

Webpage Traffic Analysis Project Phase 3 - Documentation

Phase 3: Development Part 1:

Loading and Preprocessing the Dataset

Step 1: Dataset Download

Access the website traffic dataset from Kaggle. Ensure it contains relevant information about website traffic, such as date, time, page views, unique visits, traffic sources, and other relevant metrics.

Step 2: Loading the Dataset

- Import the necessary Python libraries, including Pandas, for data analysis.
- Load the website traffic dataset into a IBM Cognos Analytics for further analysis.
- Display the first few rows of the dataset to inspect the data structure.

Step 3: Exploratory Data Analysis

- Perform an initial exploration of the dataset to understand its structure and features.
- Check for missing values, data types, and basic statistics.
- Visualize key features to gain insights into the data, such as page views over time or traffic sources.
- Identify potential relationships or correlations between features and any relevant target variable.

Step 4: Define Analysis Objectives

- Define the objectives of the analysis for this phase, which may include:
- Understanding traffic patterns.
- Identifying key drivers of website traffic.
- Analysing user engagement.
- Detecting anomalies or unusual traffic patterns.

Step 5: Data Cleaning and Preprocessing

- Clean and preprocess the data to ensure its quality and suitability for analysis:
- Handle missing values by imputing or removing them based on data exploration.
- Encode categorical variables if necessary.
- Perform feature engineering (create new features) if required.
- If you want to predict certain outcomes, create a target variable, if not already present.

Step 6: IBM Cognos for Visualization

- Utilize IBM Cognos for creating data visualizations that provide insights into the website traffic dataset:
- Create various visualizations, such as line charts, bar charts, heatmaps, and interactive dashboards.
- Visualize the distribution of traffic metrics, feature importance, and other relevant insights using IBM Cognos.

Step 7: Data Validation

- Validate the processed data to ensure its quality and accuracy:
- Perform data validation checks to confirm data consistency and accuracy.
- Identify and address any potential data quality issues.

Step 8: Documentation

- Maintain comprehensive documentation of the activities performed in this phase:
- Record the data preprocessing steps, including any transformations and cleaning.
- Document the objectives defined for this phase.
- Capture insights gained from data visualization using IBM Cognos.
- Document any observations, challenges, or discoveries made during this phase.

Adapting this plan for website traffic analysis will help us to organize and carry out the necessary steps to analyse website traffic data effectively, and create meaningful visualizations using IBM Cognos. Remember to customize the objectives, metrics, and visualizations according to your specific goals and dataset.

