PUBLIC HEALTH CAMPAIGN AWARENESS

NAAN MUDHALVAN PROJECT REPORT

Submitted by

DHARSHINI V (730321243005)

JUSVANTHIHA S (730321243008)

KAMALA V (730321243010)

NAGALAKSHMI K (730321243015)

FIFTH SEMESTER

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

BUILDERS ENGINEERING COLLEGE , KANGEYAM

ANNA UNIVERSITY :: CHENNAI 600 025

NOVEMBER 2023

BONAFIDE CERTIFICATE

Certified that this is a Bonafide record of work done by **DHARSHINI V(730321243005) , JUSVANTHIHA S(730321243008)** and **KAMALA V(730321243010) , NAGALAKSHMI K(730321243015)** in NAAN MUTHALVAN during the Academic year 2023- 24 for fifth Semester.

STAFF -INCHARGE **HEAD OF THE DEPARTMENT**

Submitted for the Naan mudhalvan viva voce held on_____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ABSTRACT

Website traffic analysis is a fundamental practice for online businesses and digital marketers, providing invaluable insights into user behavior, content performance, and overall website health. This abstract provides a concise overview of the importance, methods, and benefits of website traffic analysis.

In an increasingly digital world, websites serve as primary touchpoints for businesses and organizations. Analyzing website traffic is crucial to understanding how users interact with online content. By collecting and interpreting data, organizations can make informed decisions to optimize their online presence, improve user experience, and achieve their objectives.

Website traffic analysis, including data collection, data sources, and analysis techniques.

Data Collection: The primary data sources for website traffic analysis include web analytics tools (e.g., Google Analytics), server logs, and third-party data sources (e.g., social media insights). These sources provide information about user demographics, devices, referral sources, and behavior.

User Behavior: Understanding how users navigate a website is vital. Metrics such as pageviews, bounce rate, session duration, and conversion rates offer insights into user engagement and content effectiveness. By tracking the customer journey, organizations can identify pain points and opportunities for improvement.

Content Performance: Website traffic analysis helps assess the performance of individual web pages. Metrics like popular content, exit pages, and click-through rates shed light on which pages resonate with users and which require optimization.

Referral Sources: Analyzing where website traffic originates provides insights into the success of marketing efforts. Evaluating traffic from search engines, social media, email campaigns, and direct sources can help fine-tune marketing strategies.

Website Health: Regular website traffic analysis can uncover technical issues like broken links, slow-loading pages, and mobile responsiveness problems. Addressing these issues is essential to provide a smooth user experience.

Website traffic analysis empowers organizations to make data-driven decisions, allocate resources effectively, and continually improve their online presence. It is an indispensable tool for optimizing websites, increasing user satisfaction, and achieving business goals in the digital age.

CHAPTER 1

PROBLEM STATEMENT

In today's digital age, businesses and websites rely heavily on online presence and user engagement. Understanding and analyzing web traffic data has become paramount to make informed decisions and enhance the user experience. The problem at hand is to develop a comprehensive web traffic analysis solution that can collect, process, and derive actionable insights from vast amounts of web traffic data. This solution should address the following key challenges:

Data Collection and Storage: Implement a robust system to collect, store, and manage web traffic data efficiently. This includes capturing user interactions, page views, session durations, and other relevant metrics.

Data Processing and Real-time Analysis: Develop the capability to process the collected data in real-time or near-real-time, allowing for instant feedback on website performance and user behavior.

User Behavior Understanding: Utilize advanced analytics techniques to gain insights into user behavior, including user segmentation, navigation patterns, and conversion funnel analysis.

Performance Optimization: Identify performance bottlenecks, such as slow-loading pages or high bounce rates, and recommend optimizations to enhance the website's user experience.

Security and Privacy: Ensure data security and privacy compliance while handling sensitive user information. Implement strategies to protect user data and meet regulatory requirements.

Data Visualization and Reporting: Create intuitive dashboards and reports that enable stakeholders to easily interpret the data, track key performance indicators, and make informed decisions.

Scalability: Design the solution to be scalable, allowing it to handle increasing amounts of web traffic data as the website's user base grows.

Predictive Analysis: Implement predictive analytics to forecast future web traffic trends, enabling proactive decision-making and resource allocation.

Accessibility and User Experience: Ensure that the web traffic analysis solution is accessible to a wide range of users within the organization, providing an intuitive and user-friendly interface.

Customization and Integration: Allow for customization and integration with existing systems and tools, enabling organizations to tailor the solution to their specific needs.

The successful development and implementation of a web traffic analysis solution will empower organizations to make data-driven decisions, improve their websites, and ultimately enhance user satisfaction, leading to increased engagement and business success."

This problem statement can serve as a foundation for a project focused on web traffic analysis, helping to guide the development of a solution that meets the specific needs and goals of an organization.

CHAPTER 2

DESIGN THINKING

Real-time Analytics:

Implement real-time analytics tools that provide instant insights into user behavior. This allows website owners to make immediate changes based on current data, improving the user experience as it happens.

1.User-Centric Analytics:

Shift the focus from traditional metrics like page views to more user-centric metrics. For instance, analyze user journeys, engagement, and satisfaction rather than just raw traffic numbers. Understand why users visit the site and whether they achieve their goals.

2.AI and Machine Learning Integration:

Incorporate AI and machine learning algorithms to predict user behavior. This can help in personalizing content, recommendations, and user experiences. It can also aid in identifying patterns that lead to conversions or drop-offs.

3.Predictive Analytics:

Develop predictive analytics models to anticipate future traffic trends. This can be particularly useful for e-commerce sites to optimize inventory, plan marketing campaigns, and improve user experience based on expected traffic.

4.Multi-Channel Analysis:

Analyze traffic across various channels, including social media, email, search engines, and more. Understand how different channels contribute to overall traffic and how they affect user behavior.

5.Custom Dashboards:

Offer customizable dashboards that allow website owners and stakeholders to create personalized views of the data. This empowers users to focus on the specific metrics and KPIs that matter most to them.

6.User Segmentation:

Implement advanced user segmentation based on demographics, behavior, and other criteria. This allows website owners to tailor content and marketing strategies to specific user groups.

7.Voice and Image Recognition:

Integrate voice and image recognition for analytics. This can help identify trends related to voice searches and image searches, providing insights into emerging user behaviors.

