DATA ANALYTICS WITH COGNOS

WEBSITE TRAFFIC ANALYSIS – PHASE 3

INTRODUCTION:

In Phase 3, the given dataset of the website traffic has been loaded to the Cognos Analytics tool. Then, the visualization of the data has been done by creating dashboard and data module using the Cognos tool.

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

ABOUT DATASET:

The given data set has already been cleaned. Hence, no preprocessing is required.

➢ **Row**: An identifier or index for each row in the dataset used to identify each record uniquely.
➢ **Day**: Represents the day of the data record.
➢ **Day_of_week**: Indicates the day of the week represented as a numeric value (e.g., 1 for Sunday, 2 for Monday, etc.).
➢ **Date**: Denotes date of the entry where the entire data ranges over a period of 5 years.
➢ **Page Loads**: Number of page loads on the website for a given day. It's a count of how many times pages on the website were loaded or accessed.
➢ **Unique Visits**: Count of unique visitors to the website on a given day. It measures the number of distinct individuals whose IP addresses haven't been hits on any page in over 6 hours. A visit is classified as "unique" if a hit from same IP address has not come within the last 6 hours.
➢ **First-time Visits**: Indicates the number of first-time visitors to the website on a particular day. First-time visits typically represent new visitors who haven't visited the website before.
➢ **Returning Visits**: Count of returning visitors to the website on a specific day. Returning visits are from individuals who have visited the website before and are returning for subsequent visits.

ANALYSIS PURPOSE:

• **Analyze Traffic Patterns:** Allows to understand how website traffic varies on a daily basis, with insights into daily, weekly, and possibly academic calendar-related fluctuations.
• **Forecasting:** The data can be used to build forecasting models to predict future website traffic. For example, you can develop models for 1-day-ahead forecasts, 7-day-ahead forecasts, and forecasts for the entire next week.
• **Visitor Segmentation:** By distinguishing between unique, first-time, and returning visitors, you can gain insights into the behavior of different types of website users.
• **Identify Trends and Seasonal Effects:** These data sets help in identifying trends and seasonality in website traffic, which can inform content strategies, marketing efforts, and resource allocation.

VISUALIZATION USING COGNOS:

1.PAGE LOADS AND UNIQUE VISITS BY DATE AND DAY :

Page.Loads and Unique.Visits by Date and Day
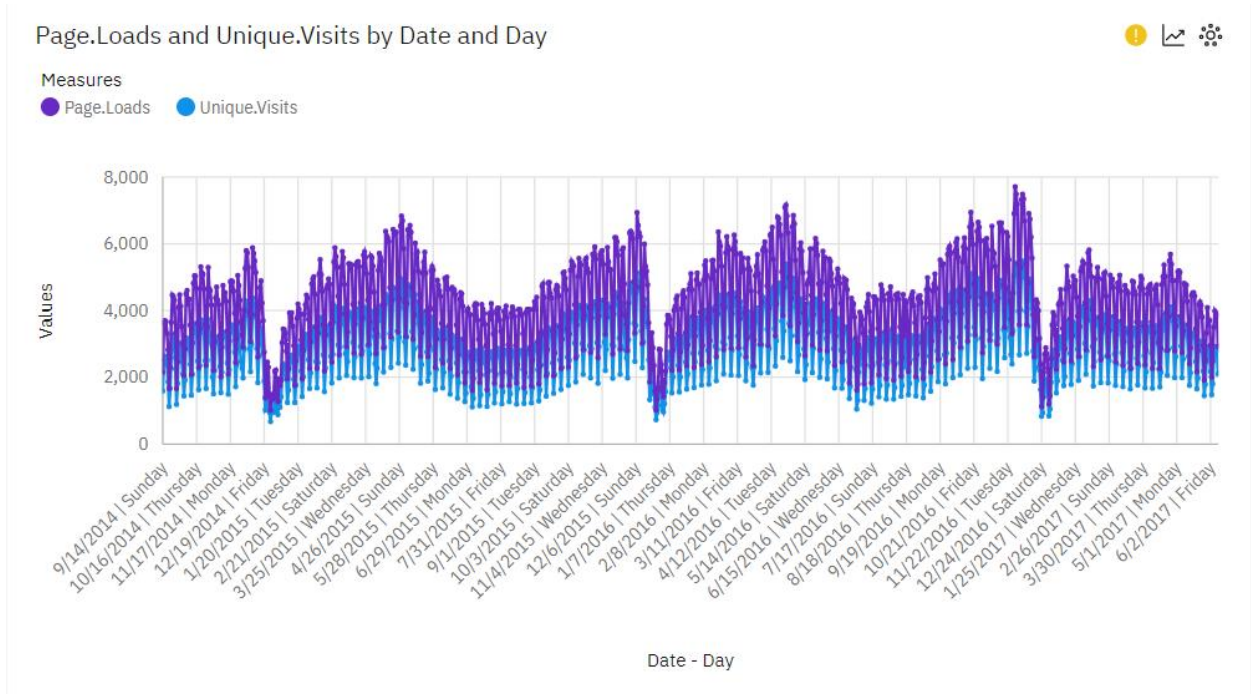
INSIGHTS:

