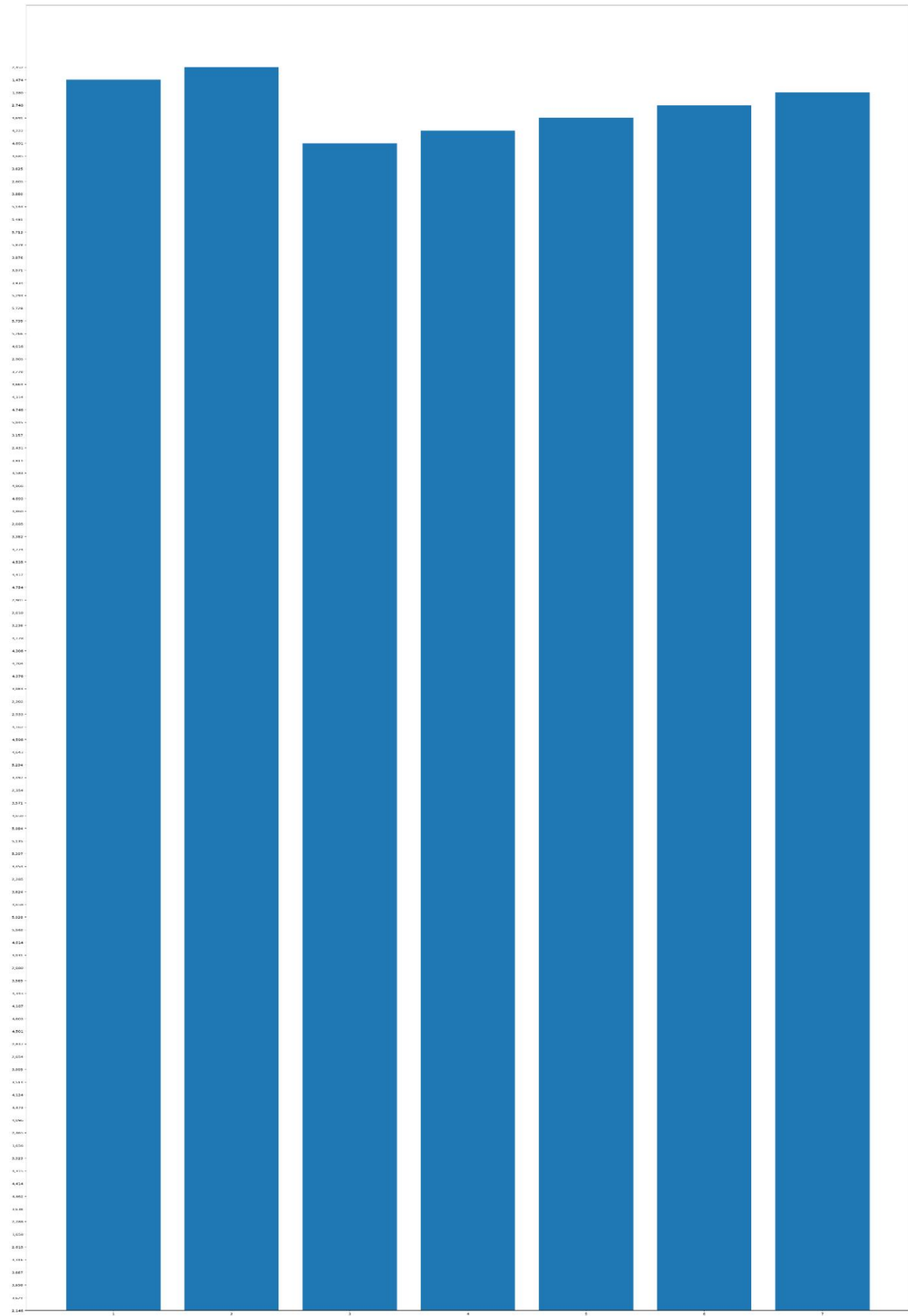
Conclusion based on segmentation:

The user segmentation analysis of our website traffic has unveiled a wealth of information that can significantly impact our digital strategy. By categorizing our visitors into distinct segments based on days, we've gained a deeper understanding of our audience. This knowledge can be leveraged to tailor content, marketing efforts, and user experiences to better meet the expectations of our users.