8.Data Visualization:

Use a line chart to show how your traffic changes over time, a pie chart to show the distribution of your traffic sources, or a heatmap to show the areas of your website that get the most clicks. By using visual tools, you can make your data more accessible, engaging, and actionable.

CHAPTER 3

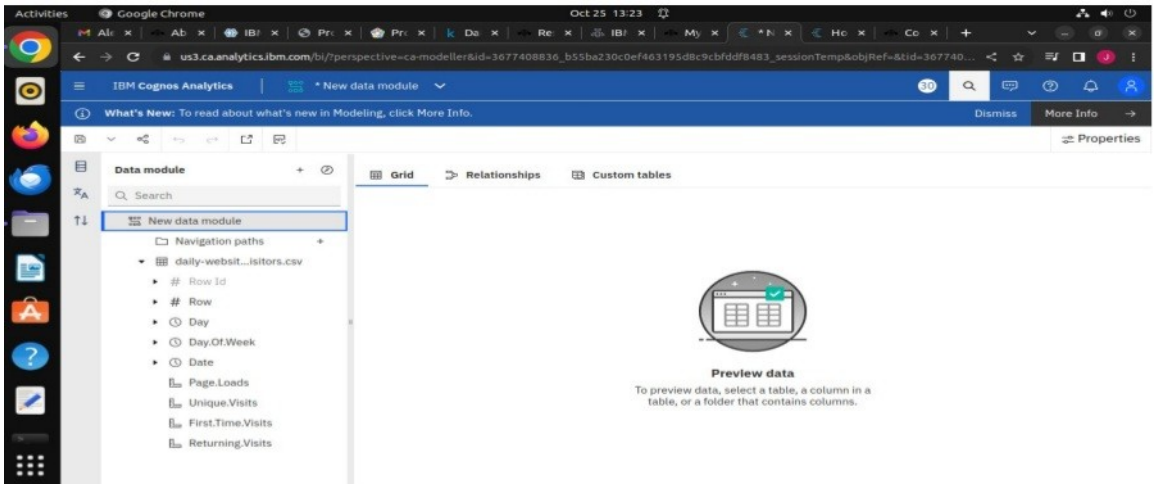
DATASET DEFINITION

OBJECTIVES

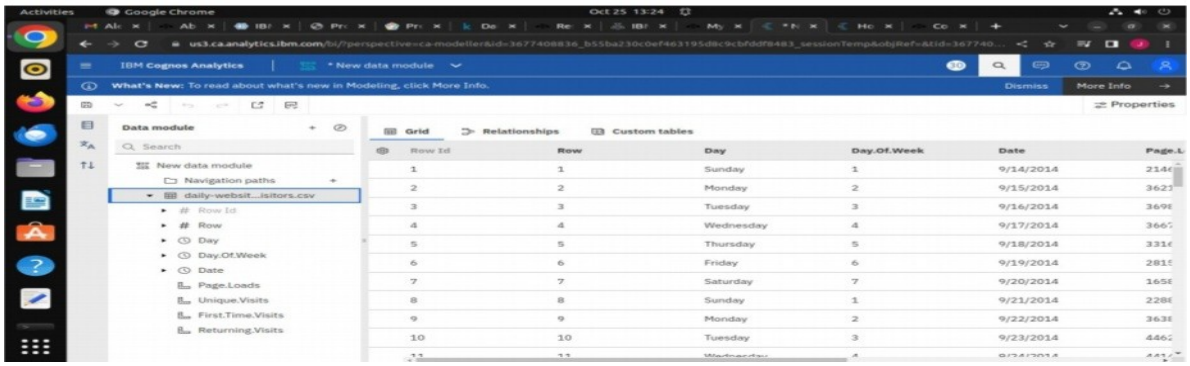
- Data collection
- Data importing
- Data analyzing
- Data preprocessing
- Data visualization

STEP-1: Downloading and uploading the dataset

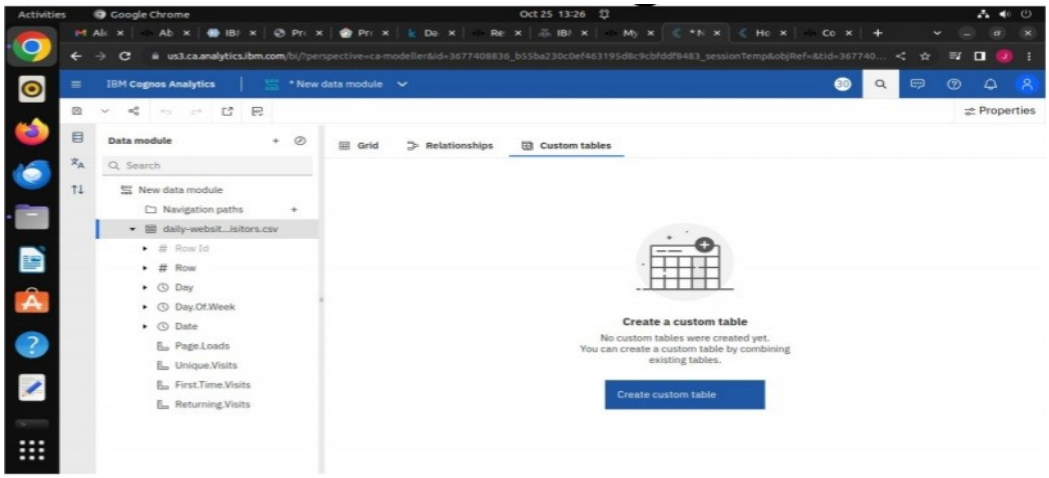
The Following dataset is downloaded and it is saved in file manage and it is uploaded

