WEBSITE TRAFFIC ANALYSIS

PHASE 3 : DEVELOPMENT PART 1

PRIYADARSHINI ENGINEERING COLLEGE

SHREEVATHSHAN R / 511921104065

1.0 INTRODUCTION:

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.

1.1 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.

1.2 Problem Statement:

The problem of "Website Traffic Analysis" lies in the need for organizations to effectively understand and leverage user behavior on their websites

2.0 Dataset

The dataset given is "Daily web visitors" from kaggle.com

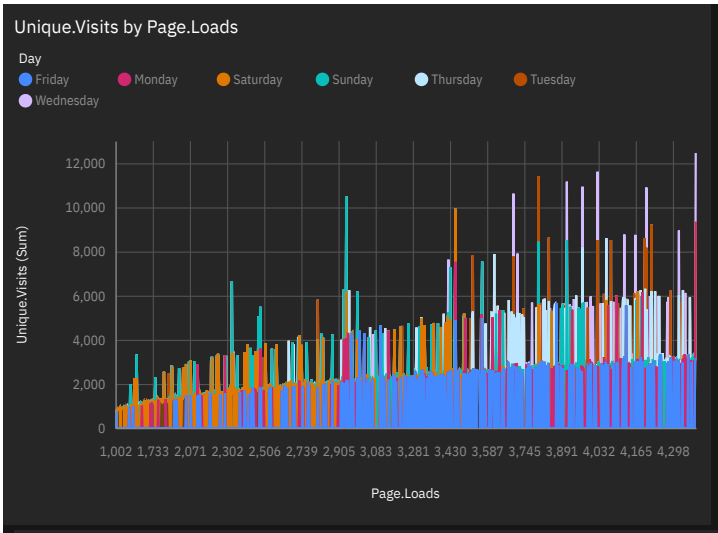
Link:https: //www.kaggle.com/datasets/bobnau/daily-website-visitors

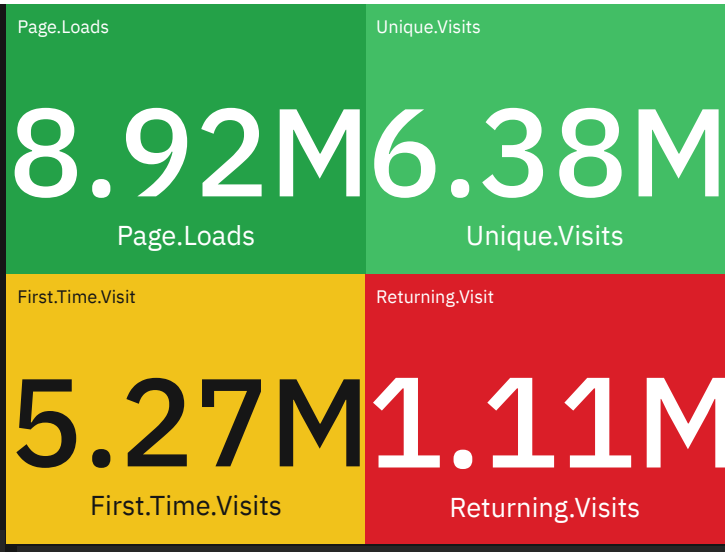
Row	Day	Day.Of.W	Date	Page.Load	Unique.Vi	First.Time	Returning.Visits
1	Sunday	1	9/14/2014	2,146	1,582	1,430	152
2	Monday	2	9/15/2014	3,621	2,528	2,297	231
3	Tuesday	3	9/16/2014	3,698	2,630	2,352	278
4	Wednesd	4	9/17/2014	3,667	2,614	2,327	287
5	Thursday	5	9/18/2014	3,316	2,366	2,130	236
6	Friday	6	9/19/2014	2,815	1,863	1,622	241
7	Saturday	7	9/20/2014	1,658	1,118	985	133
8	Sunday	1	9/21/2014	2,288	1,656	1,481	175
9	Monday	2	9/22/2014	3,638	2,586	2,312	274