- Based on the current forecasting, Page Loads may reach nearly four thousand by Date-Day 2021-10-27-Wednesday.
- Over all date - days, the average of Page Loads is 2.997.
- Across all date - days, the average of Unique Visits is over four thousand.
- The total number of results for Page Loads, across all date - days, is over two thousand.
- The total number of results for Unique Visits, across all date - days, is over two thousand.
- Page Loads ranges from over a thousand, when Date - Day is 2014-12-25|Thursday, to nearly eight thousand, when Date - Day is 2018-04-25|Wednesday.
- Unique Visits ranges from 667, when Date - Day is 2014-12-25|Thursday, to over 5500, when Date - Day is 2018-04-25|Wednesday.

2.ROW BY PAGE LOADS AND FIRST TIME VISITS :

Row by Page.Loads and First.Time.Visits colored by Date
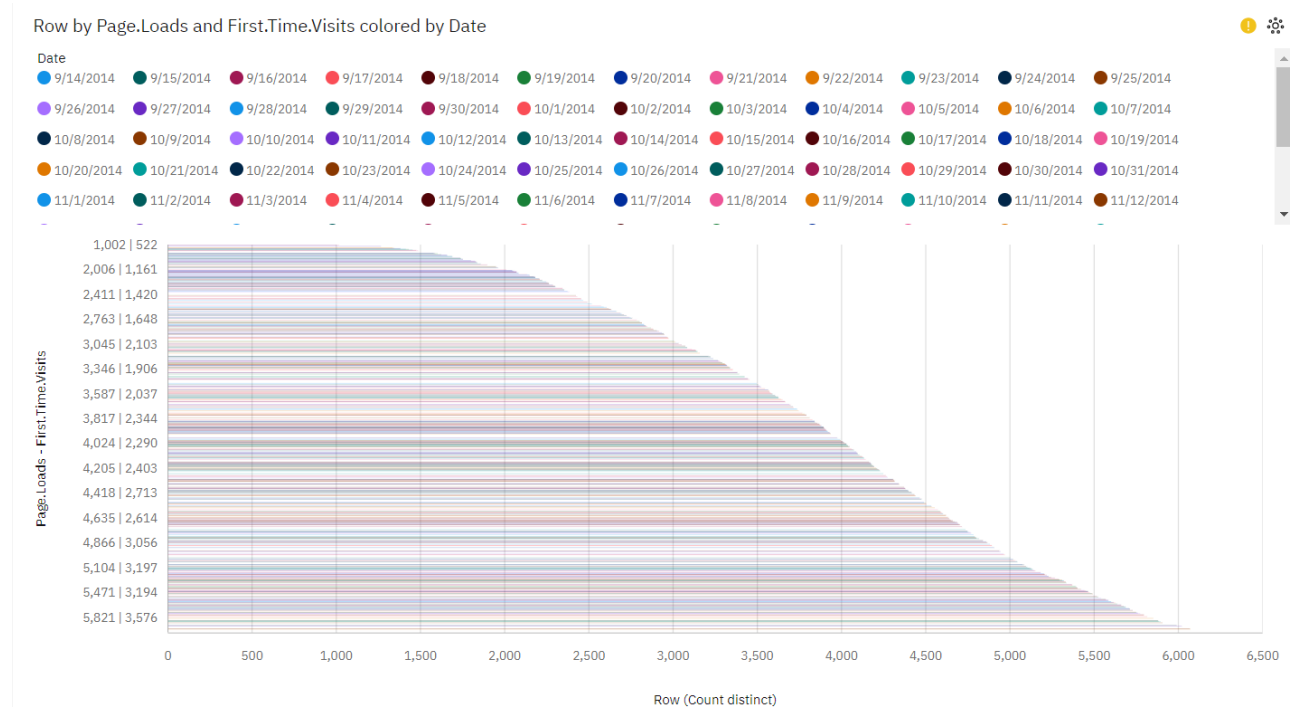
INSIGHTS:

- Day Tuesday has the highest Page Loads at almost 365 thousand, out of which First Time Visits 4369 contributed the most at nearly 2 thousand.
- 4616 has a Page Loads of nearly two thousand for Day Wednesday.
- First Time Visits 3146 Page Loads from Day Monday is almost 1500, whereas 3133 is only over a thousand.
- First Time Visits 3146 has the highest Page Loads due to Day Monday.
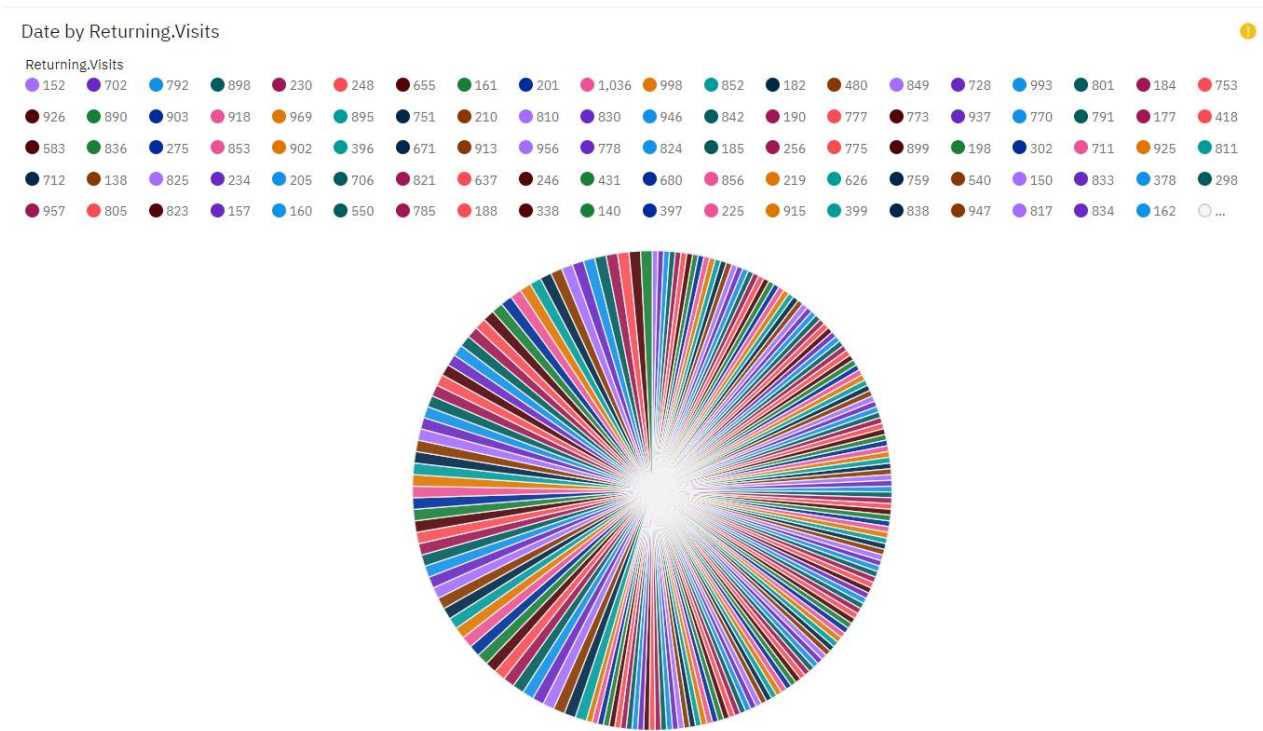- The overall number of results for Row is over two thousand.

