



DATA ANALYTICS

Most sought after skills in coming years
- World Economic Forum

Uncovering User Insights: Website Traffic Analysis with IBM Cognos

P. BAKKIYAM

CK COLLEGE OF ENGINEERING AND TECHNOLOGY

ABSTRACT:

The goal of this project is to analyse website traffic data in order to acquire important insights into user behaviour, popular sites, and traffic sources. The major purpose is to allow website owners to improve the user experience by learning how users interact with their site. The project entails setting clear analytical objectives, gathering important website traffic data from various sources, utilising IBM Cognos for data visualisation to build meaningful dashboards and reports, and integrating Python code for advanced data analysis. The project's goal in applying these tactics is to give practical recommendations that can lead to improvements in user engagement and website performance.

OBJECTIVE:

In this website traffic analysis project, the primary objectives are to gain comprehensive insights into user behavior, popular website pages, and traffic sources, enabling data-driven enhancements for user experiences. We aim to understand how users engage with the website, their navigation patterns, and session durations (Objective 1). Identifying and prioritizing high-traffic pages for content optimization (Objective 2) and categorizing traffic sources (Objective 3) will guide our focus. Additionally, profiling visitors based on demographics (Objective 4) will allow for tailored experiences. Leveraging these insights, we will generate practical recommendations for website improvements (Objective 5). We'll use IBM Cognos for data visualization (Objective 6) and implement Python for advanced analysis, including predictive modeling (Objective 7), to forecast future trends. The goal is to provide actionable insights (Objective 8) for continuous monitoring (Objective 9) to ensure sustained website performance enhancements.

Data Collection Process:

Data collection is a fundamental step in the process of working with the dataset .

1. Data Source:

- The data in this dataset likely comes from a website analytics tool, such as Google Analytics or a similar platform. These tools automatically collect data as visitors interact with a website.

2. Data Columns:

- The dataset consists of several columns, each representing a different aspect of website traffic:
 - **Row:** An identifier for each row of data.
 - **Day:** The day of the week (e.g., Sunday, Monday).
 - **Day of Week:** Numeric representation of the day of the week (e.g., 1 for Sunday, 2 for Monday).

The header features a dark background with various colorful geometric shapes including circles, rings, and a gear, some connected by dotted lines. The text 'DATA ANALYTICS' is centered in a large, white, sans-serif font.

DATA ANALYTICS

Most sought after skills in coming years
- World Economic Forum

- **Date:** The date of data collection.
- **Page Loads:** The total number of pages viewed on the website on that specific day.
- **Unique Visits:** The count of distinct visitors to the website on that day.
- **First Time Visits:** The number of visitors who visited the website for the first time on that day.
- **Returning Visits:** The number of visitors who had previously visited the website and returned on that day.

3. Data Frequency:

- It appears that the data is collected on a daily basis, with each row representing a specific date.

4. Data Accuracy and Completeness:

- Data collected by analytics tools is typically considered accurate and complete. However, data may sometimes contain anomalies or errors that require validation and cleansing.

5. Data Privacy and Compliance:

- Ensure that the data collection process complies with data privacy regulations and standards, especially if you are collecting personally identifiable information (PII) from website visitors.

6. Data Preprocessing:

- Depending on the analysis you intend to perform, you may need to preprocess the data. This can include handling missing values, converting data types, and aggregating data as needed.

7. Data Storage:

- The dataset should be stored securely and accessible for analysis. It's essential to maintain data integrity and confidentiality.

8. Data Documentation:

- Proper documentation of the dataset is crucial. This includes descriptions of each column, units of measurement, and any transformations or calculations performed on the data.

9. Data Retention:

- Depending on your organization's policies, you may need to decide how long to retain the data. This can impact storage requirements and compliance considerations.

10. Data Validation:

- Periodically validate the data to ensure its accuracy and consistency. This may involve comparing it with other sources or conducting data quality checks.

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

Create Visualizations

1. **Access the Report Authoring Tool:** Open the IBM Cognos Report Authoring Tool or web-based interface that you are using for report creation.
2. **Select or Create a Report:** Choose an existing report or create a new one, depending on your needs.
3. **Add Data:** Connect your report to the data source that contains the relevant information for your visualizations.
4. **Select Visualization Type:** Within the report, select the area or section where you want to add a visualization (e.g., a chart, graph, or table).

DATA ANALYTICS

Most sought after skills in coming years
- World Economic Forum

5. **Drag and Drop:** Drag and drop the data elements (fields or columns) from your dataset into the appropriate sections of the visualization. For example, place data elements on the x-axis and y-axis for a chart.
6. **Customize Visualizations:** Customize the appearance of your visualizations by adjusting settings such as colors, labels, legends, and tooltips. Use the properties panel to make these customizations.
7. **Apply Filters:** If needed, apply filters to the visualization to focus on specific data subsets. Filters allow you to interactively control what data is displayed.
8. **Group and Aggregate:** Group and aggregate data within the visualization to create summaries or aggregates for better insights. For instance, you can create bar charts with sums or averages.
9. **Add Interactivity:** Enhance interactivity by adding features like drill-through options or tooltips to provide users with more information when they interact with the visualization.
10. **Test and Preview:** Test your visualization within the report authoring tool to ensure it accurately represents your data and conveys the intended message.
11. **Save and Publish:** Save your report and publish it to make it available to users who need to access the visualizations.
12. **Documentation:** Document the visualization details, including data sources, customization choices, and any special considerations.

