PHASE 3 : DEVELOPMENT PART 1

WEBSITE TRAFFIC ANALYSIS

PRIYADARSHINI ENGINEERING COLLEGE

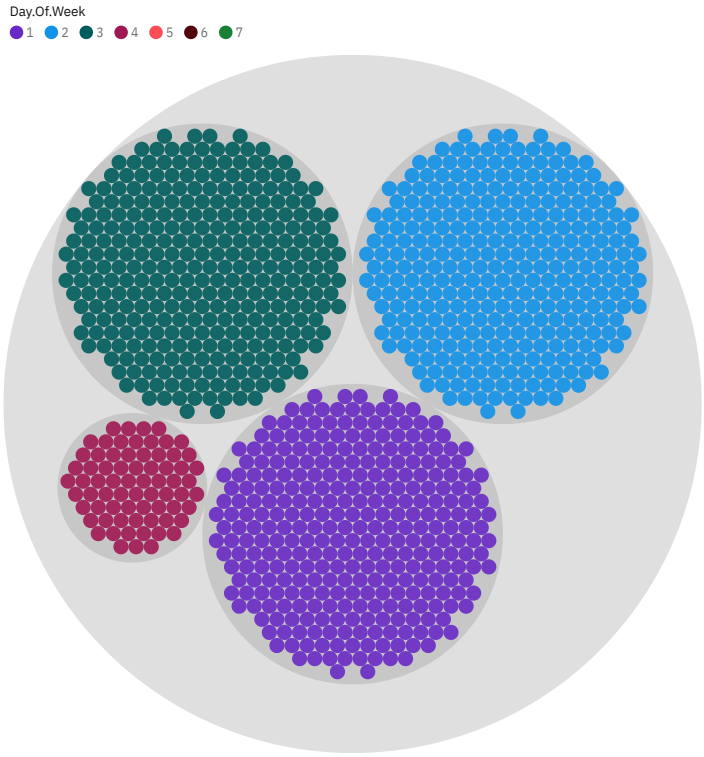
Website traffic analysis is the process of collecting, examining, and interpreting data related to the visitors and interactions on a website. It provides invaluable insights into user behavior, preferences, and trends, helping organizations make informed decisions, optimize their online presence, and enhance user experiences.

Abstract:

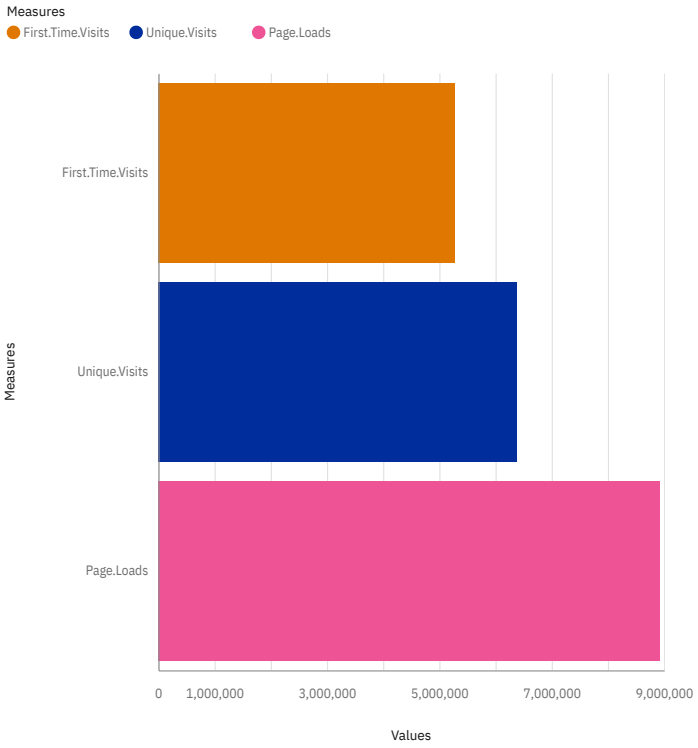
This project aims to analyze website traffic data for insights into user behavior, popular pages, and traffic sources. It involves data collection, visualization using IBM Cognos, and Python for advanced analysis. The goal is to optimize user experiences and enhance website performance.