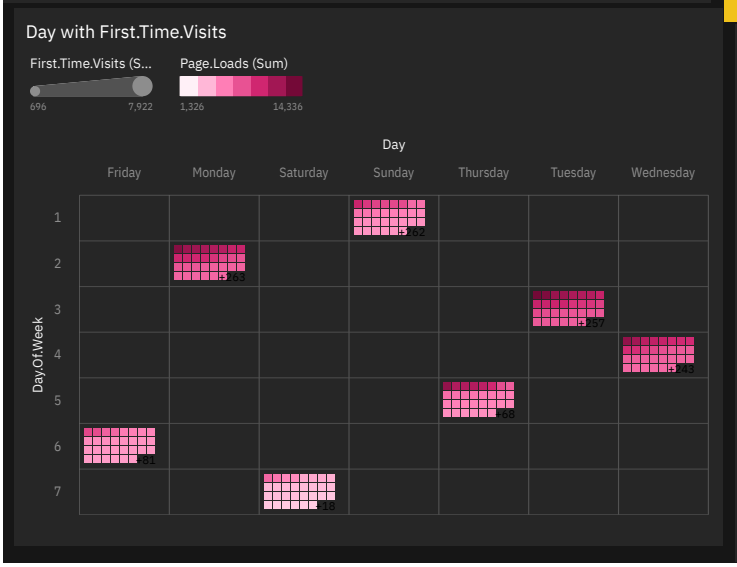
3.0 Analytics using IBM cognos tools

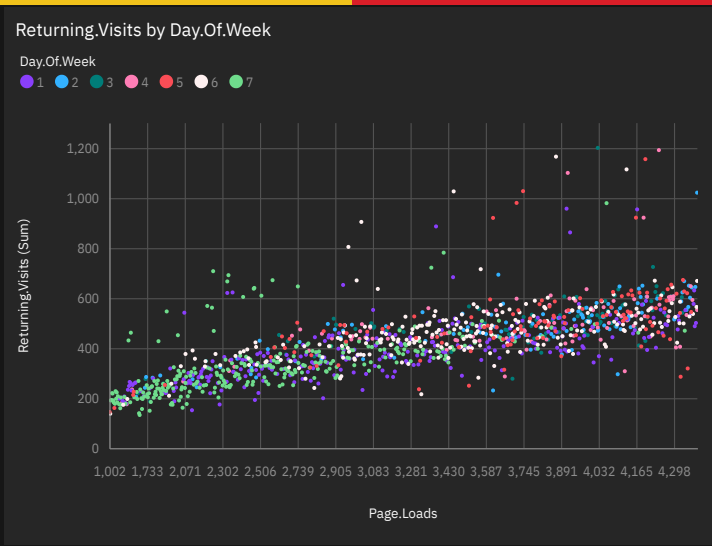
Tab 1

3.1 overall



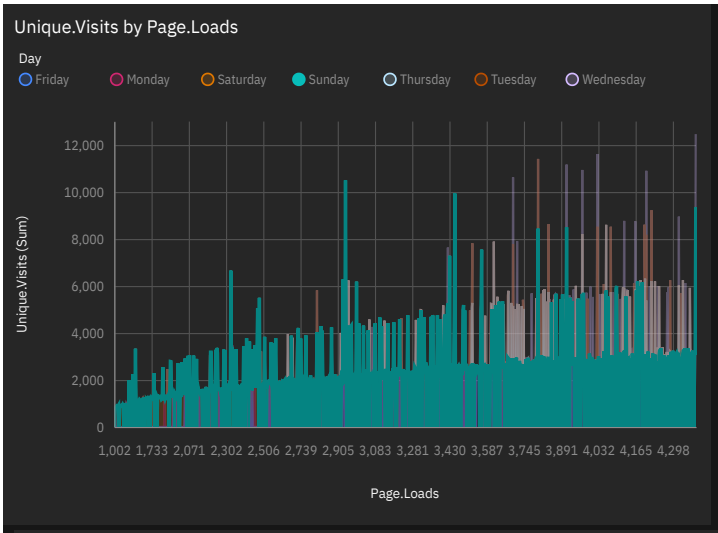


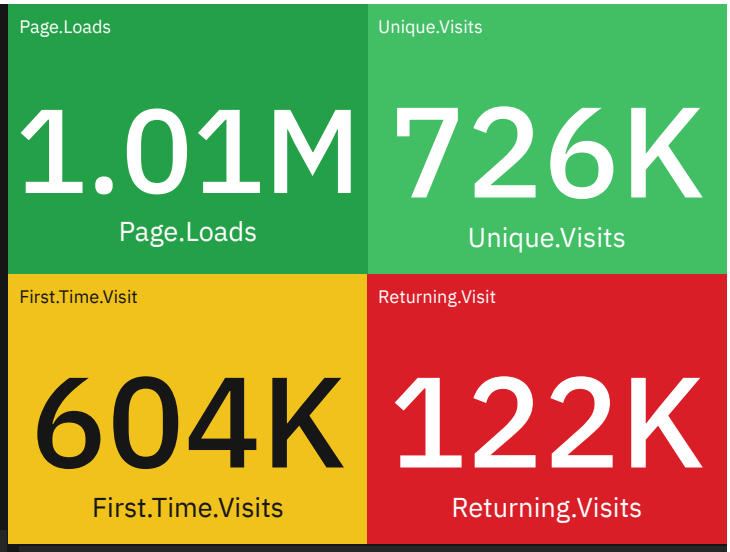


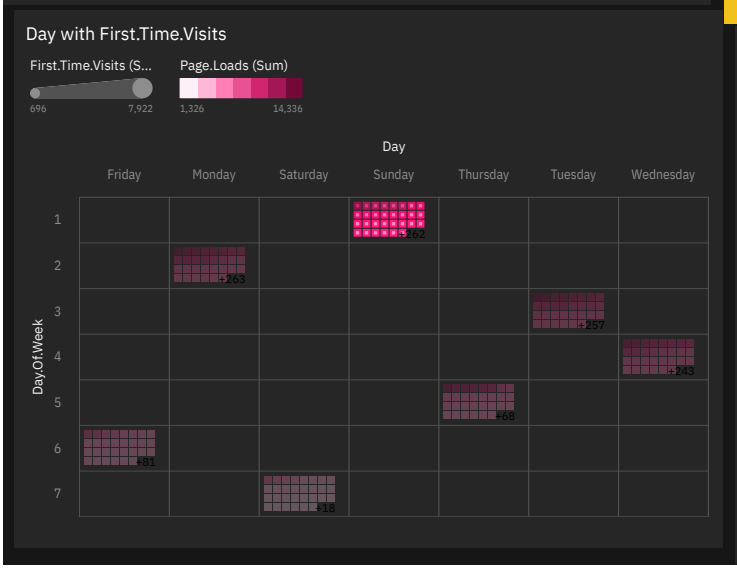


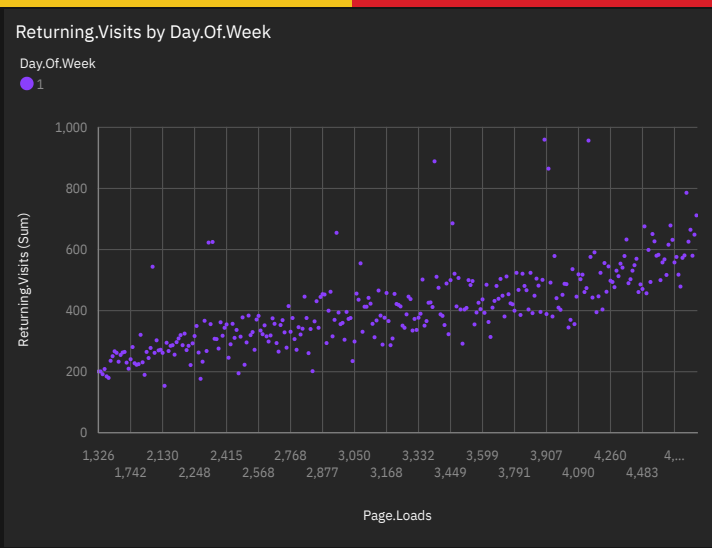
3.2 sunday visitors

Tab 1



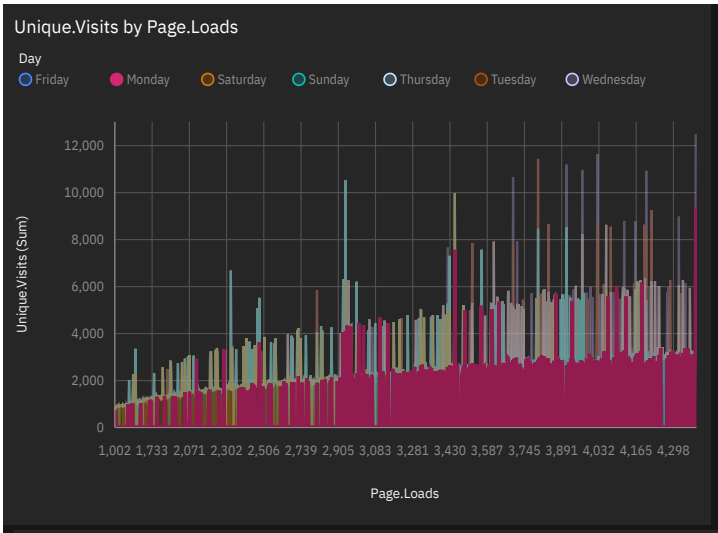


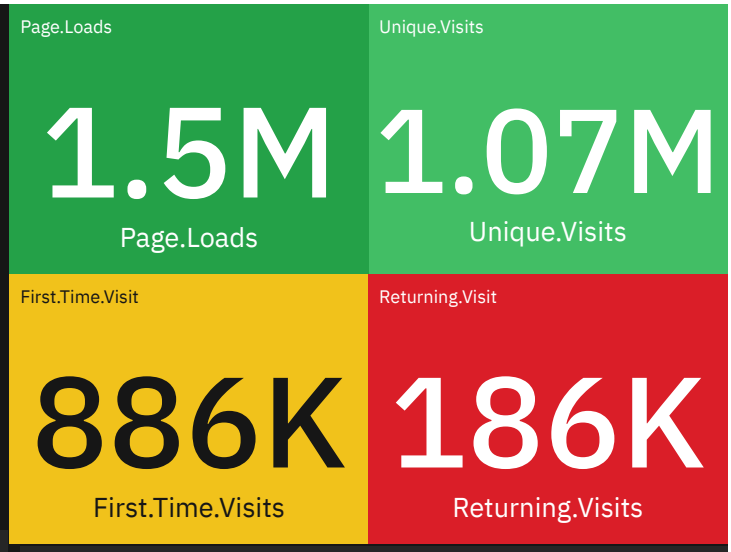


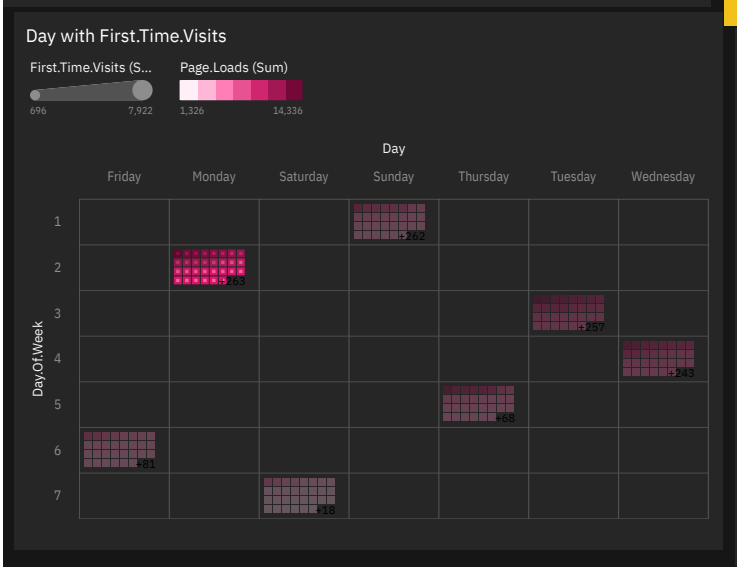


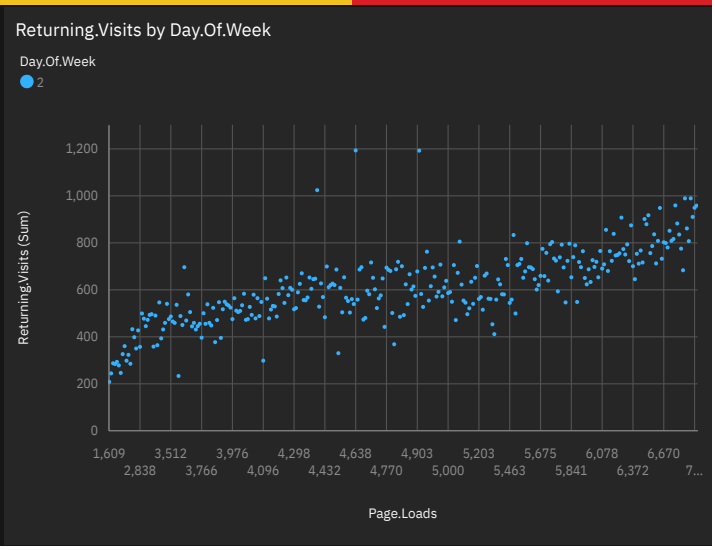
Tab 1

3.3 monday visitors



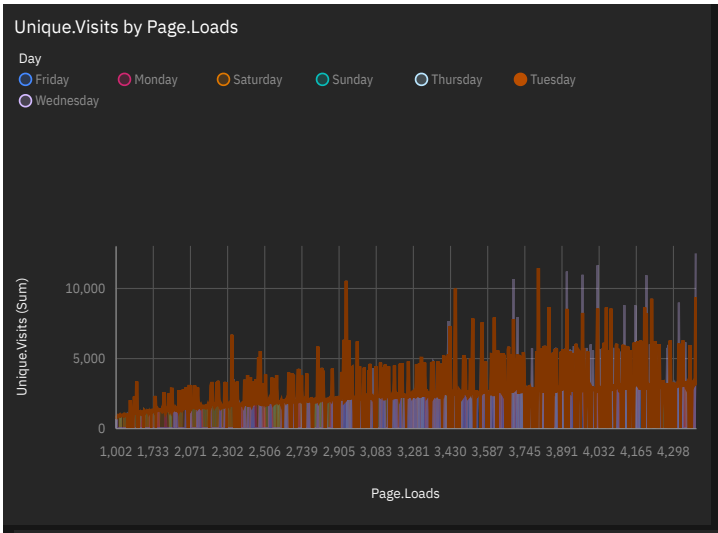


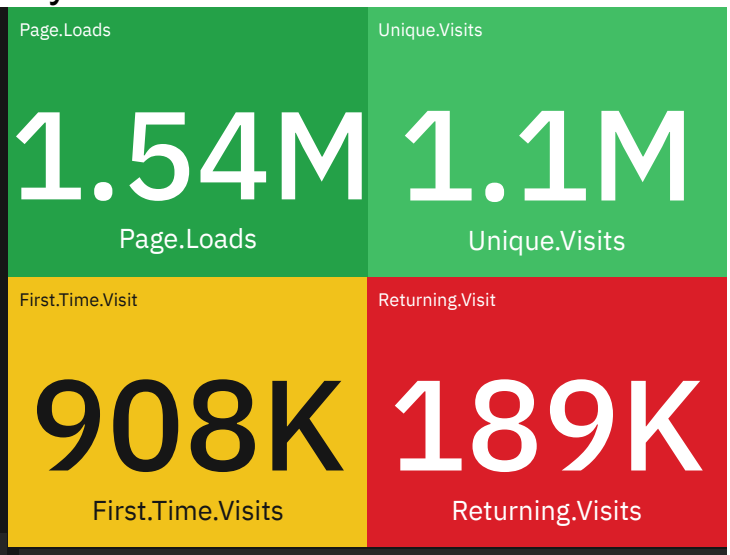


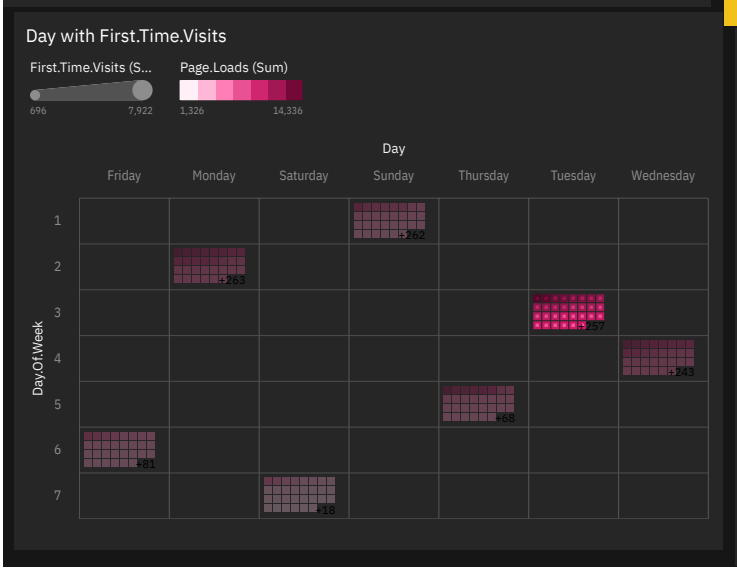


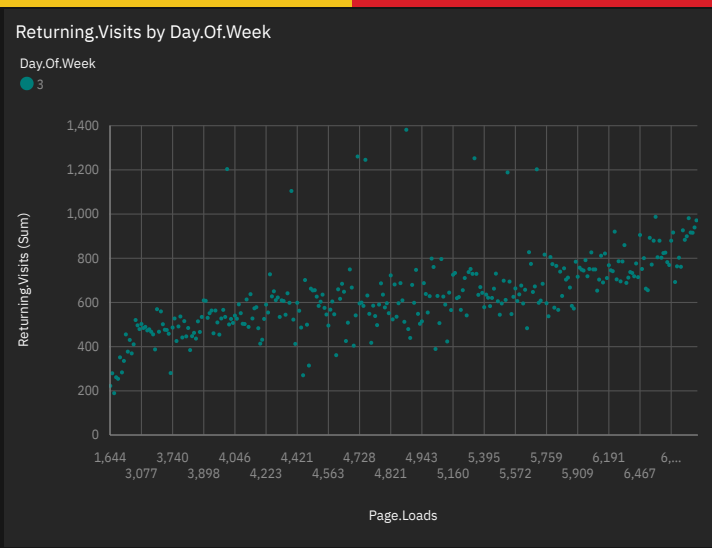
Tab 1