User segmentation not only allows us to identify high-value customer groups but also pinpoint areas for improvement in our website's performance. By customizing our approach for different segments, we can enhance user engagement, increase conversion rates, and ultimately drive business growth.

Moving forward, we should continue to refine our user segmentation strategies and keep a close eye on how these segments evolve over time. This ongoing analysis will enable us to stay agile and responsive to changing user needs, ensuring that our website remains a valuable asset in achieving our objectives.

ANALYSIS USING IBM COGNOS ANALYTICS :

       IBM Cognos Analytics with Watson is a web-based integrated business intelligence suite by IBM. It provides a toolset for reporting, analytics, scorecarding, and monitoring of events and metrics.

       IBM Cognos Analytics integrates **reporting**, modeling, analysis, **dashboards**, stories, and event management so that you can understand your organization's data, and make effective business decisions.
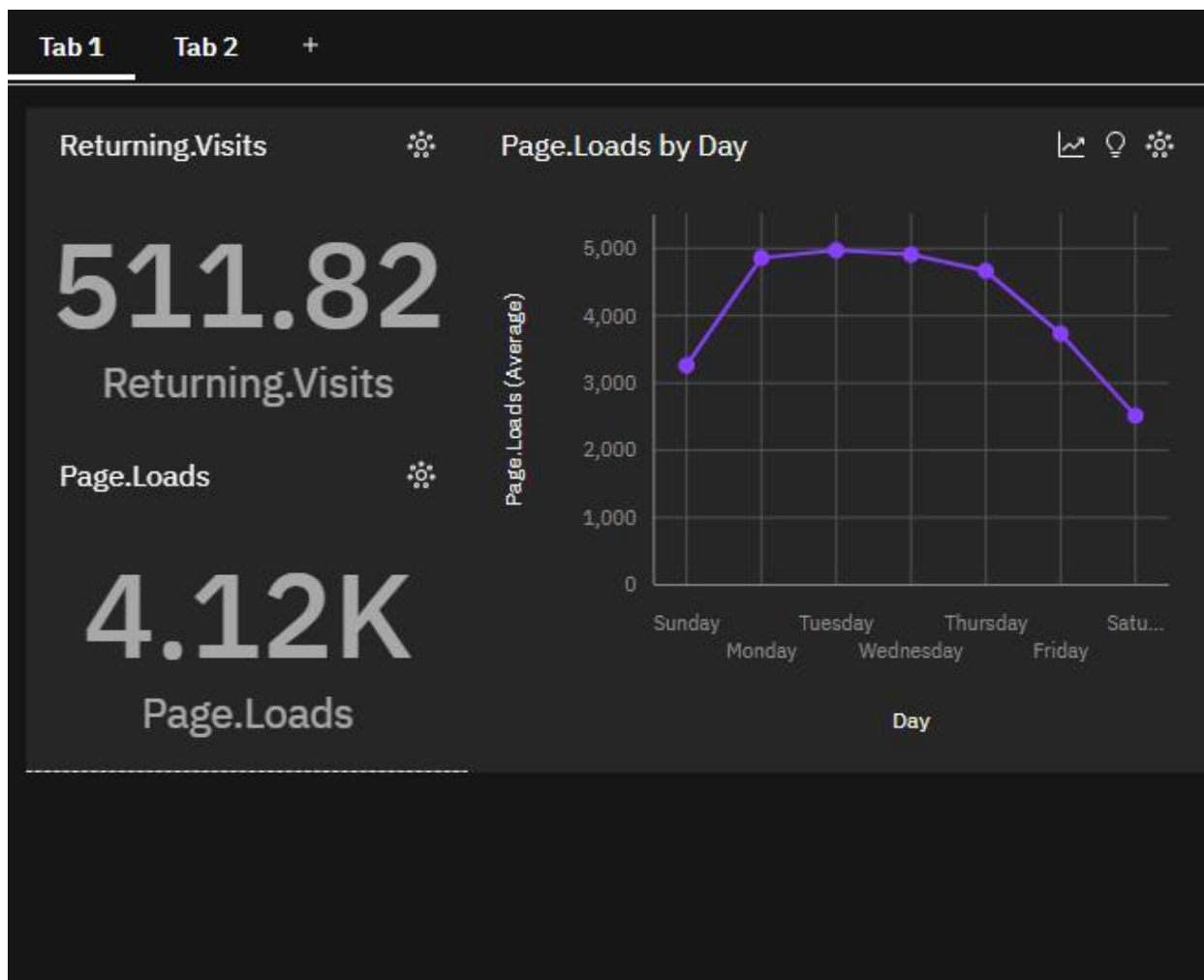
Dashboards:

       A dashboard helps you to monitor events or activities at a glance by providing key insights and analysis about your data on one or more pages or screens.