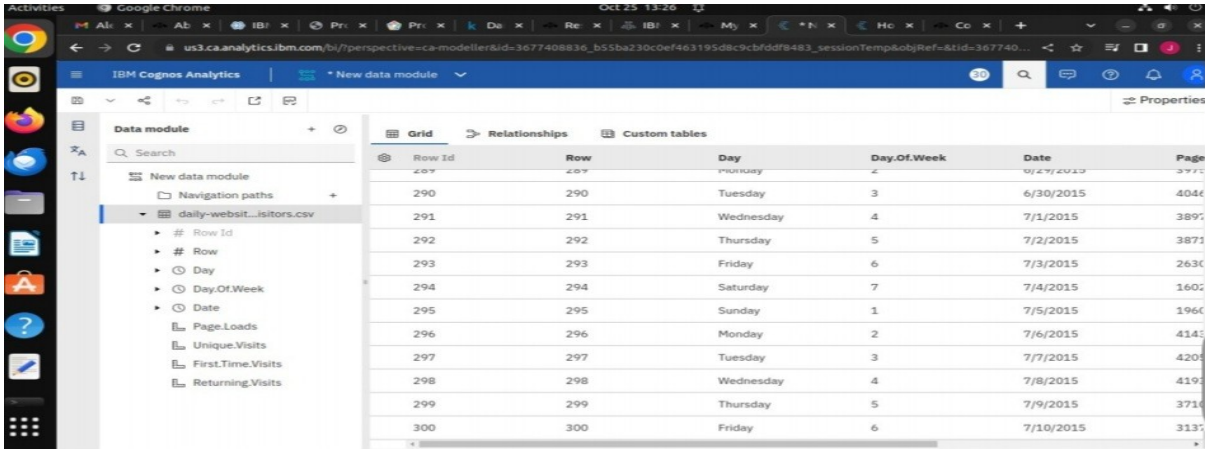


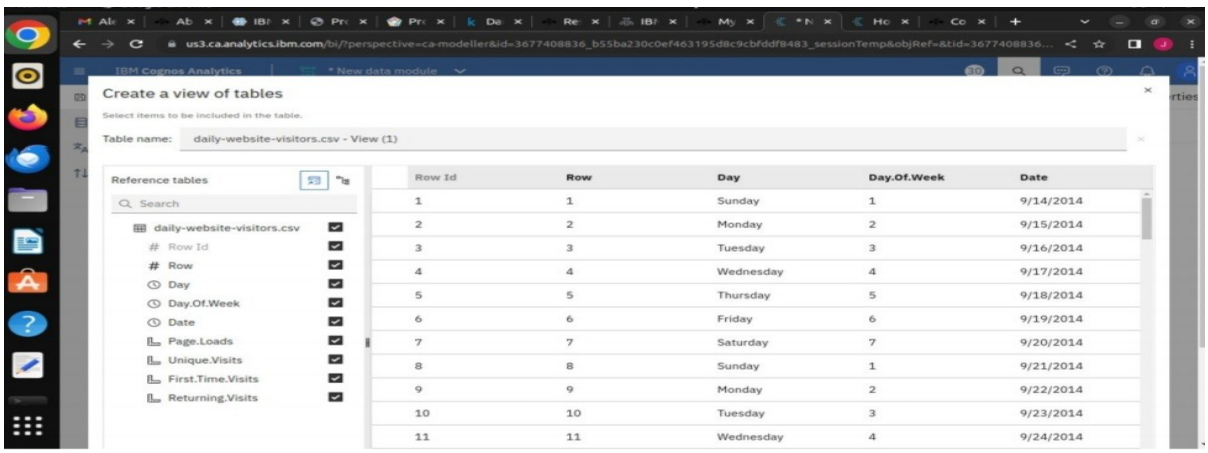
Step-2:Data collection and analysis ● In this stage,data is collected and analysis using IBM cognos analytics software