3.PAGE LOADS AND UNIQUE VISITS :

Page.Loads and Unique.Visits by Date

Measures
● Page.Loads ● Unique.Visits
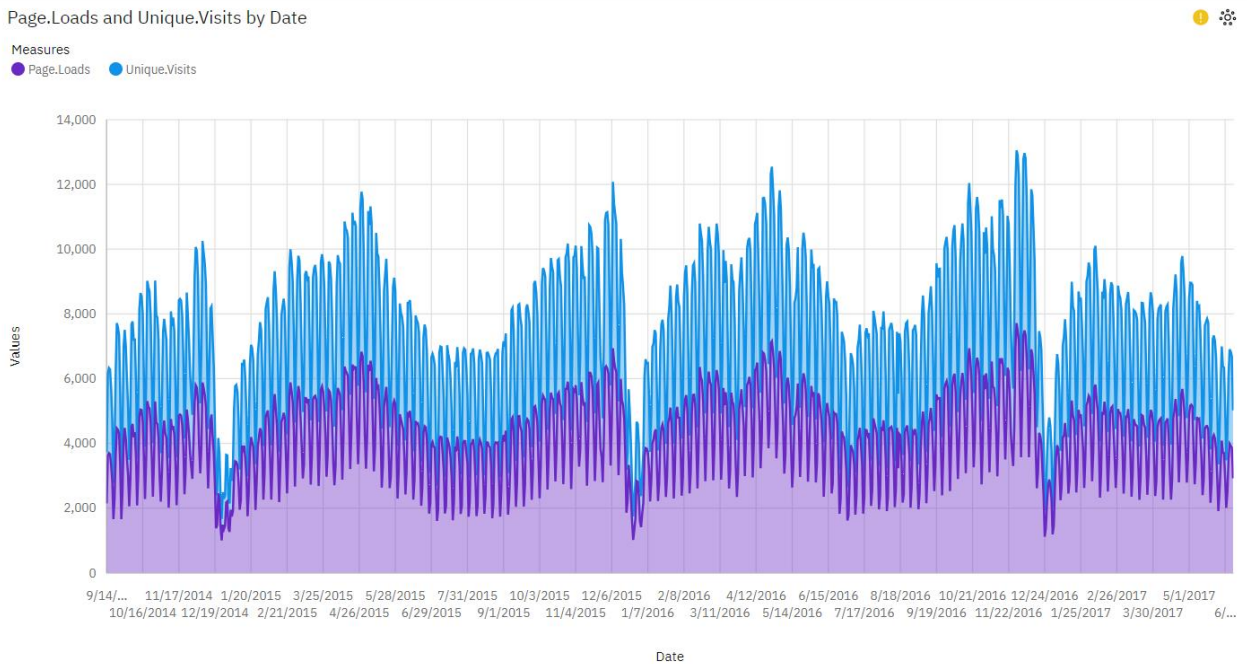
INSIGHTS:

- Page Loads has a strong weekly trend. The largest values typically occur on Tuesday, whereas the smallest values on Saturday.
- Based on the current forecasting, Page Loads may reach nearly four thousand by Date 2021-10-27.
- The value of Page Loads at the last observed time point 2020-08-19 is unusual. This may indicate incomplete data or a recent event that might require investigation.
- Page Loads has most unusual values at 24 time points, the most notable of which are 2020-03-16, 2019-01-03, 2020-08-19, 2016-11-24, and 2019-01-02.
- Across all dates, the average of Page Loads is over four thousand.
- Over all dates, the average of Unique Visits is nearly three thousand.
- The total number of results for Page Loads, across all dates, is over two thousand.
- The total number of results for Unique Visits, across all dates, is over two thousand.
- Page Loads ranges from over a thousand, when Date is 2014-12-25, to nearly eight thousand, when Date is 2018-04-25.
- Unique Visits ranges from 667, when Date is 2014-12-25, to over 5500, when Date is 2018-04-25.
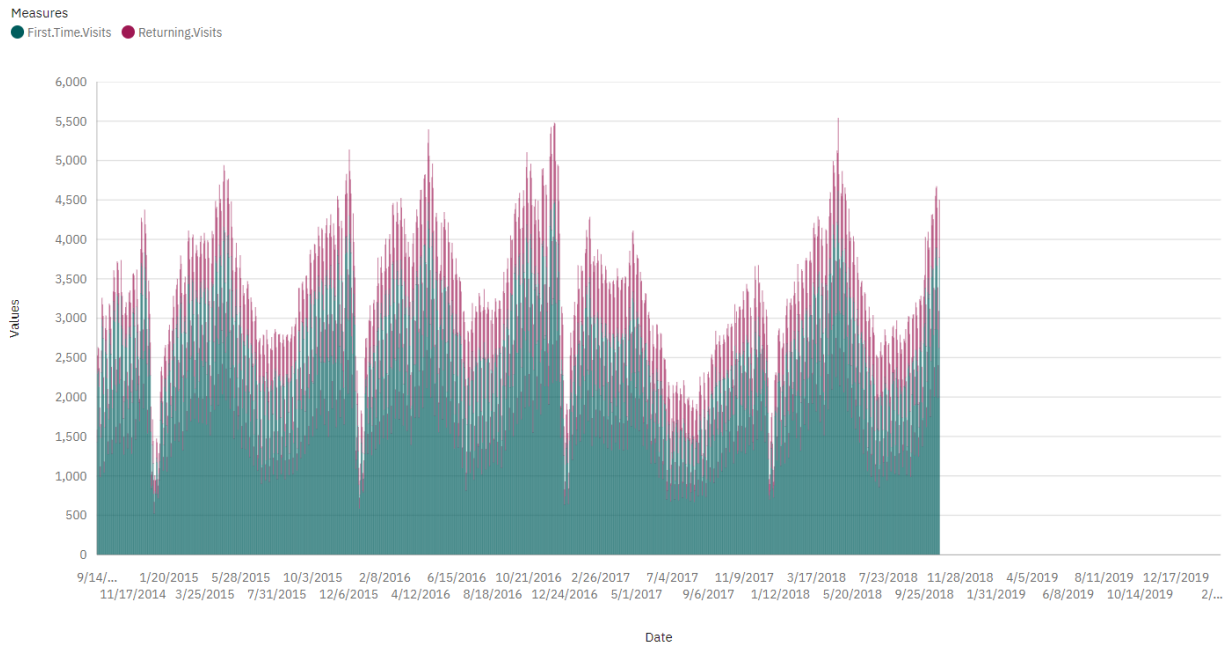
4.DATE BY RETURNING VISITS :

Date by Returning.Visits

Returning.Visits

152	702	792	898	230	248	655	161	201	1,036	998	852	182	480	849	728	993	801	184	753
926	890	903	918	969	895	751	210	810	830	946	842	190	777	773	937	770	791	177	418
583	836	275	853	902	396	671	913	956	778	824	185	256	775	899	198	302	711	925	811
712	138	825	234	205	706	821	637	246	431	680	856	219	626	759	540	150	833	378	298
957	805	823	157	160	550	785	188	338	140	397	225	915	399	838	947	817	834	162	...

INSIGHTS:

- The total number of results for Date, across all Returning Visits, is over two thousand.

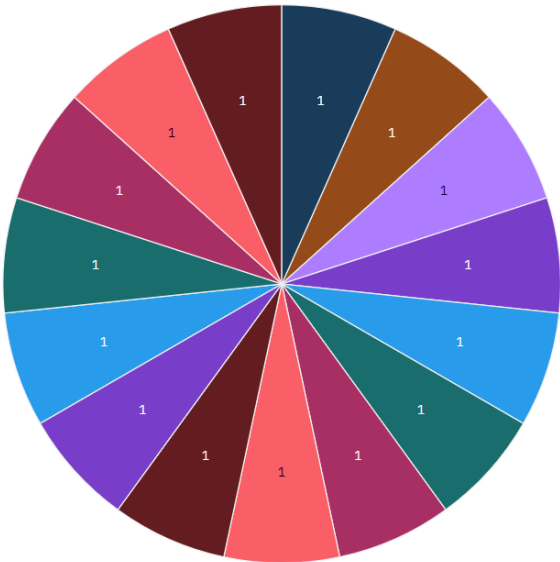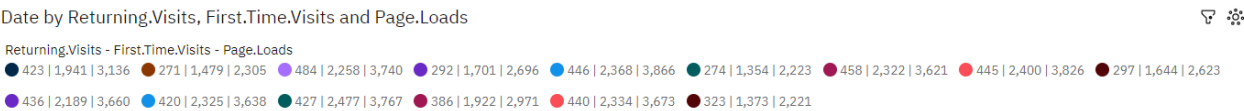**5.FIRST TIME VISITS AND RETURNING VISITS BY DATE :**