       Here we created a dashboard using the cognos analytics tool with the factors such as pageloads,uniquevisits,etc… From the given dataset.
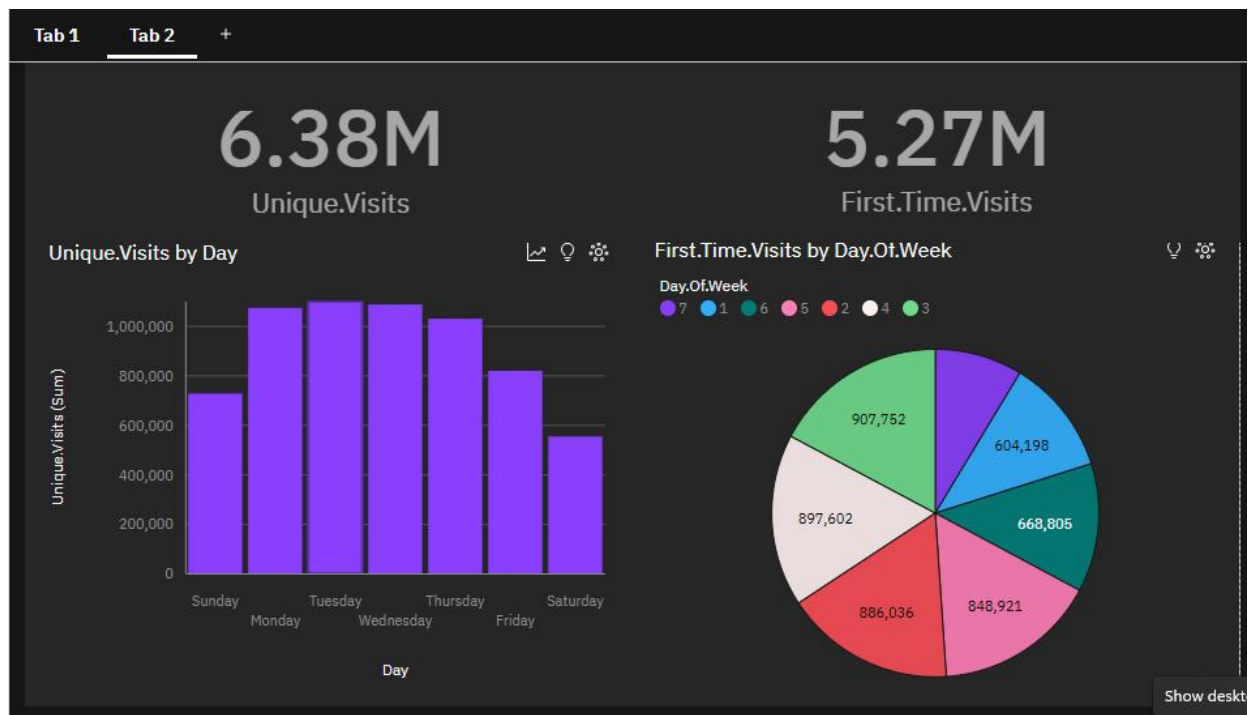
Dashboard Link:

https://ap2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2Fdashboard%2528pageloads%2529&action=view&mode=dashboard&subView=model0000018b84d98916_00000000

Dashboard 1 preview:

On the above dashboard we have three separate sections where each section indicates specific information which will be used to understand the dataset .The first section of dashboard we have displayed the average returning visits from the total Page Loads which will specifies how many users where returned to the website again .This information will provides the community or the organization how familiar does the user uses the website in average of the day. On the second section we included the average page loads. On comparing the total page loads with the returning visits we can validate how often the page is visited . The third section is based on the page loads by week days in the line graph visualization we have plotted the page loads with the days. The graph shows how many users where visited the website on the specific day.