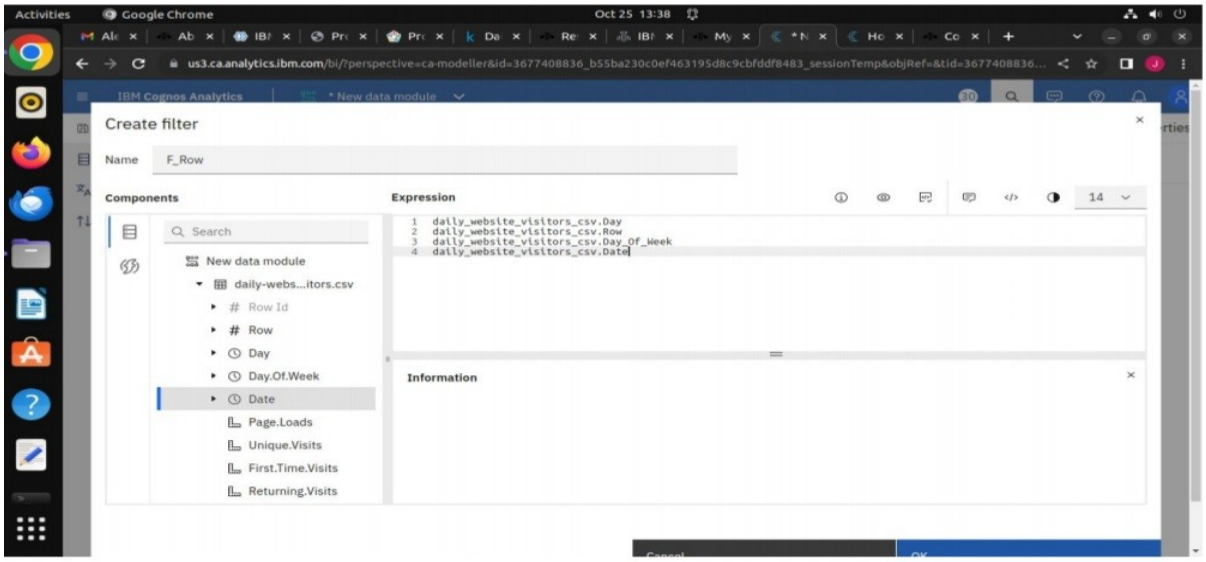


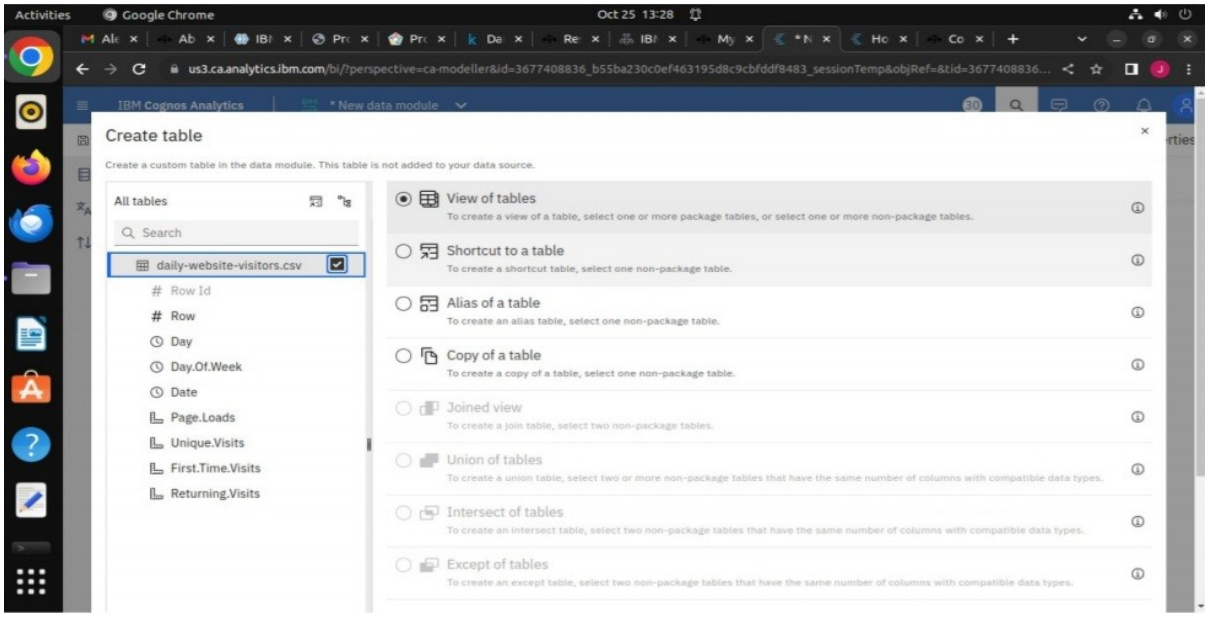
Step-3:Data preprocessing and cleaning








Activities
Google Chrome
Oct 25 13:38
us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=3677408836_b55ba230c0ef463195d8c9cbfddf8483_sessionTemp&objRef=&tid=3677408836...
IBM Cognos Analytics
* New data module
Create filter
Name
F_Row
Components
Expression
Search
New data module
daily-webs...itors.csv
Row Id
Row
Day
Day.Of.Week
Date
Page.Loads
Unique.Visits
First.Time.Visits
Returning.Visits
1 daily_website_visitors_csv.Day
2 daily_website_visitors_csv.Row
3 daily_website_visitors_csv.Day_Of_Week
4 daily_website_visitors_csv.Date
Information


Activities
Google Chrome
Oct 25 13:28
us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=3677408836_b55ba230c0ef463195d8c9cbfddf8483_sessionTemp&objRef=&tid=3677408836...
IBM Cognos Analytics
* New data module
Create table
Create a custom table in the data module. This table is not added to your data source.
All tables
Search
daily-website-visitors.csv
Row Id
Row
Day
Day.Of.Week
Date
Page.Loads
Unique.Visits
First.Time.Visits
Returning.Visits
View of tables
To create a view of a table, select one or more package tables, or select one or more non-package tables.
Shortcut to a table
To create a shortcut table, select one non-package table.
Alias of a table
To create an alias table, select one non-package table.
Copy of a table
To create a copy of a table, select one non-package table.
Joined view
To create a join table, select two non-package tables.
Union of tables
To create a union table, select two or more non-package tables that have the same number of columns with compatible data types.
Intersect of tables
To create an intersect table, select two non-package tables that have the same number of columns with compatible data types.
Except of tables
To create an except table, select two non-package tables that have the same number of columns with compatible data types.

CHAPTER 4

DATA PREPROCESSING AND FEATURE EXTRACTION

MODULES:

1)SPLITTING THE TRAINING AND TESTING DATA

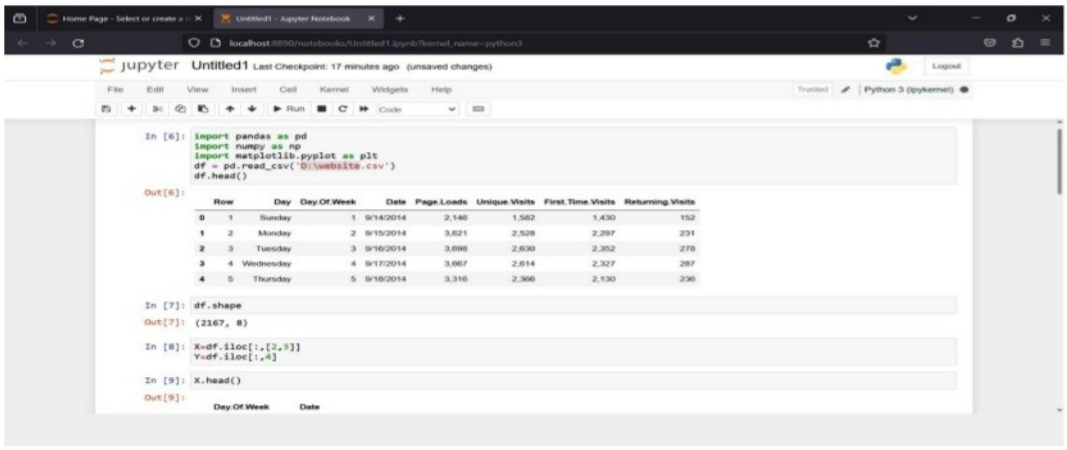
2) APPLYING SUPPORT VECTOR MACHINES

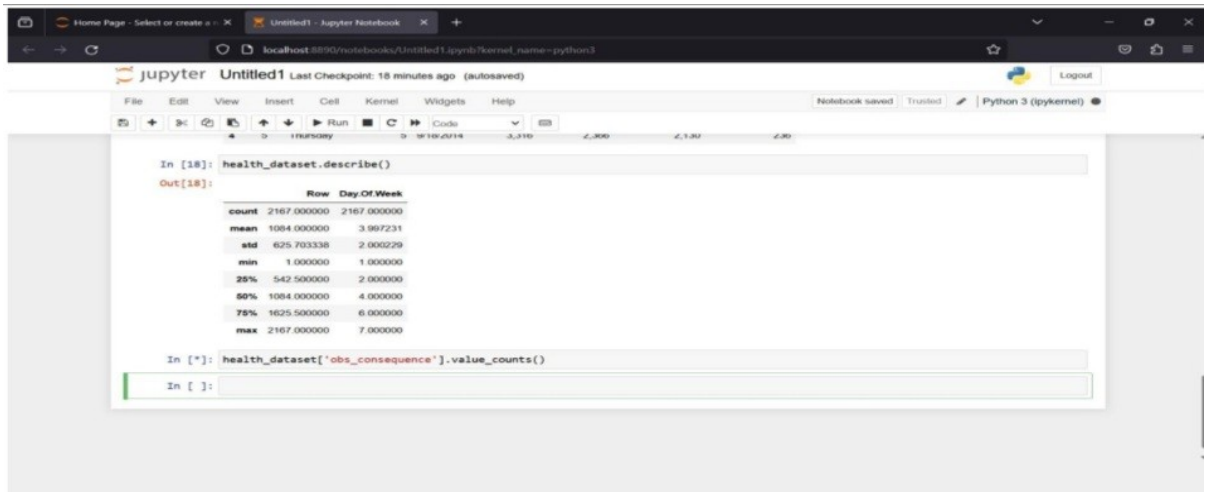
3)CREATING DASHBOARD USING IBM COGNOS

4)CREATING REPORT USING IBM COGNOS

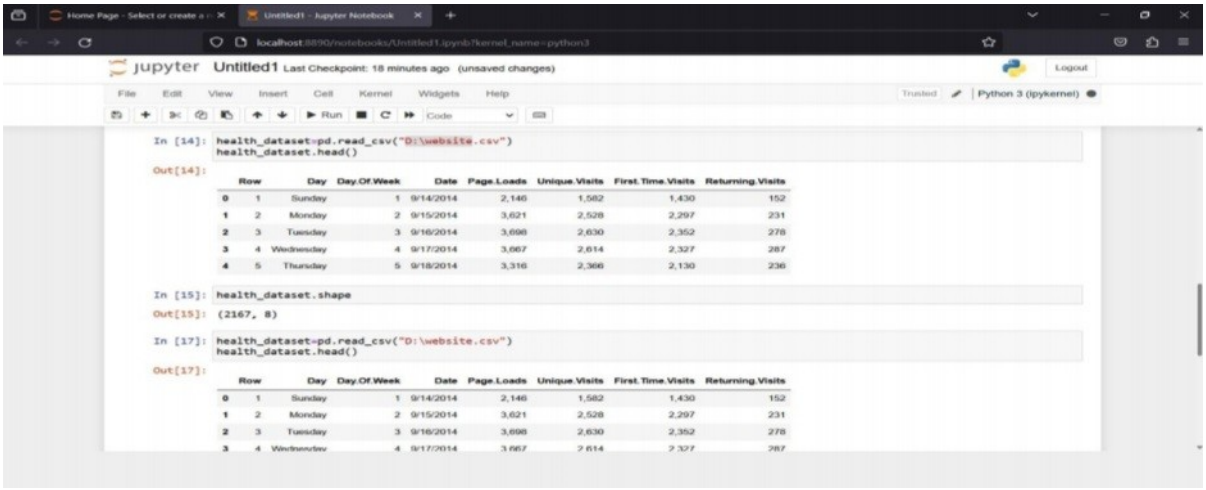
5)STATISTICAL TESTS USING TREEMAP

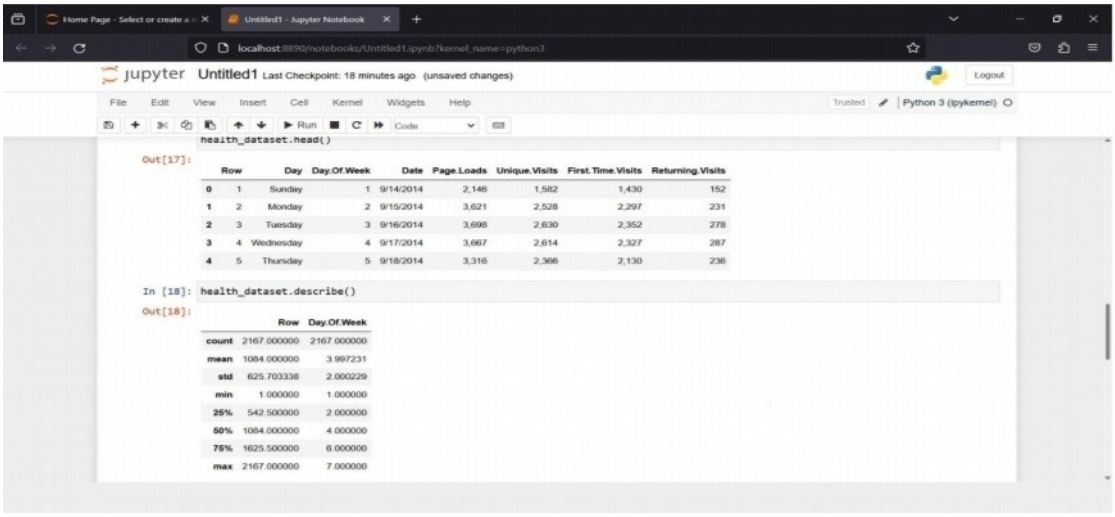
Step1: Splitting the testing and training data.