Tab 6

Page.Loads, Day

8.92M

Page.Loads

Unique.Visits, Day

2.94K

Unique.Visits

First.Time.Visits, Day

4.62K

First.Time.Visits

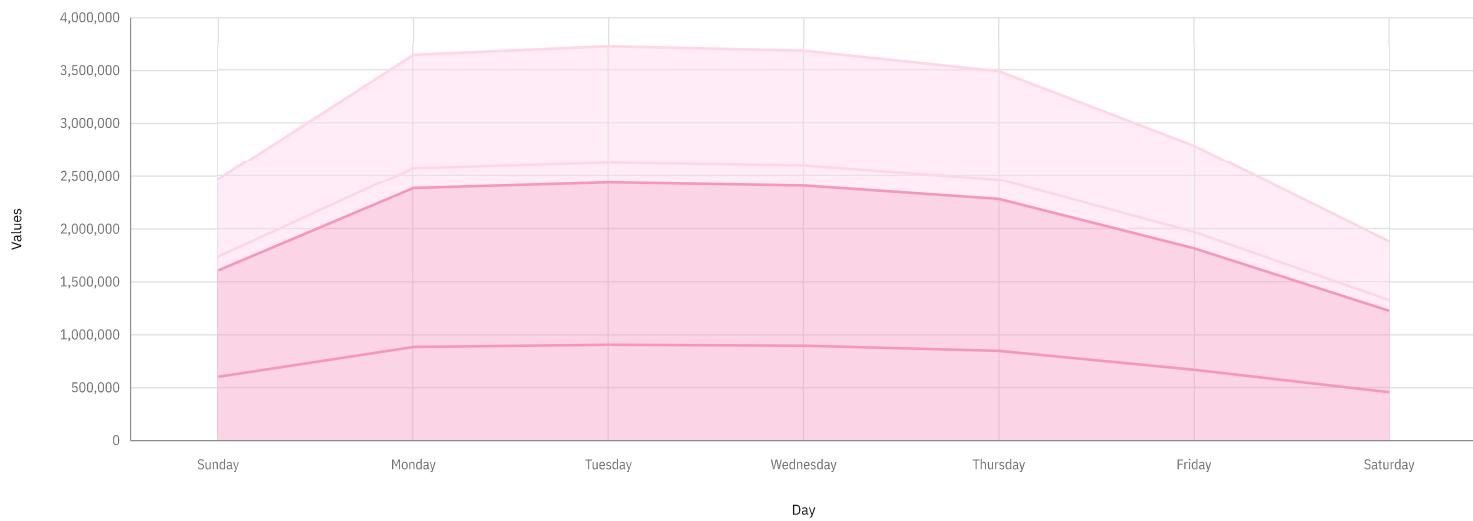
Returning.Visits, Day

511.82

Returning.Visits

First.Time.Visits, Page.Loads, Returning.Visits and Unique.Visits by Day

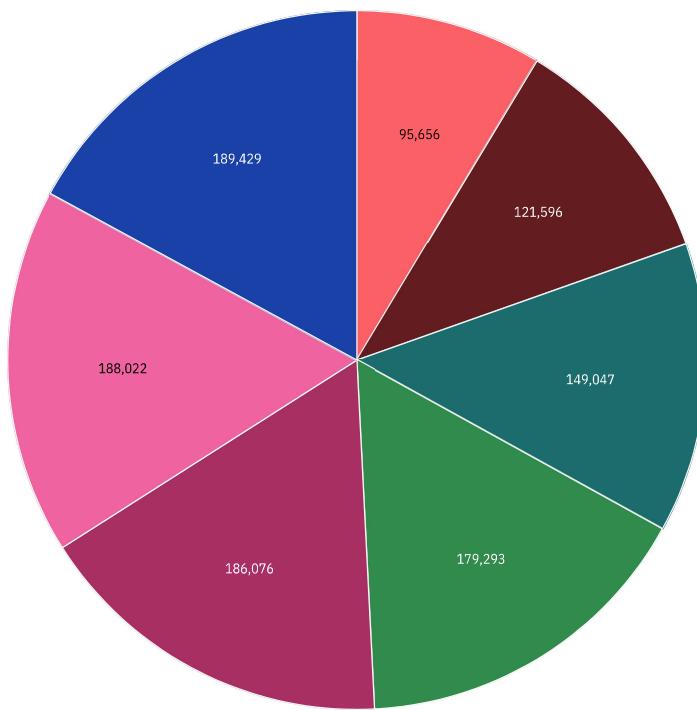