3.4 tuesday visitors



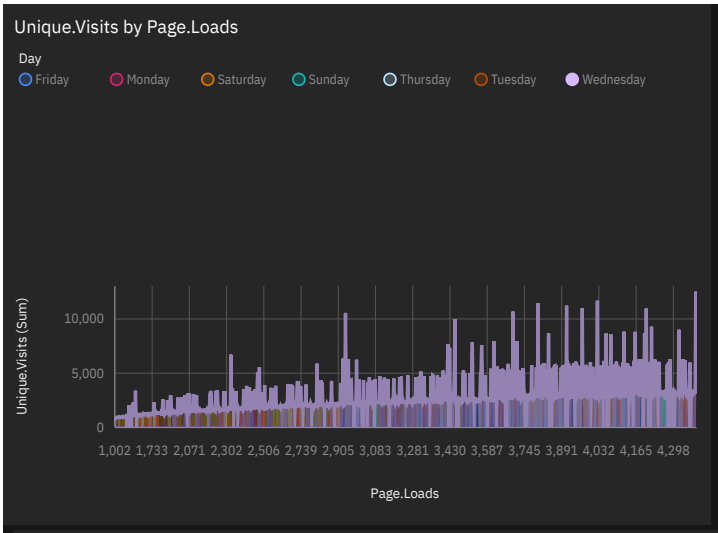


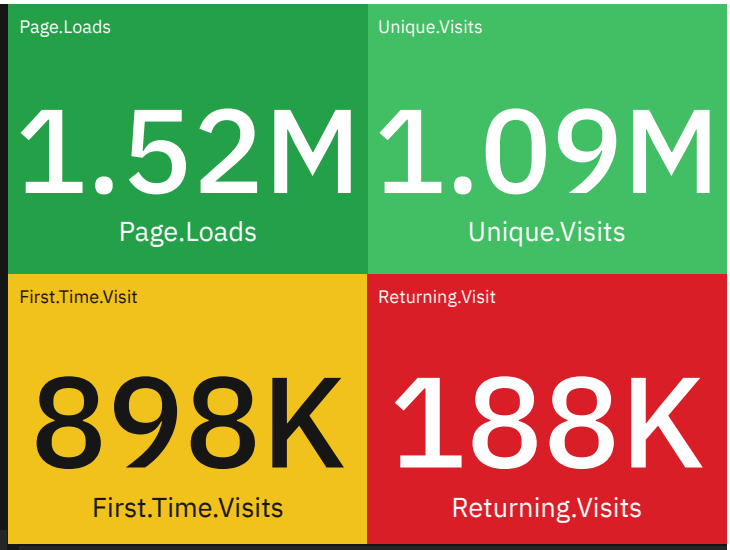


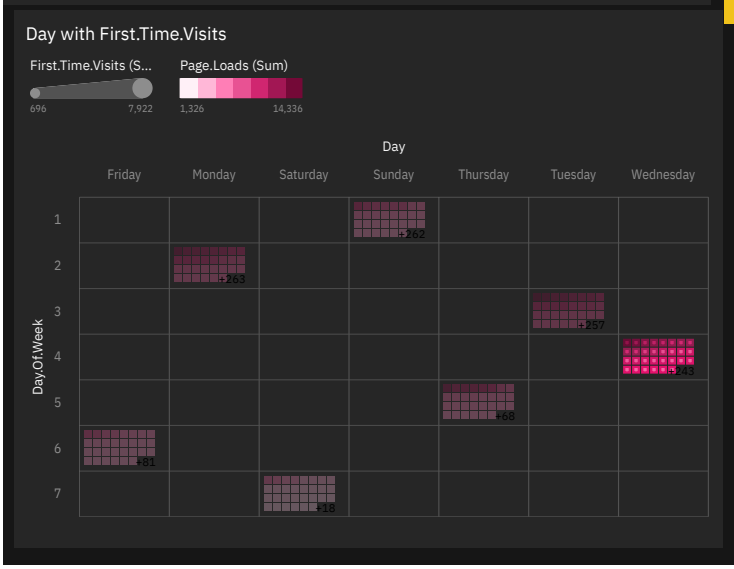


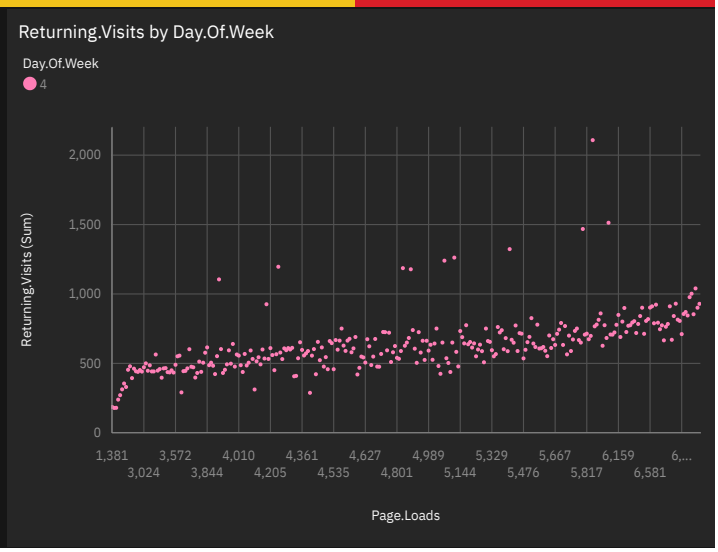
3.5 wednesday visitors

Tab 1



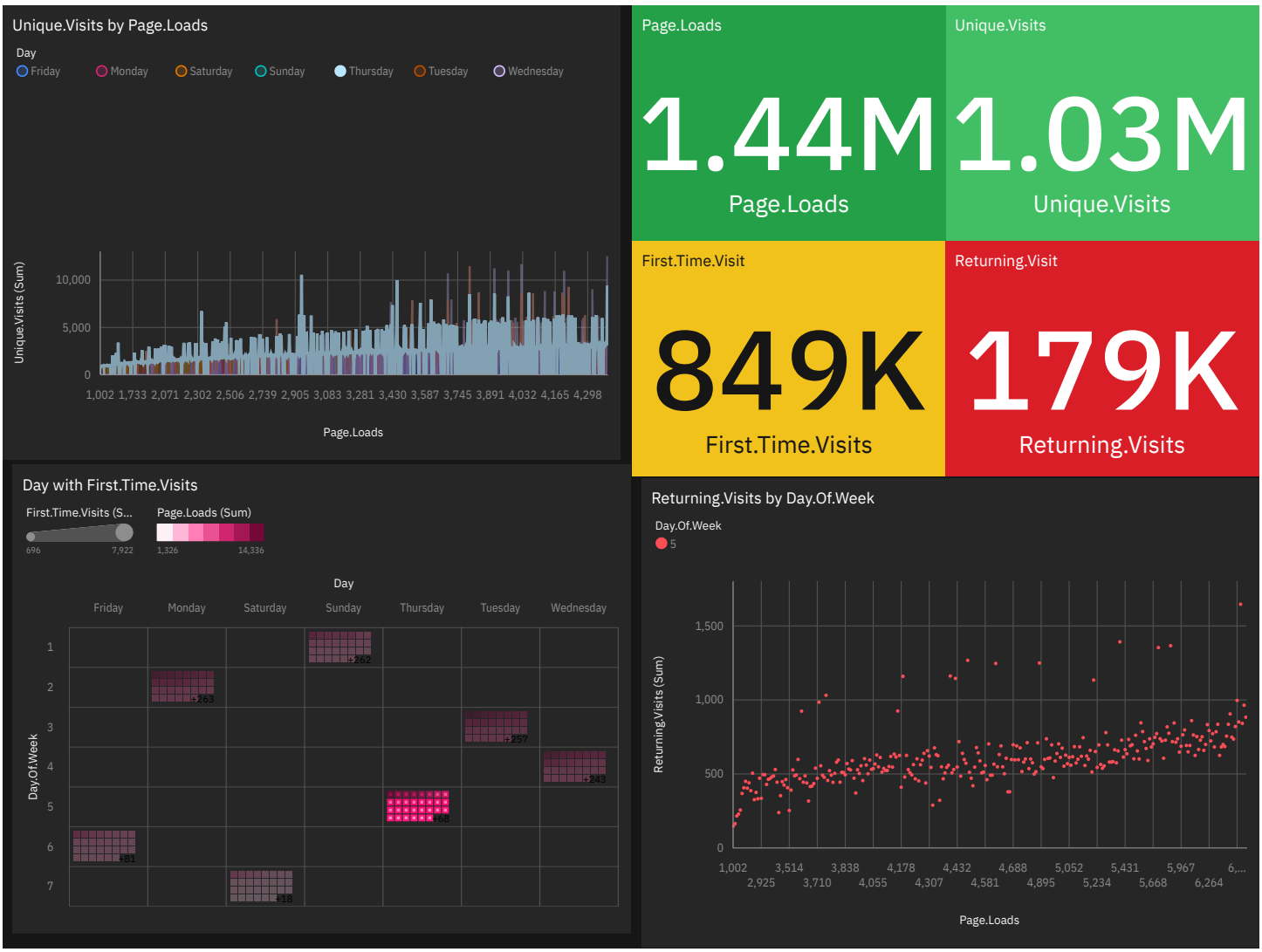






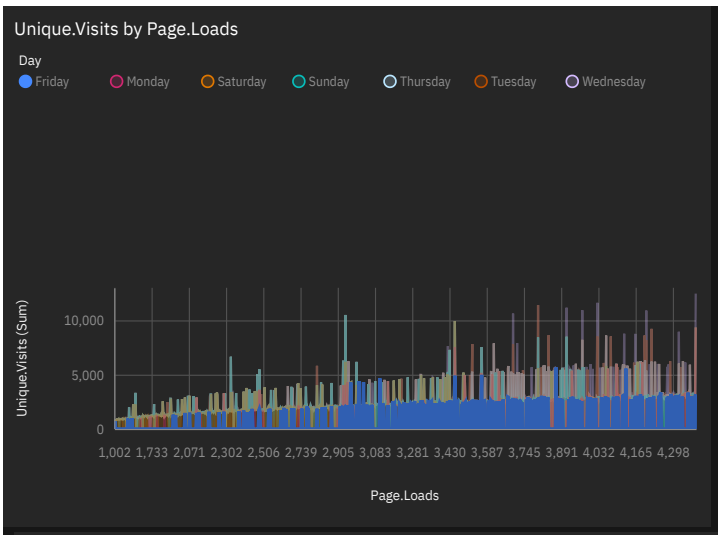
Tab 1

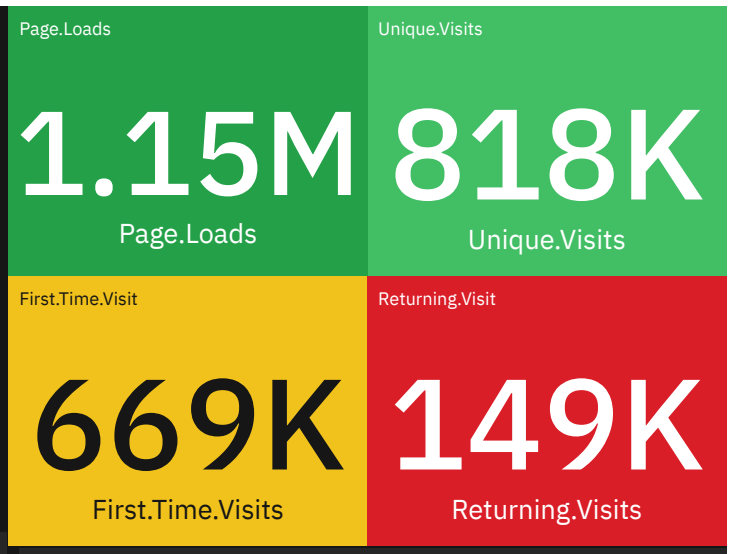
3.6 thursday visitors

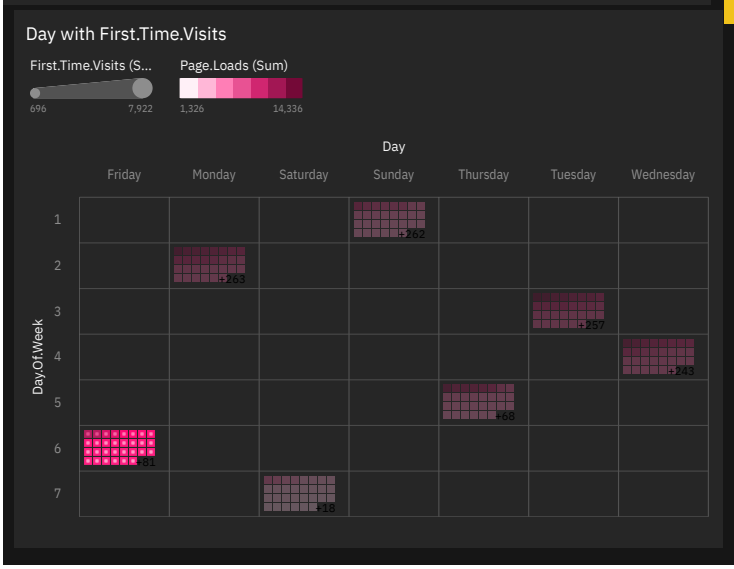


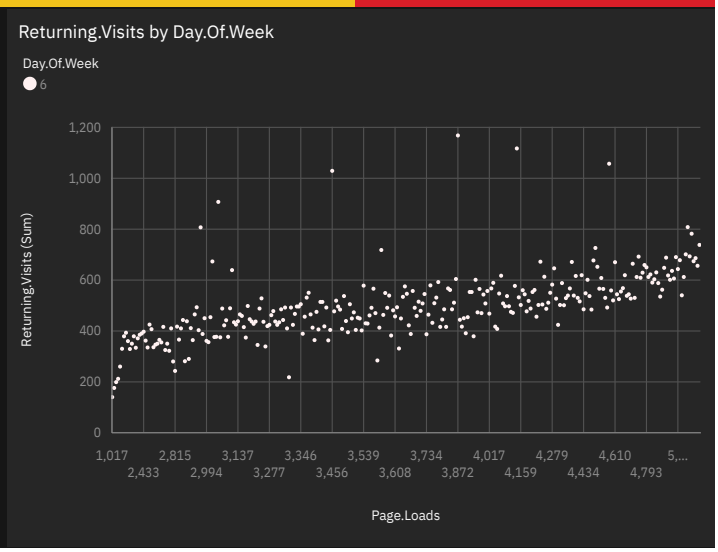
Tab 1

3.7 friday visitors



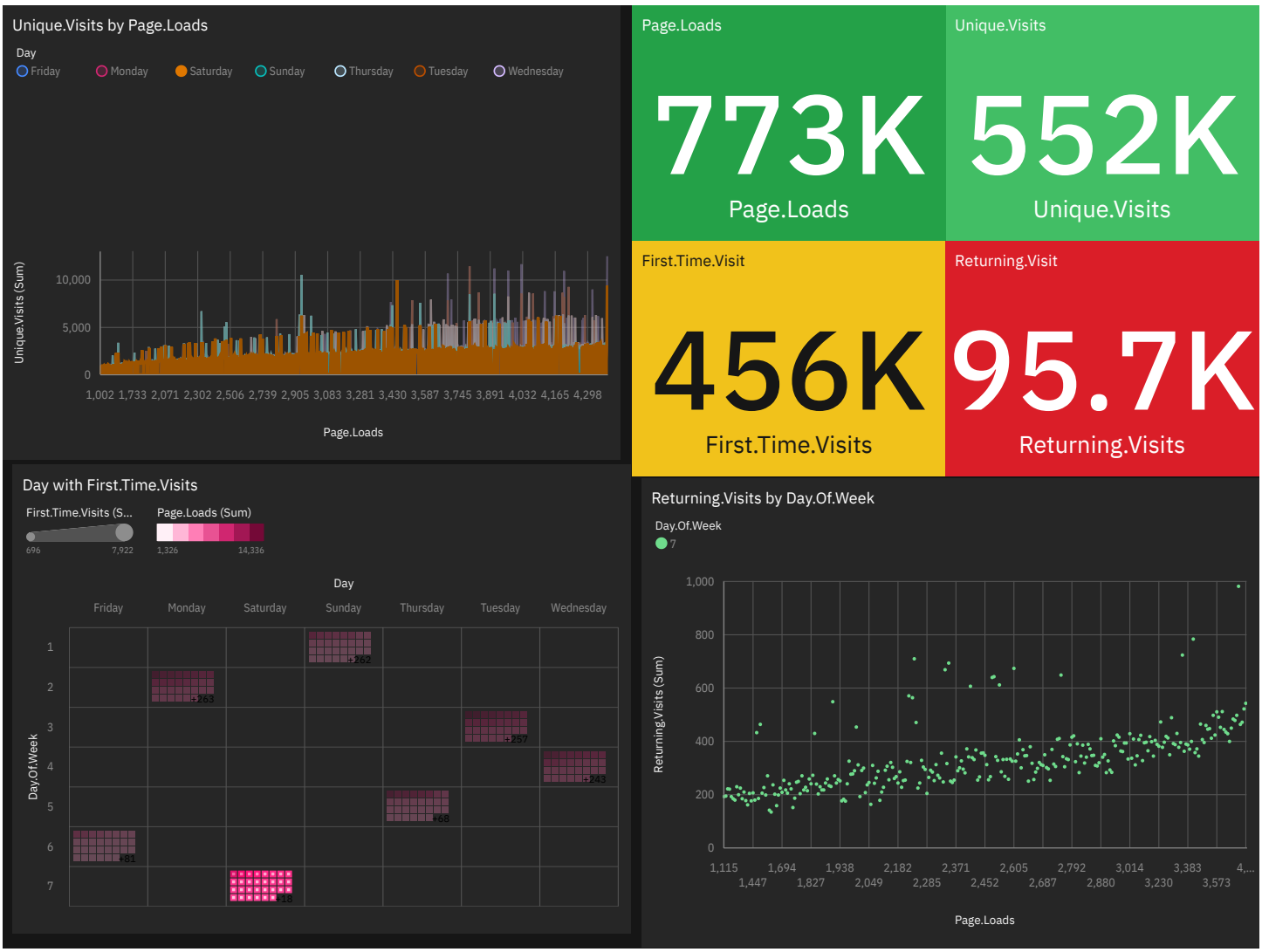




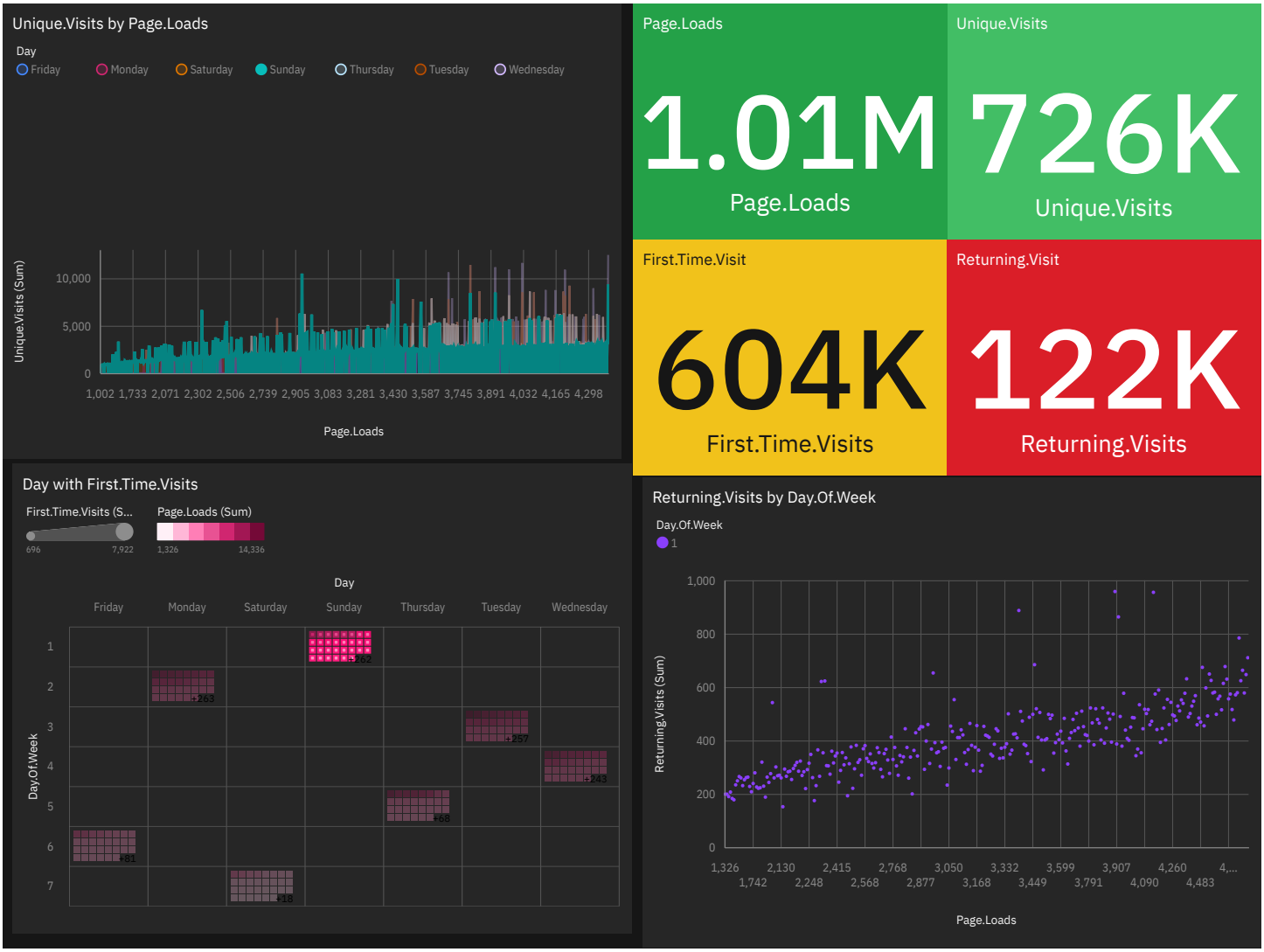


Tab 1

3.8 saturday visitors



Tab 1


Unique.Visits by Page.Loads
Day
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Unique.Visits (Sum)
