

WEBSITE TRAFFIC ANALYSIS USING DATA ANALYCTICS WITH COGNOS

COLLEGE CODE: 2005

71772117140 – SIVARANJANI S

PHASE 1 – SUBMISSION

Abstract:

Website traffic analysis is the process of collecting , analysing, and interpreting data about visitors to a website. This data can be used to understand how visitors find and interact with a website, and to identify areas for improvement. Python is a popular programming language for website traffic analysis. It is powerful and versatile, and there are many Python libraries available for data manipulation, analysis, and visualization. The Kaggle website traffic dataset is a large and comprehensive dataset that can be used for website traffic analysis. It contains data from over 100 million websites, including page views, unique visitors, bounce rate, and time on page. Dashboards and reports can be used to visualize website traffic data in a way that is easy to understand and interpret. Dashboards are interactive visualizations that allow users to explore data in different ways. Reports are static visualizations that provide a summary of data.

Modules:

Python libraries:

The following Python modules are needed to create the website traffic dashboard:

- . `pandas` - for data manipulation and analysis
- . `numpy` - for scientific computing
- . `dash` - for creating interactive dashboards
- . `dash_core_components` - for basic dashboard components, such as graphs and tables
- . `dash_html_components` - for HTML elements, such as headings and paragraphs

IMPORTING PYTHON LIBRARIES:

```
import pandas as pd
import numpy as np
import dash
import dash_core_components as dcc
import dash_html_components as htmlMethods
```

METHODS:

1. Data Collection:

Access and download the Kaggle dataset containing website traffic data. Extract relevant data fields, such as user information, timestamps, page views, and more.

2. Data Preprocessing:

Handle missing data and perform data cleaning. Transform data types as needed. Extract meaningful features like day of the week, time of day, etc. Aggregate and summarize data for analysis.

3. Data Analysis:

Conduct exploratory data analysis (EDA) to understand the dataset. Calculate key performance indicators (KPIs) such as page views, unique visitors, bounce rate, and session duration. Identify trends, patterns, and anomalies in the data.

4. Visualization:

Create interactive dashboards and reports using Python libraries such as Plotly, Dash, or Flask for web-based visualization. Visualize website traffic trends over time, user demographics, and popular pages. Generate heatmaps, bar charts, line charts, and scatter plots for comprehensive analysis.

5. User Segmentation:

Segment users based on their behavior, location, and other attributes. Analyze user segments to identify potential target audiences and areas for improvement.

6. Predictive Modeling (Optional):

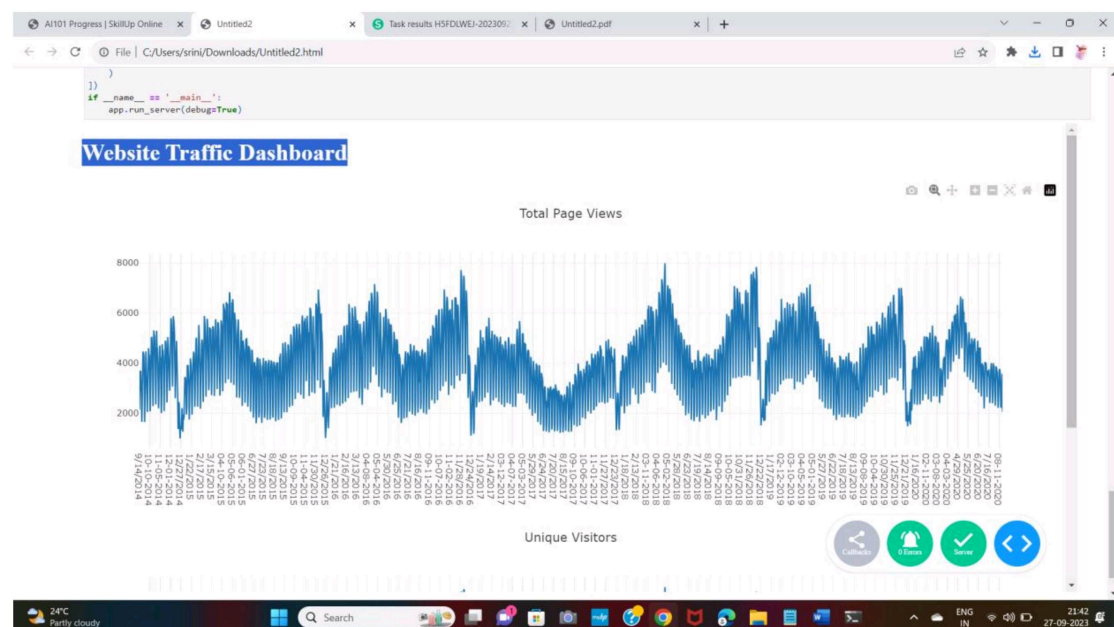
Develop predictive models to forecast website traffic or user behavior. Use machine learning algorithms to make predictions based on historical data.