Measures

● First.Time.Visits ● Page.Loads ● Returning.Visits ● Unique.Visits

Tab 4

Returning.Visits by Day

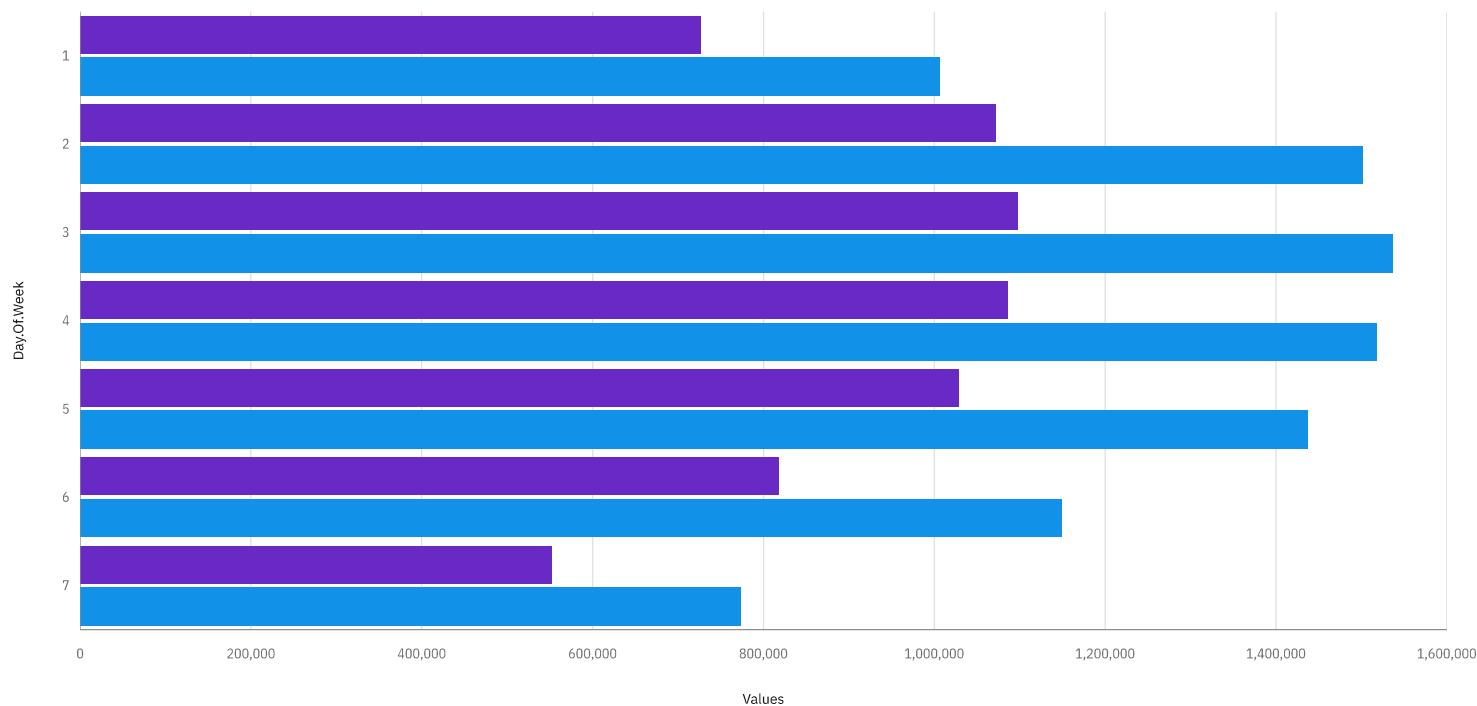
Day
Saturday Sunday Friday Thursday Monday Wednesday Tuesday



Tab 2

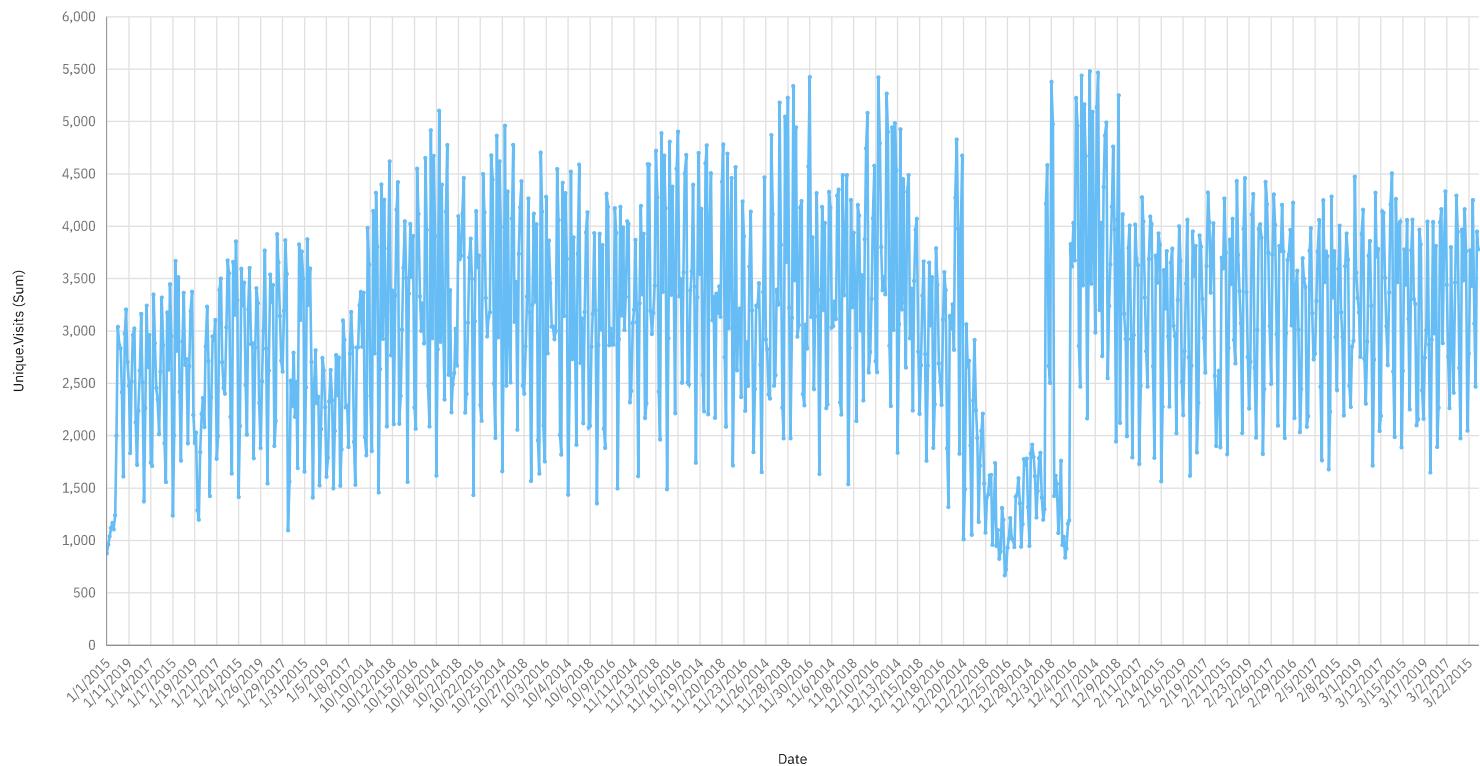
Unique.Visits and Page.Loads by Day.Of.Week