10,000
5,000
0
1,002 1,733 2,071 2,302 2,506 2,739 2,905 3,083 3,281 3,430 3,587 3,745 3,891 4,032 4,165 4,298
Page.Loads
Page.Loads
1.01M
Page.Loads
Unique.Visits
726K
Unique.Visits
First.Time.Visit
604K
First.Time.Visits
Returning.Visit
122K
Returning.Visits
Day with First.Time.Visits
First.Time.Visits (S...
696
7,922
Page.Loads (Sum)
1,326
14,336
Day
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Day.Of.Week
1
2
3
4
5
6
7
Returning.Visits by Day.Of.Week
Day.Of.Week
1
Returning.Visits (Sum)
1,000
800
600
400
200
0
1,326 2,130 2,415 2,768 3,050 3,332 3,599 3,907 4,260
1,742 2,248 2,568 2,877 3,168 3,449 3,791 4,090 4,483
Page.Loads

4.0 Insights for above **“IBM Cognos Analytics”**

•unique visits:

Unique.Visits is unusually low when Day is Saturday.

Based on the current forecasting, Unique.Visits may reach almost 481 thousand by Day Monday+1.

It is projected that by Monday+1, 4205 will exceed 3973 in Unique.Visits by almost 1500.

Page.Loads 4376 has the highest Total Returning.Visits but is ranked #5 in Total Unique.Visits.

Page.Loads 4638 has the highest Total Unique.Visits but is ranked #3 in Total Returning.Visits.

Over all values of Page.Loads and Day, the sum of Unique.Visits is almost 6.4 million.

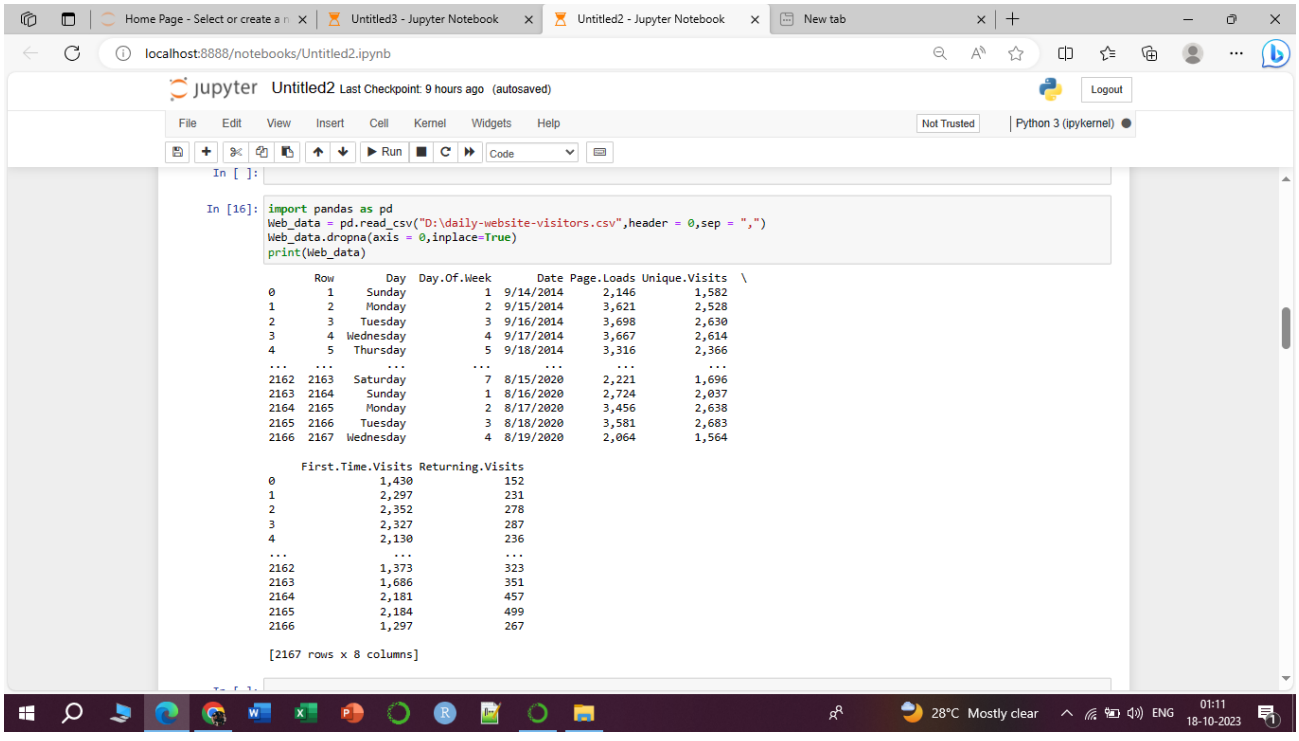
The summed values of Unique.Visits range from 667 to nearly 13 thousand.

For Unique.Visits, the most significant values of Day are Tuesday, Wednesday, Monday, Thursday, and Friday, whose respective Unique.Visits values add up to over 5.1 million, or 80 % of the total.