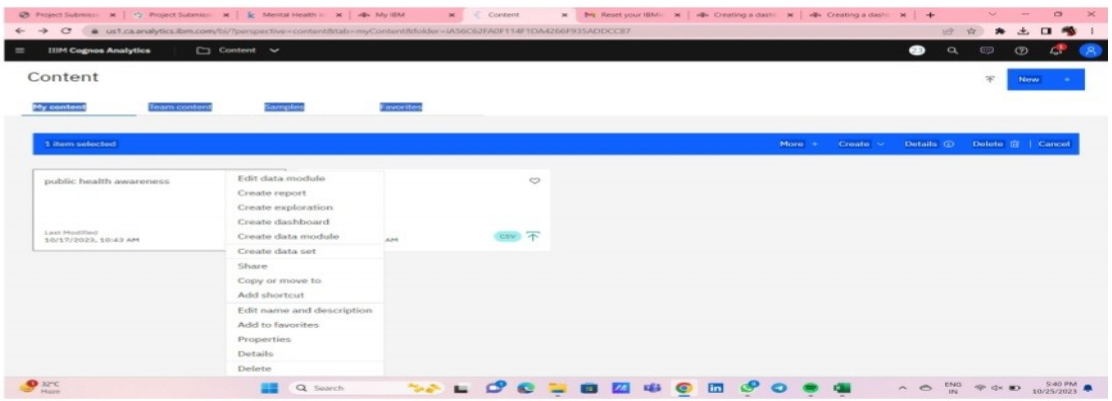


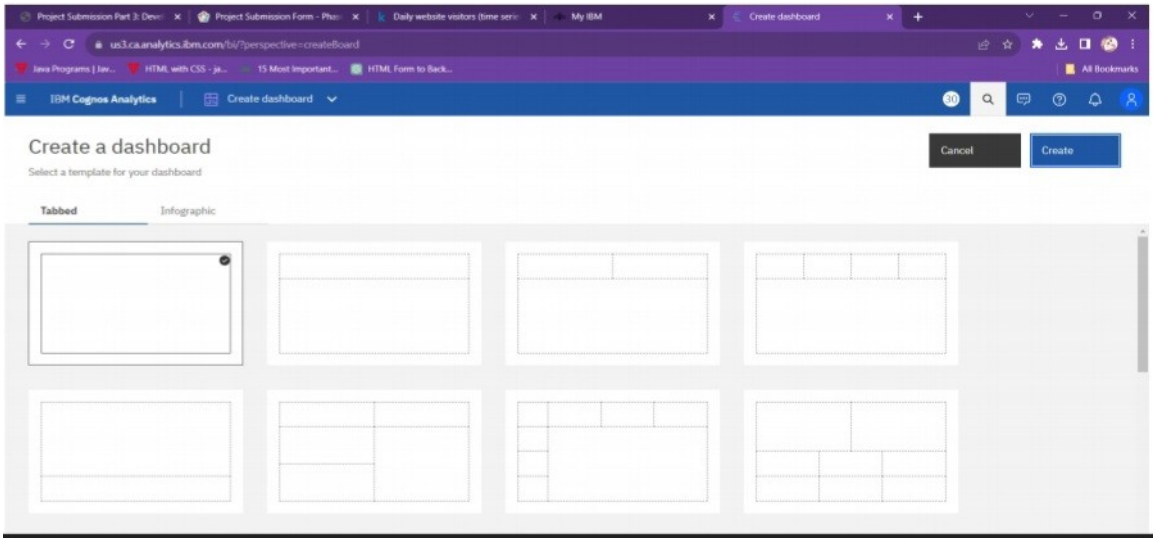
Step 2: Applying support vector machines and loss calculation.

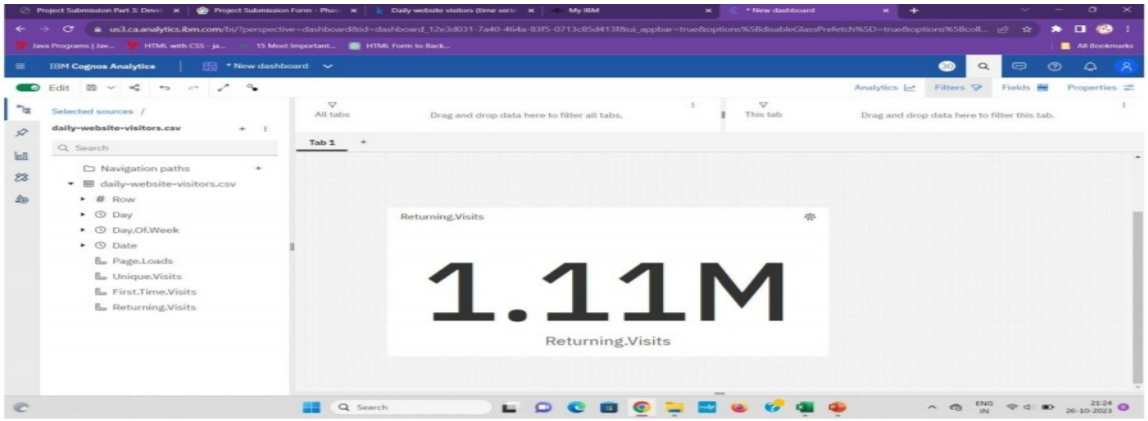




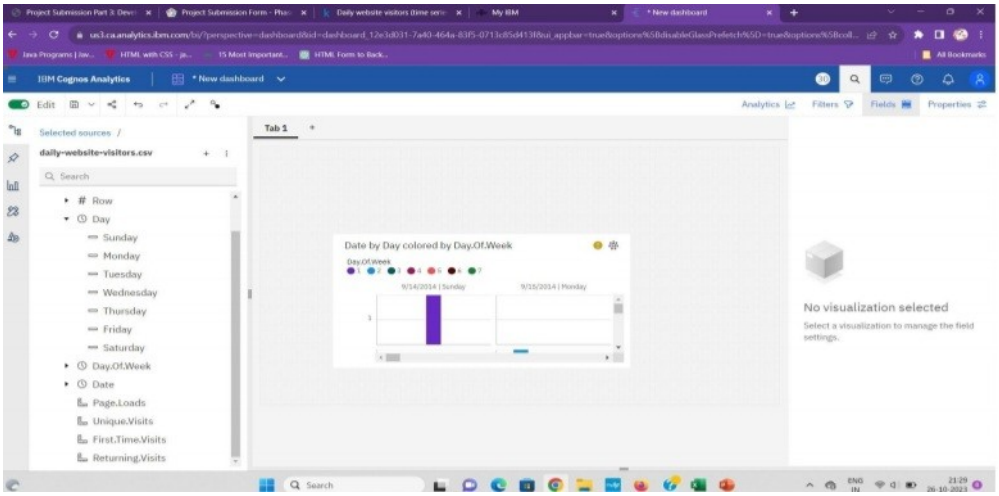
Step 3: Creating dashboard using IBM COGNOS