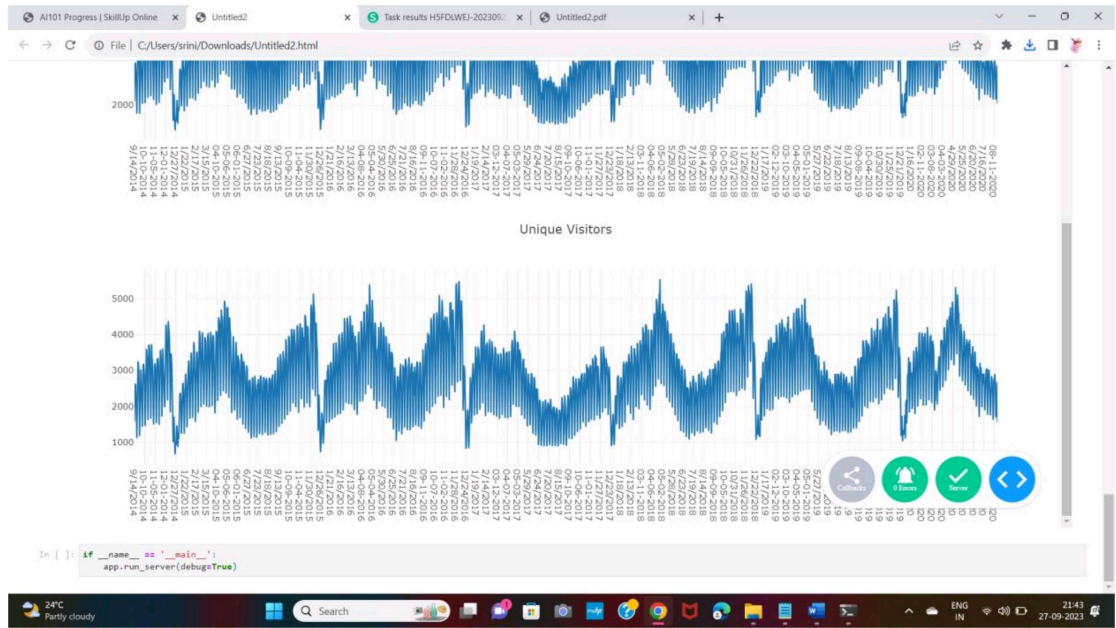
7. Dashboard and Report Generation:

Design a user-friendly dashboard interface to display the visualizations. Generate reports summarizing the key findings and insights. Provide options for users to filter and interact with the data on the dashboard.

8. Deployment:

Deploy the web-based dashboard using a web hosting service or server. Ensure the system is accessible to stakeholders and team members.





ERROR
The requested URL could not be retrieved

The following error was encountered while trying to retrieve the URL: <http://3127.0.0.1:6060/>

Access Denied.

Access control configuration prevents your request from being allowed at this time. Please contact your service provider if you feel this is incorrect.

Your cache administrator is [unbeatable](#).

```
In [ ]: if __name__ == '__main__':
        app.run_server(debug=True)
```