Tab 1

Day and Date hierarchy colored by Day.Of.Week



First.Time.Visits, Unique.Visits, Page.Loads

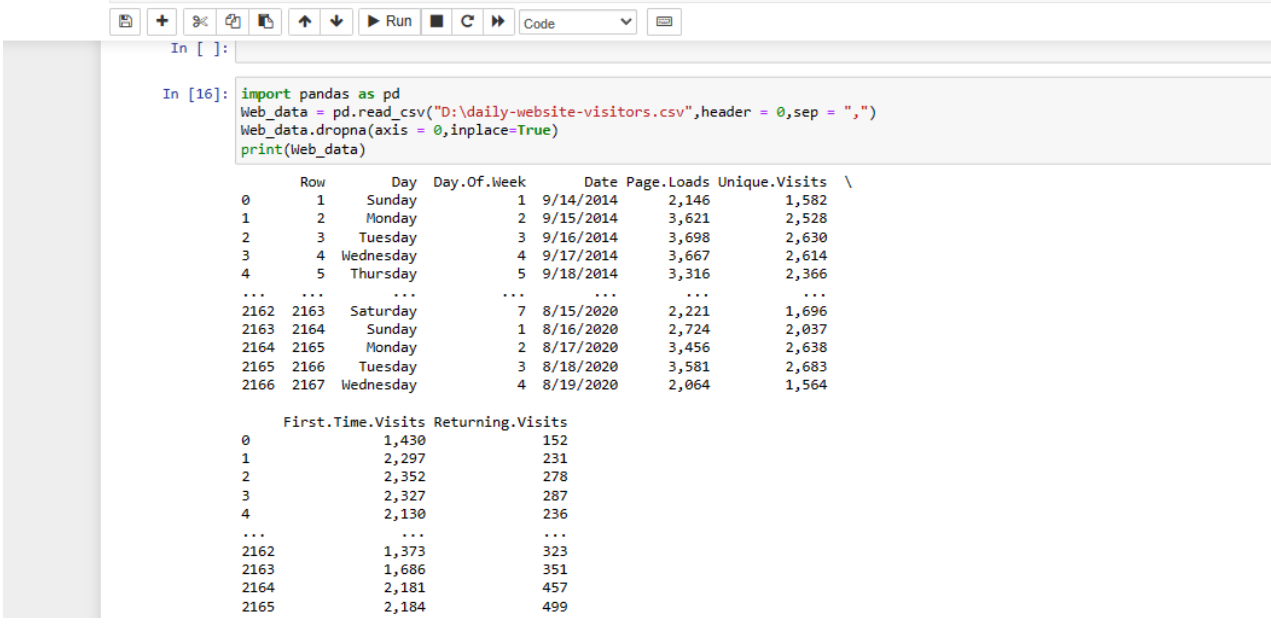


Insights :

Based on the current forecasting, **First.Time.Visits** may reach **over 395 thousand** by **Day Monday+1**.
The overall number of results for **First.Time.Visits** is **over two thousand**.
The overall number of results for **Page.Loads** is **over two thousand**.
The overall number of results for **Unique.Visits** is **over two thousand**

'Python Integration' for Website Traffic Analysis :

Including the Insights into the Python to know the actual data that provided by Kaggle



```
In [ ]:

In [16]: import pandas as pd
         Web_data = pd.read_csv("D:\daily-website-visitors.csv",header = 0,sep = ",")
         Web_data.dropna(axis = 0,inplace=True)
         print(Web_data)
```

```
       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
...    ...        ...          ...        ...        ...           ...
2162  2163   Saturday            7  8/15/2020      2,221         1,696
2163  2164     Sunday            1  8/16/2020      2,724         2,037
2164  2165     Monday            2  8/17/2020      3,456         2,638
2165  2166    Tuesday            3  8/18/2020      3,581         2,683
2166  2167  Wednesday            4  8/19/2020      2,064         1,564

     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
...                ...              ...
2162             1,373              323
2163             1,686              351
2164             2,181              457
2165             2,184              499
```

```
2166            1,297            267

[2167 rows x 8 columns]
```

Getting of data set information using info function.

```
In [17]: print(Web_data.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2167 entries, 0 to 2166
Data columns (total 8 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Row                2167 non-null   int64
 1   Day                2167 non-null   object
 2   Day.Of.Week        2167 non-null   int64
 3   Date               2167 non-null   object
 4   Page.Loads         2167 non-null   object
 5   Unique.Visits      2167 non-null   object
 6   First.Time.Visits  2167 non-null   object
 7   Returning.Visits   2167 non-null   object
dtypes: int64(2), object(6)
memory usage: 135.6+ KB
None

In [ ]:

In [18]: print(Web_data.describe())

                Row  Day.Of.Week
count  2167.000000  2167.000000
mean   1084.000000     3.997231
std     625.703338     2.000229
min       1.000000     1.000000
25%     542.500000     2.000000
50%    1084.000000     4.000000
75%    1625.500000     6.000000
max    2167.000000     7.000000

In [ ]:
```

Printing the values of data set using Pandas Library

```
In [3]: df = pd.read_csv("D:\daily-website-visitors.csv")
        df
```

Out[3]:

	Row	Day	Day.Of.Week	Date	Page.Loads	Unique.Visits	First.Time.Visits	Returning.Visits
0	1	Sunday	1	9/14/2014	2,146	1,582	1,430	152
1	2	Monday	2	9/15/2014	3,621	2,528	2,297	231
2	3	Tuesday	3	9/16/2014	3,698	2,630	2,352	278
3	4	Wednesday	4	9/17/2014	3,667	2,614	2,327	287
4	5	Thursday	5	9/18/2014	3,316	2,366	2,130	236
...	...	...	...	...	...	...	...	...
2162	2163	Saturday	7	8/15/2020	2,221	1,696	1,373	323
2163	2164	Sunday	1	8/16/2020	2,724	2,037	1,686	351
2164	2165	Monday	2	8/17/2020	3,456	2,638	2,181	457
2165	2166	Tuesday	3	8/18/2020	3,581	2,683	2,184	499
2166	2167	Wednesday	4	8/19/2020	2,064	1,564	1,297	267

2167 rows × 8 columns