For Unique.Visits, the most significant value of Page.Loads is 4638, whose respective Unique.Visits values add up to over thirteen thousand, or 0.2 % of the total.

5.0 'Python Integration' for Website Traffic Analysis

Extraction the Data file from the directory to the python text editor to excecute the data set .



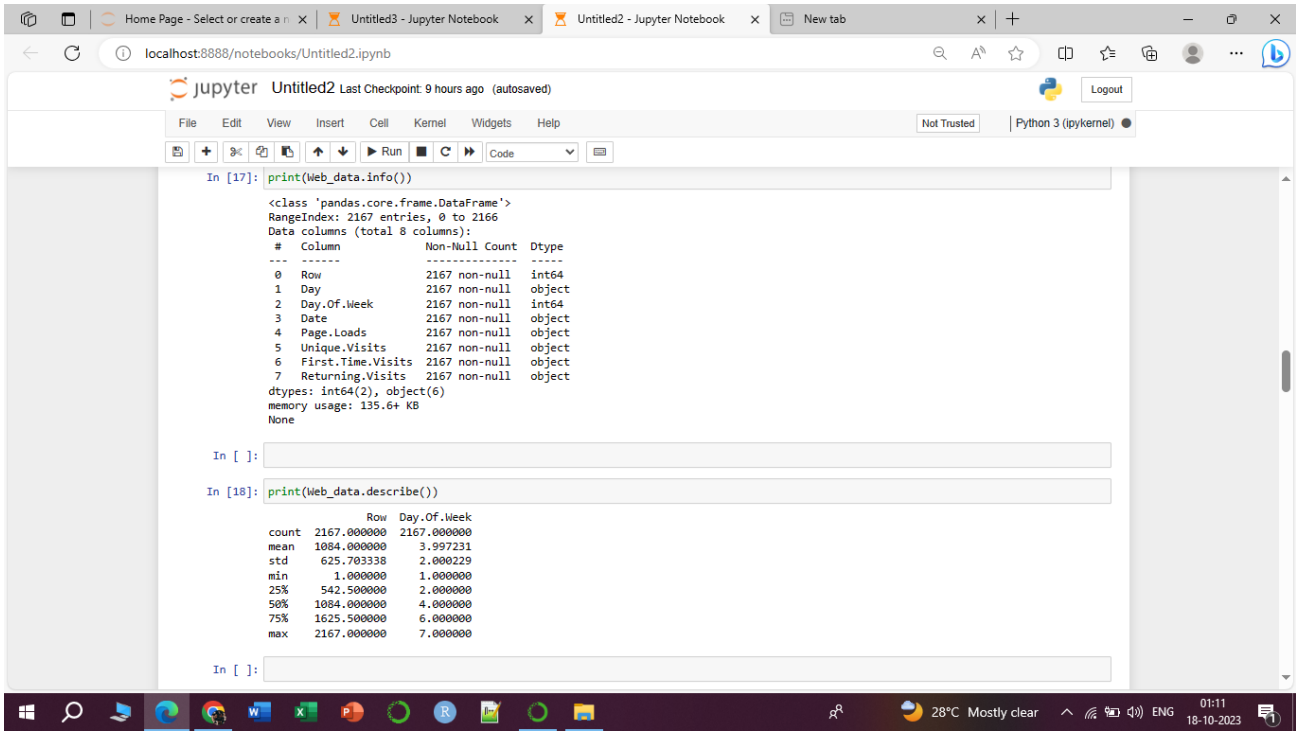
```
In [16]: import pandas as pd
         Web_data = pd.read_csv("D:\daily-website-visitors.csv",header = 0,sep = ",")
         Web_data.dropna(axis = 0,inplace=True)
         print(Web_data)
```

```
       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
```