Measures

● Unique.Visits ● Page.Loads

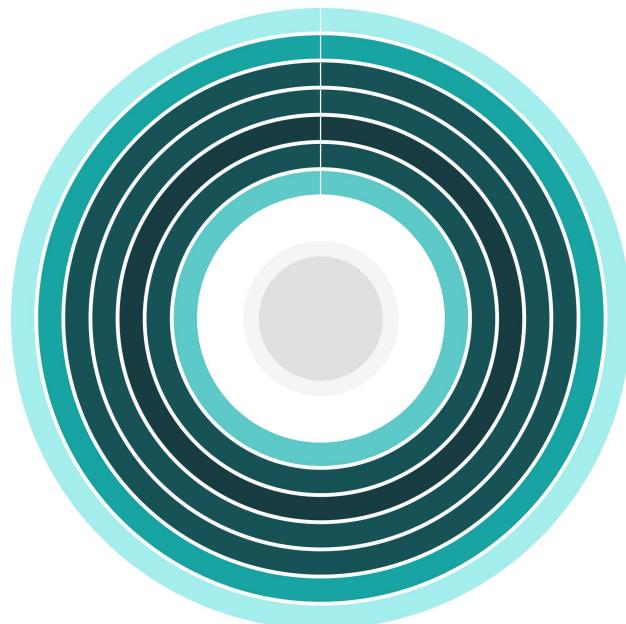
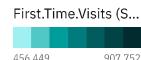
Tab 1

Unique.Visits by Date



Tab 5

Page.Loads - Day colored by First.Time.Visits



Day.Of.Week

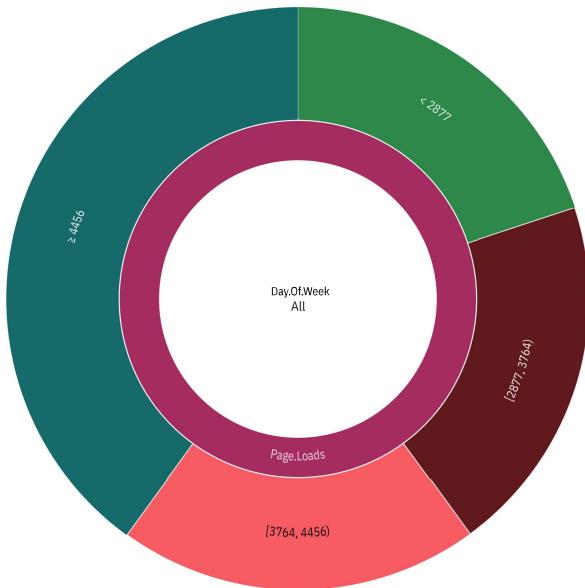
Nodes

All

Day.Of.Week



Target category



Tab 3

Day.Of.Week by Page.Loads with points for First.Time.Visits