Dashboard 2 preview:

We have created an another dashboard which will contain the two graphs where one of the graph will contain the total unique visits by day and other one will have first time visits by day. We have initially calculated the total unique visits and total first time visits overall period of time. Considering the bar graph it will show the total unique visits on each day. From the bar graph Saturday of the week will have very low unique visits compared to other days of the week and Tuesday of the will have the highest unique visits . The pie chart includes the total first time visits per day of the week. Each segment in the pie chart will represent the day of the week. Pie chart shows that Saturday has the lowest first time visitors and Tuesday has the highest first time visitors respectively.

Report:

Reporting a web-based report authoring tool that professional report authors and developers use to build sophisticated, multiple-page, multiple-query reports against multiple databases .

In the below report we have represented the count of page loads for the given period of time. On the other side of the report we have the same Total page loads based on the day of week.

Report Link :

https://ap2.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2Fwebsite-visitors&action=run&prompt=false

Reportpreview:
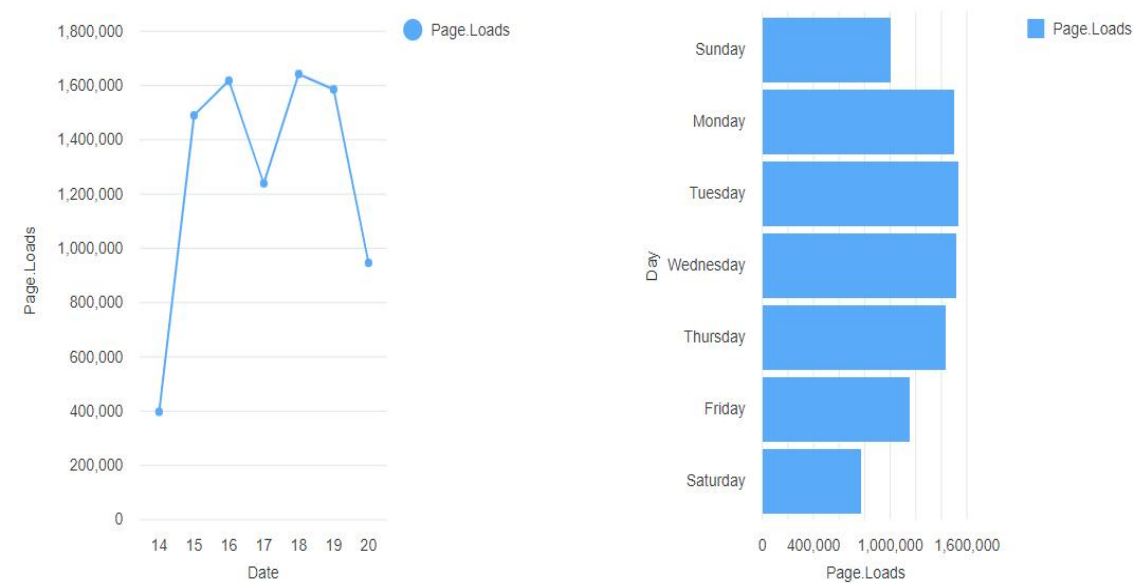
8,921,516
Page.Loads

6,378,882
Unique.Visits

5,269,763
First.Time.Visits

1,109,119
Returning.Visits

Day.Of.Week	Unique.Visits	First.Time.Visits	Returning.Visits	Page.Loads
2	1,072,112	886,036	186,076	4,845.68064516
4	1,085,624	897,602	188,022	4,893.91612903
5	1,028,214	848,921	179,293	4,651.35598706
1	725,794	604,198	121,596	3,246.98064516
6	817,852	668,805	149,047	3,719.86084142
3	1,097,181	907,752	189,429	4,955.33548387
7	552,105	456,449	95,656	2,501.02588997

Conclusion:

In conclusion, the website traffic analysis has provided valuable insights into the performance and user engagement on the given website. By examining key metrics such as page views, unique visitors, returning visit, and first time visitors, we have gained a better understanding of how our online presence is faring. These insights can help us make informed decisions to optimize the website, refine content strategy, and enhance the overall user experience. It is clear that continuous monitoring and analysis of website traffic data is essential for staying competitive in the digital landscape and ensuring that our online presence remains effective in meeting our goals and objectives.

Team members:

Sakthivel.R

Girivasan.k

Poovarasan.M

Saravanan.C