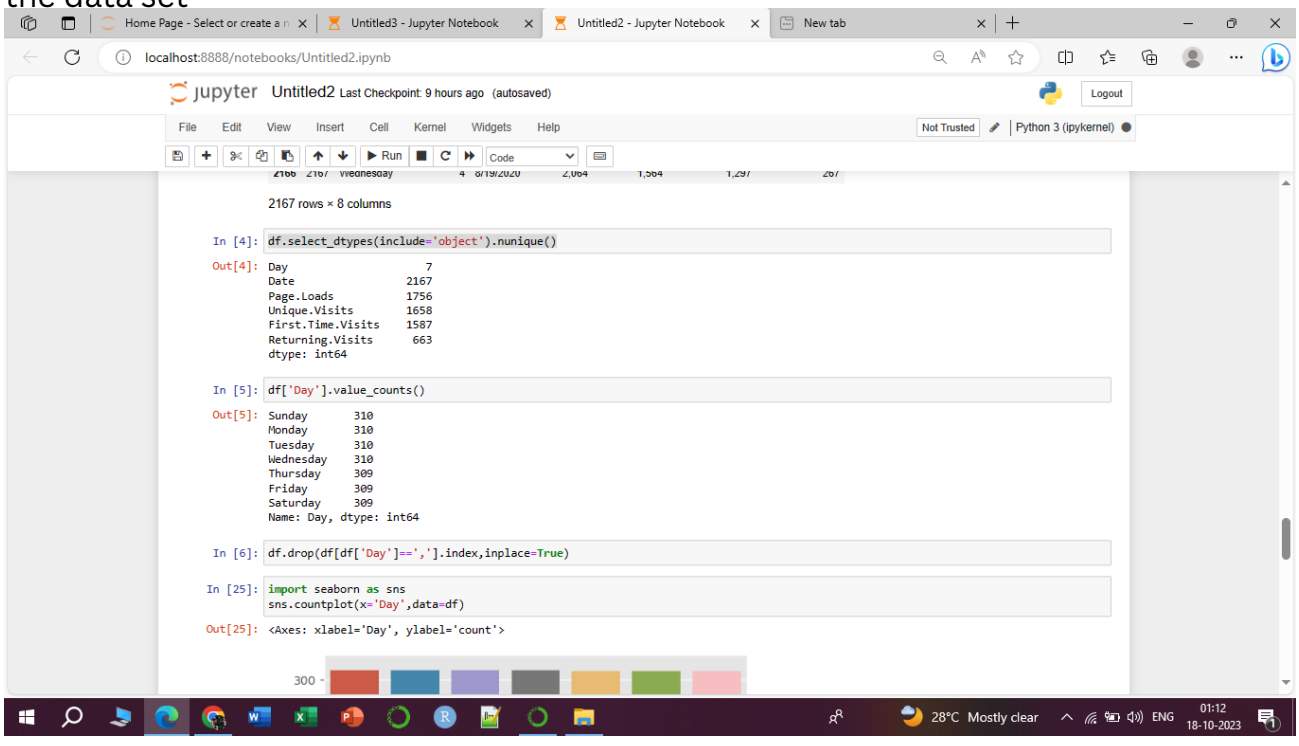
```
In [ ]:
```

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

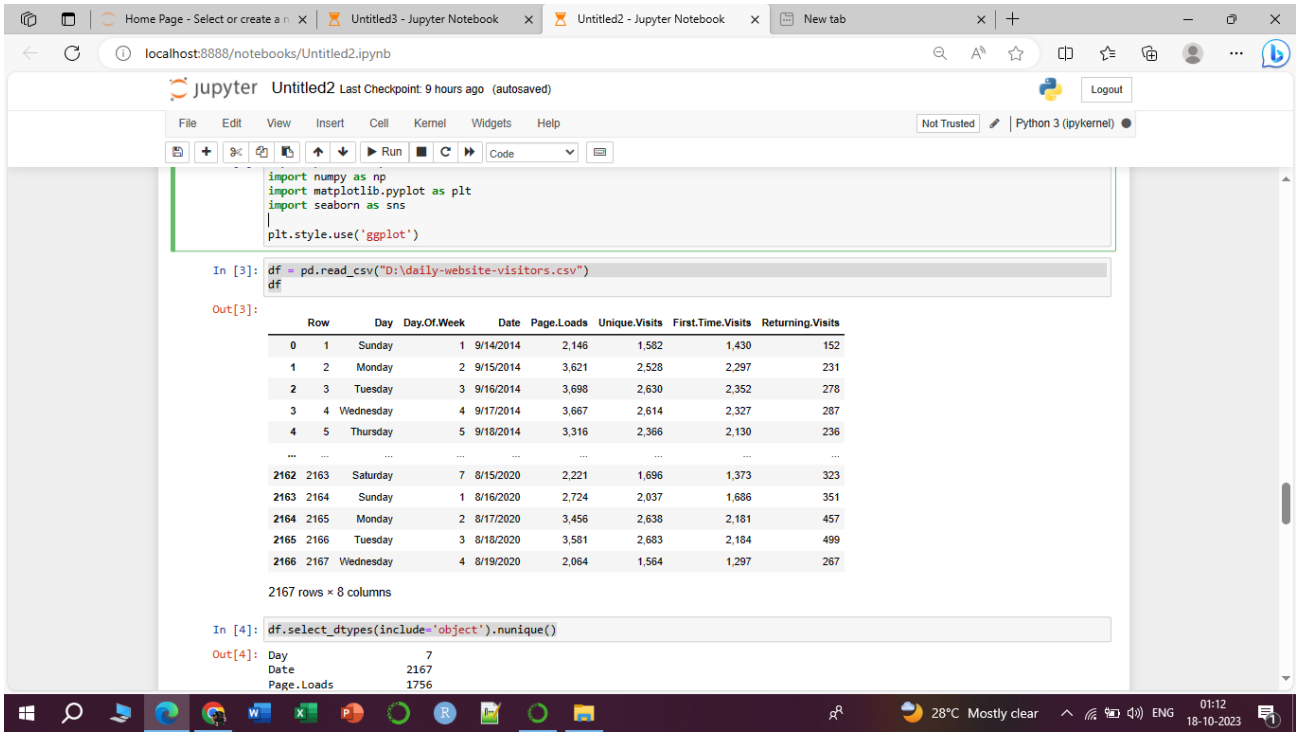
```
                Row  Day.Of.Week
count  2167.000000  2167.000000
mean   1084.000000     3.997231
std     625.703338     2.000229
min       1.000000     1.000000
25%     542.500000     2.000000
50%    1084.000000     4.000000
75%    1625.500000     6.000000
max    2167.000000     7.000000
```

```
In [ ]:
```