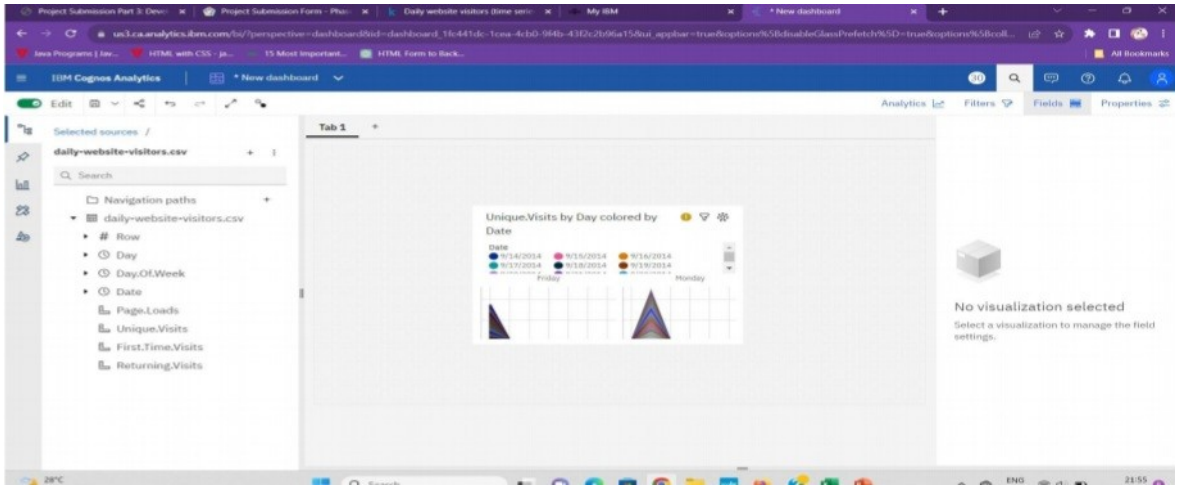




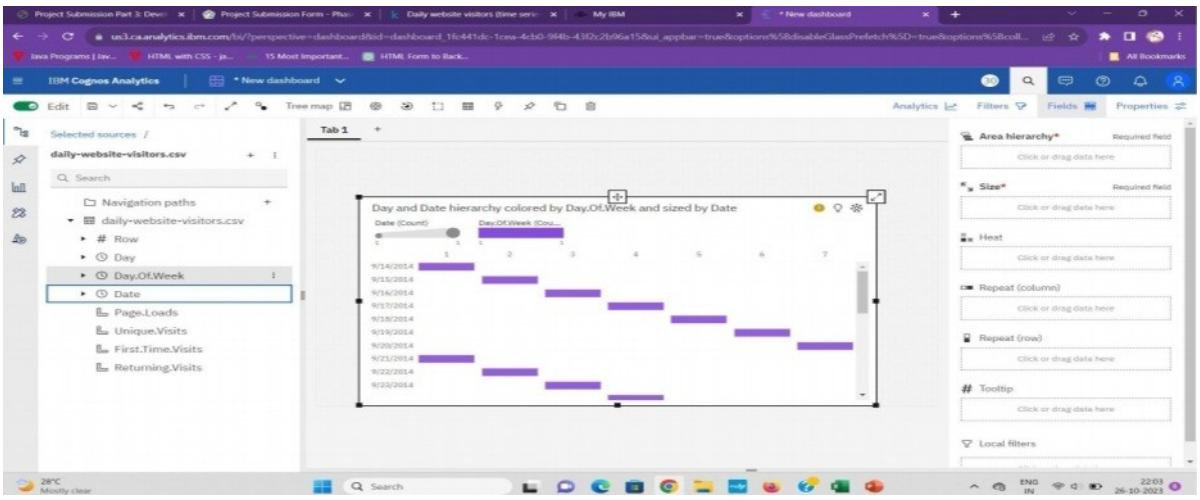
Creating bar graph using dashboard



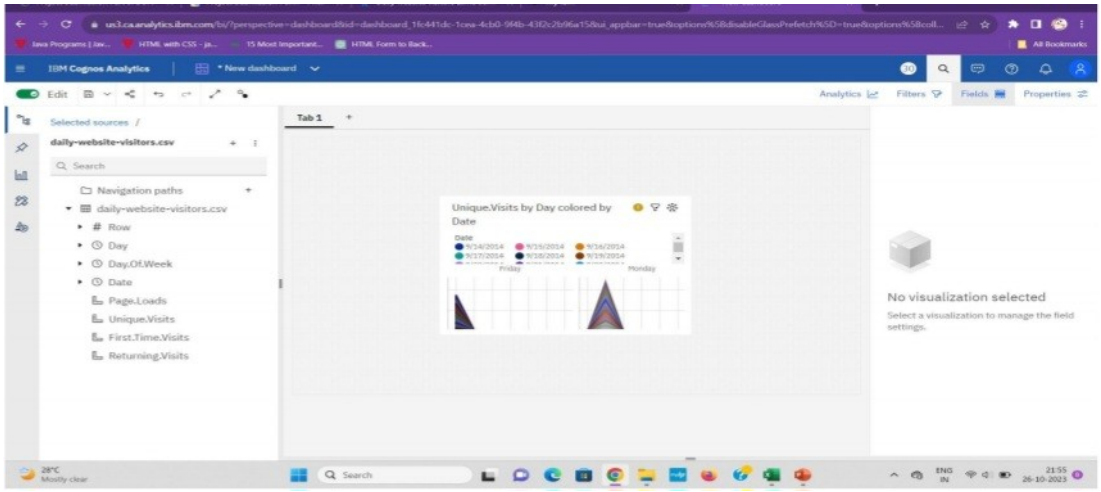
Creating area using dashboard



Creating bar graph using report



Statistical tests using treemap



CHAPTER 5

PROPOSED ALGORITHM

SUPPORT VECTOR MACHINES (SVM)

Support Vector Machines (SVM) can be used in website traffic analysis for various purposes, such as classification and anomaly detection. Here's how SVM can be applied in this context:

User Behavior Classification: SVM can be used to classify user behavior on a website. For example, it can classify users into categories like new visitors, returning visitors, or potential buyers based on their interaction patterns with the website. This can help tailor the user experience for different segments and improve conversion rates.

Bot Detection: SVMs can be used to detect and classify bot traffic. Bots can skew website traffic analysis and can be detrimental to the user experience. SVMs can be trained to distinguish between legitimate human users and bots based on various features such as user-agent strings, IP addresses, request patterns, and more.

Anomaly Detection: SVMs can be used to identify unusual or anomalous traffic patterns on a website. By training an SVM on historical data, it can learn what "normal" traffic patterns look like and flag any deviations from this norm. This can be crucial in identifying potential security threats or performance issues.

Click Fraud Detection: For e-commerce websites, SVMs can be used to detect click fraud. By analyzing the behavior of users who click on ads or affiliate links, SVMs can identify patterns that suggest fraudulent activities.

Conversion Rate Optimization: SVMs can be employed to optimize website conversion rates. By analyzing user behavior and characteristics, SVMs can help identify which elements of a website (e.g., content, layout, buttons) influence users to convert (e.g., make a purchase, sign up for a newsletter). This information can be used to make data-driven design decisions.

Predictive Analysis: SVMs can be used to predict future website traffic based on historical data. This can be valuable for capacity planning, content scheduling, and resource allocation.

Content Recommendation: SVMs can be used to personalize content recommendations for website users. By analyzing the browsing behavior and preferences of users, SVMs can suggest relevant articles, products, or services, thus improving user engagement and satisfaction.