First.Time.Visits and Returning.Visits by Date

Measures
● First.Time.Visits ● Returning.Visits

INSIGHTS:

- Over all dates, the average of First Time Visits is almost 2500.
- Over all dates, the average of Returning Visits is 511.8.
- The total number of results for First Time Visits, across all dates, is over two thousand.
- The total number of results for Returning Visits, across all dates, is over two thousand.
- First Time Visits ranges from 522, when Date is 2014-12-25, to over 4500, when Date is 2018-04-25.
- Returning Visits ranges from 133, when Date is 2014-09-20, to over a thousand, when Date is 2016-11-30.

**6.DATE BY RETURNING VISITS, FIRST TIME VISITS AND PAGE LOADS :**

Date by Returning.Visits, First.Time.Visits and Page.Loads

Returning.Visits - First.Time.Visits - Page.Loads

● 423 | 1,941 | 3,136 ● 271 | 1,479 | 2,305 ● 484 | 2,258 | 3,740 ● 292 | 1,701 | 2,696 ● 446 | 2,368 | 3,866 ● 274 | 1,354 | 2,223 ● 458 | 2,322 | 3,621 ● 445 | 2,400 | 3,826 ● 297 | 1,644 | 2,623
● 436 | 2,189 | 3,660 ● 420 | 2,325 | 3,638 ● 427 | 2,477 | 3,767 ● 386 | 1,922 | 2,971 ● 440 | 2,334 | 3,673 ● 323 | 1,373 | 2,221

INSIGHTS:

- The total number of results for Date, across all Returning Visits – First Time Visits – Page Loads, is 15.

7.PAGE LOADS BY ROW :

Page.Loads by Row colored by Day
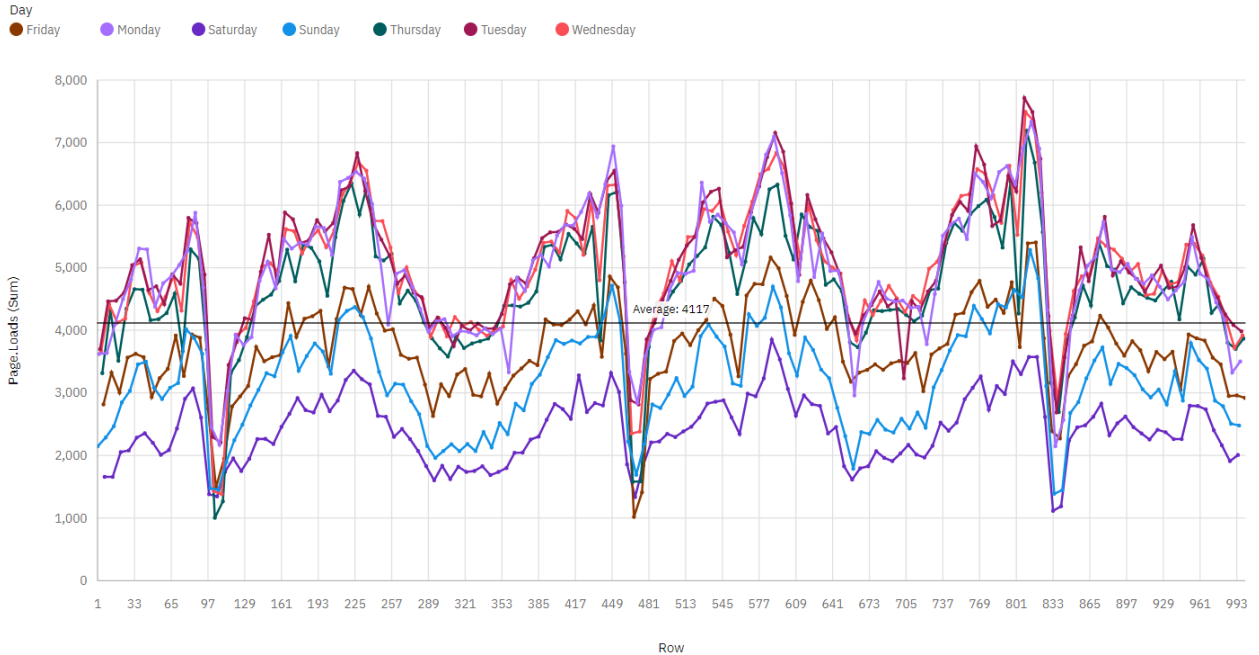
INSIGHTS:

- Row 1320 has the highest total Page Loads due to Day Wednesday.
- Page Loads is unusually low when Day is Saturday.
- Based on the current forecasting, Page Loads may reach over 675 thousand by Day Monday+1.
- Day Tuesday has the highest Page Loads at over 1.5 million, out of which Row 808 contributed the most at over 7500.
- Across all rows and days, the sum of Page Loads is over 8.9 million.
- The summed values of Page Loads range from over a thousand to nearly eight thousand.
- For Page Loads, the most significant values of Day are Tuesday, Wednesday, Monday, Thursday, and Friday, whose respective Page Loads values add up to over 7.1 million, or 80.1 % of the total.
- For Page Loads, the most significant value of Row is 1320, whose respective Page Loads values add up to nearly eight thousand, or 0.1 % of the total.

8.DAY OF WEEK BY RETURNING VISITS AND UNIQUE VISITS :

Day.Of.Week by Returning.Visits and Unique.Visits colored by Day

Day
● Monday ● Saturday ● Sunday
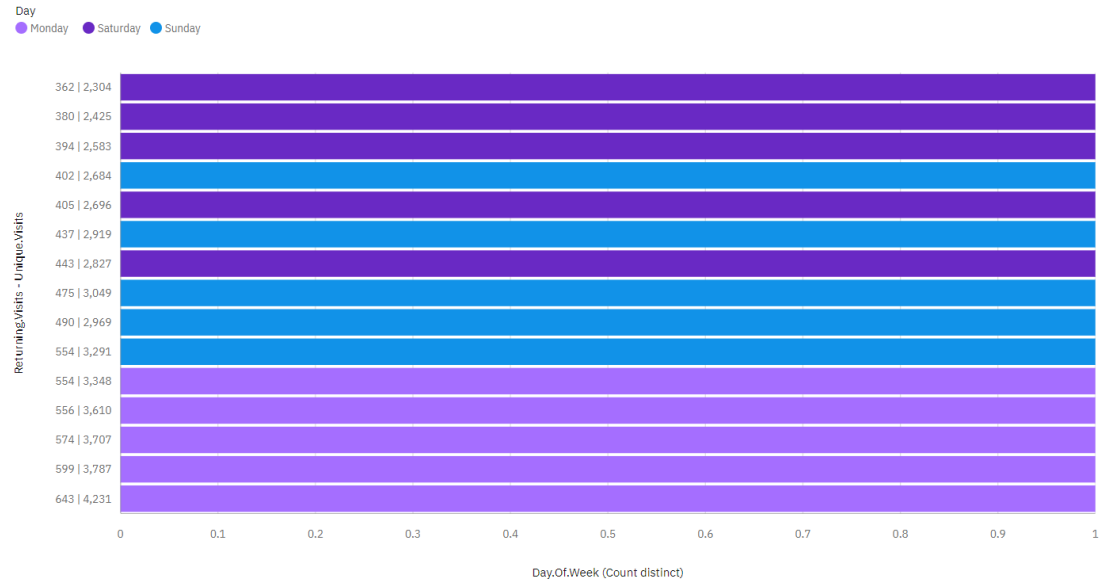
INSIGHTS:

- Day Monday has the highest values of both Returning Visits and First Time Visits.
- The overall number of results for Day Of Week is 15.

CONCLUSION:

      Data visualisation for the time series analysis were obtained using IBM Cognos tools and various analytics insights were inferred. By collecting and analyzing data related to how users interact with a website, this would help stakeholders gain valuable insights and make informed decisions to improve visitors online presence.