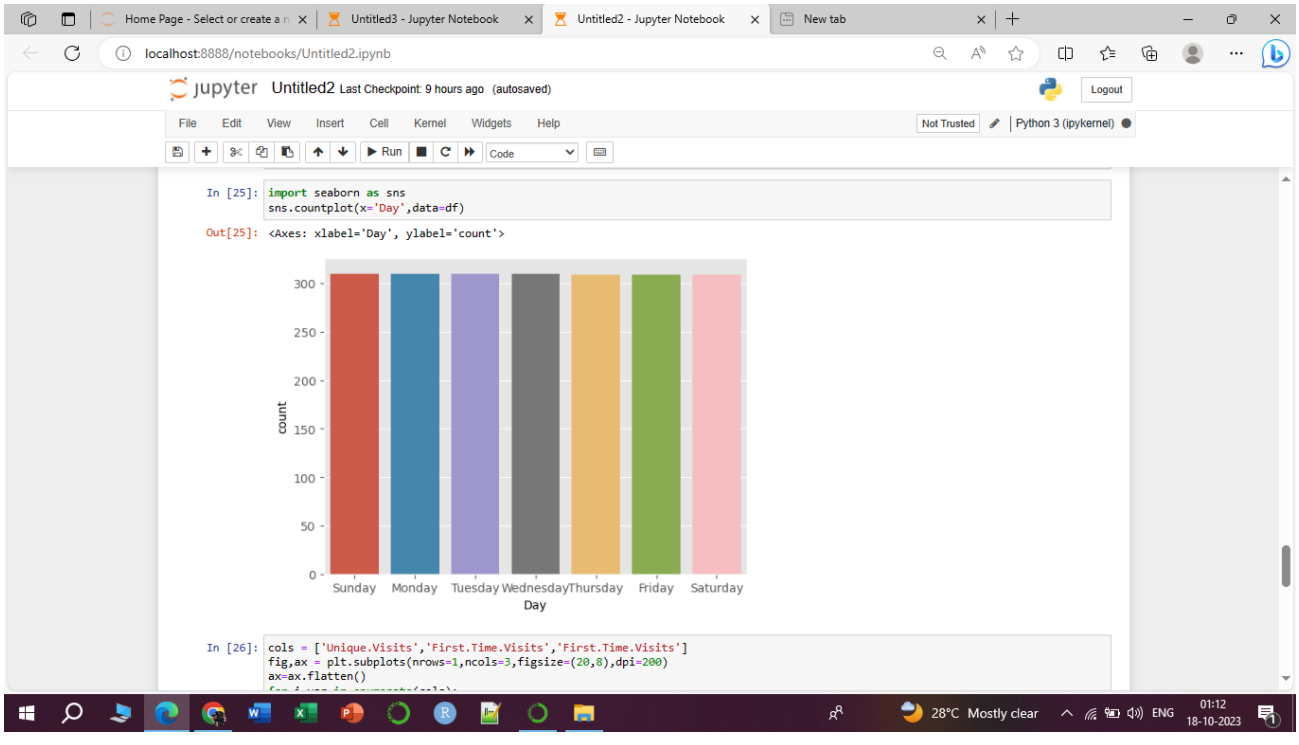
Value Counts of each Insights of the data set content and Object Integration of the data set



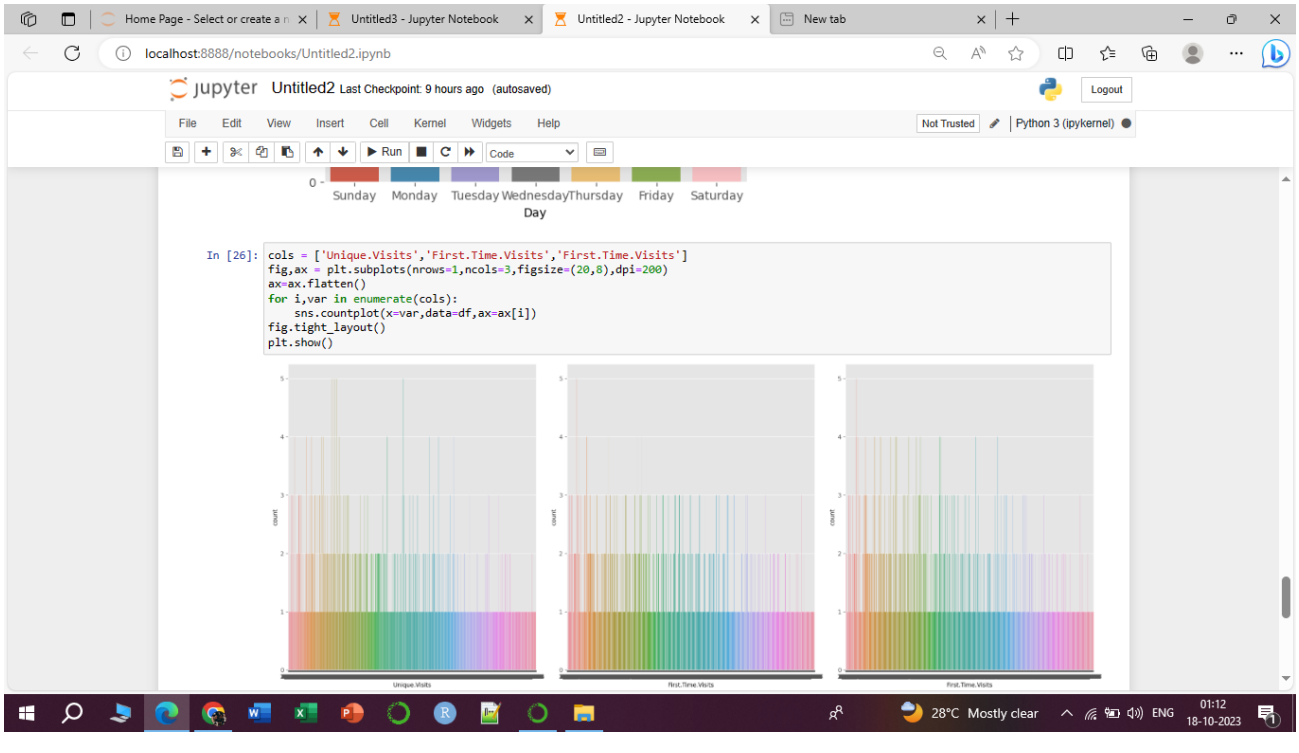
Again pointing out the data set with the help of Pandas Library.

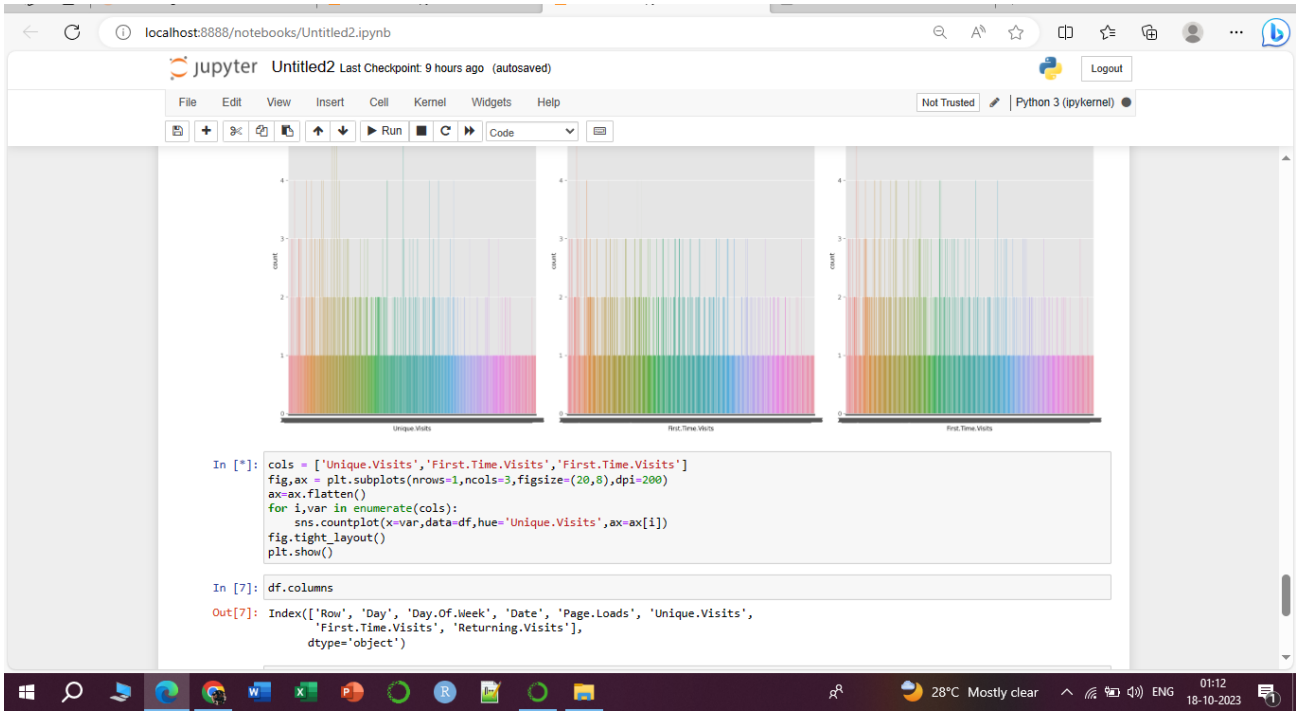


Listing the Number of Days that had been observed by the data set using Seaborn Library.



Plotting of Three important Stuffs like Unique visitor,First Time Visitor,Returning Visitor.





6.0 conclusion and future process:

Thus in this documentation phase1 we have visualized the given data overall and day wise also we have integrated python in it in further process our team has planned to use ibm cognos tools much more deeper and make a document out of the same