In all these applications, SVMs require labeled training data to learn the patterns or boundaries that separate different classes of data. Feature engineering is also a critical aspect, where relevant features such as page views, click-through rates, session duration, and more need to be extracted from the website's logs or database.

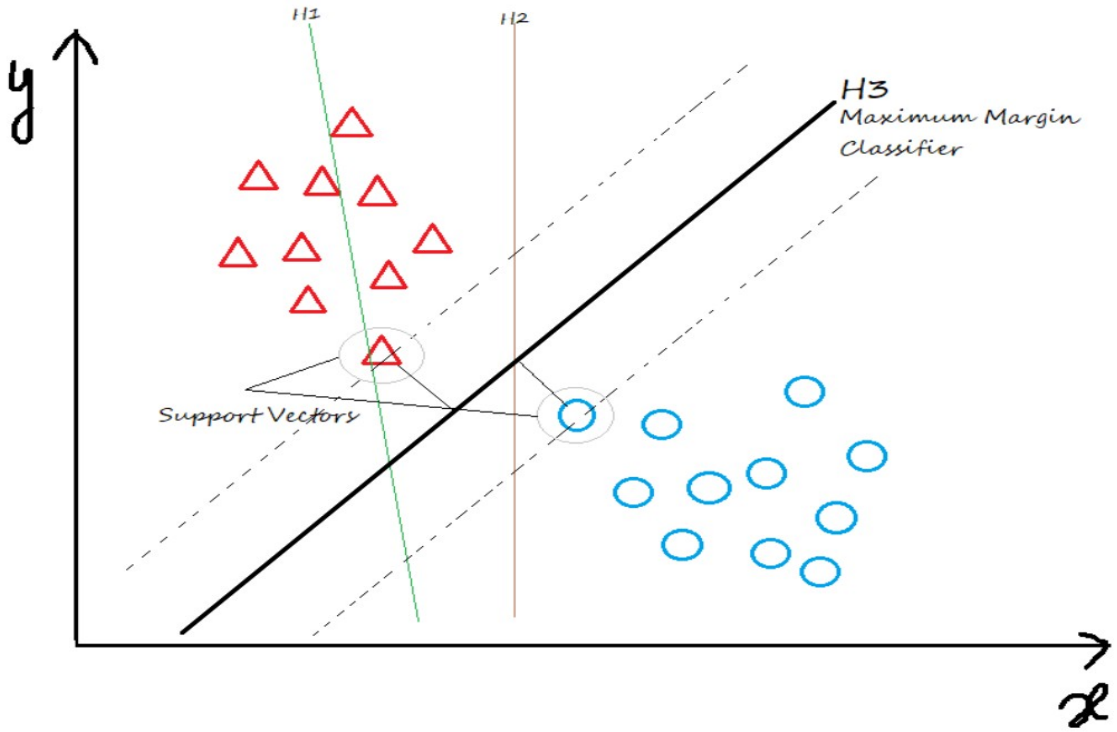
```
>>> from sklearn import svm
>>> X = [[0, 0], [1, 1]]
>>> y = [0, 1]
>>> clf = svm.SVC()
>>> clf.fit(X, y)
SVC()


>>> clf.predict([[2., 2.]])
array([1])


>>> # get support vectors
>>> clf.support_vectors_
array([[0., 0.],
       [1., 1.]])
>>> # get indices of support vectors
>>> clf.support_
array([0, 1]...)
>>> # get number of support vectors for each class
>>> clf.n_support_
array([1, 1]...)
```

CHAPTER 6

PROPOSED INNOVATION TECHNIQUE

Support vector machine

Support vector machine SVM have been one of the most successful machine learning techniques in recent years, applied successfully to many engineering related applications including those of the petroleum and mining. In this chapter, attempts were made to indicate how an SVM works and how it can be structured to provide reliable results. Few issues were raised including selection of kernel functions and other parameters of SVMs. By providing examples from different applications of SVMs, it was concluded that the RBF (Gaussian) kernel function is perhaps the best kernel to have an efficient SVM. To select the parameters of SVMs, though, it seems that the cross-validation approach would be the best choice based on the studies carried out so far. One should always remember that although SVMs are very good approaches in resolving the issues raised by having a limited number of data for training, they might be a time-consuming approach if applied to a huge database.



CHAPTER 7

CONCLUSION AND FUTURE SCOPE

CONCLUSION

Website traffic analysis is a critical component of web management, digital marketing, and user experience optimization. It provides valuable insights into how users interact with a website, enabling data-driven decision-making and helping organizations achieve their online goals. In conclusion, some key takeaways from website traffic analysis are:

Understanding User Behavior: Website traffic analysis allows you to gain insights into how users behave on your site. You can track page views, visitor demographics, geographical locations, and user journeys to understand their preferences and needs.

Performance Evaluation: It helps in assessing the performance of your website, identifying slow-loading pages, and pinpointing technical issues that may impact user experience.

Conversion Optimization: Analyzing user behavior and traffic patterns can lead to improved conversion rates by identifying bottlenecks and areas for improvement in the sales funnel.

Content Strategy: Data from website traffic analysis can inform content strategy. You can identify popular pages, user engagement with specific content, and the effectiveness of marketing campaigns.

SEO Improvement: You can use website traffic data to enhance your search engine optimization (SEO) efforts. It helps in identifying keywords that drive traffic, improving meta tags, and monitoring the impact of SEO changes on your site's visibility in search engines.

User Experience Enhancement: By analyzing how users navigate your website, you can enhance the user experience. This includes improving site navigation, reducing bounce rates, and providing content that matches user interests.

Security and Anomaly Detection: Website traffic analysis can be used to identify and mitigate security threats such as bot attacks and anomalous user behavior.

Data-Driven Decision Making: The data gathered from website traffic analysis should inform strategic decisions. It helps in setting goals, evaluating the success of marketing campaigns, and allocating resources effectively.

Personalization: Understanding user behavior allows for the implementation of personalized user experiences, such as product recommendations and tailored content.

Capacity Planning: For websites with variable traffic loads, analysis can help with capacity planning, ensuring that the infrastructure can handle traffic spikes.

In conclusion, website traffic analysis is a dynamic and continuous process. Regular monitoring and analysis are essential to adapt to changing user behavior, technology trends, and market conditions. With the right tools and methodologies, organizations can leverage website traffic data to optimize their online presence, improve user satisfaction, and achieve their online objectives.