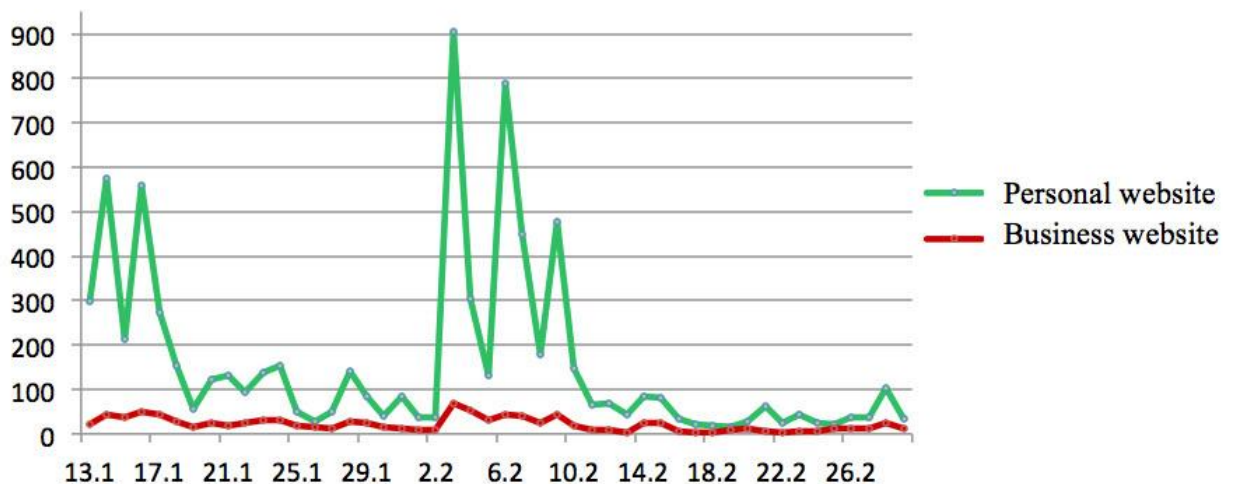


Fig 1 sample image

Python Integration:

Incorporating Python integration for machine learning into this website traffic analysis project offers a robust framework for gaining valuable insights and predicting future traffic trends and user behavior patterns.

Predictive Analytics with Time Series Forecasting:

Time series data, such as website traffic data over time, is well-suited for predictive analytics. By using Python libraries like pandas for data manipulation and prophet or ARIMA for time series forecasting, you



DATA ANALYTICS

Most sought after skills in coming years
- World Economic Forum

can develop models that predict future traffic trends. These models can capture seasonality, trends, and cyclical patterns in your data, providing you with forecasts that help you allocate resources more effectively and plan for traffic spikes or downtrends.

User Behavior Modeling:

Understanding user behavior is critical for improving the user experience and optimizing conversion rates. Python's machine learning libraries like scikit-learn and XGBoost can be employed to build predictive models. For instance, you can develop classification models to predict whether a user will convert (e.g., make a purchase or sign up) based on their historical behavior, demographics, or interaction with the website. This information can guide personalized content recommendations and marketing strategies.

Anomaly Detection:

Python offers advanced techniques for anomaly detection, which is crucial for identifying unusual or suspicious activities on your website. Using algorithms like Isolation Forest or One-Class SVM, you can detect anomalies such as fraudulent transactions or abnormal traffic spikes that may indicate a security breach or technical issue.

Personalized Content Recommendations:

Machine learning models can analyze user behavior data to make real-time content recommendations. Python-based recommendation systems, including collaborative filtering and content-based filtering, can help improve user engagement by suggesting relevant articles, products, or services based on individual preferences and browsing history.

A/B Testing and Experimentation:

Python can be used to design and analyze A/B tests to evaluate the impact of changes to your website or marketing strategies. By comparing user interactions and conversions between control and experimental groups, you can make data-driven decisions to optimize user experiences and increase conversions.

Integration with Web Analytics Tools:

Python's flexibility allows you to integrate machine learning models seamlessly with web analytics tools like Google Analytics or Adobe Analytics. This integration enables you to combine insights from both sources for a comprehensive understanding of user behavior.

Scalability and Customization:

Python's extensive ecosystem of libraries and frameworks offers scalability and customization options. You can choose the right machine learning algorithms and techniques that fit your specific project requirements and data size. Whether it's deep learning with TensorFlow or ensemble methods with scikit-learn, Python provides the flexibility to tailor your models to your needs.

Continuous Improvement:

Machine learning models can be retrained periodically to adapt to changing user behavior patterns and evolving traffic trends. This ensures that your predictive capabilities remain accurate and relevant over time.

Conclusion:

In conclusion, this website traffic analysis project aims to uncover valuable insights into user behavior, popular website content, and effective traffic sources. By understanding how users interact with the site and tailoring content and experiences accordingly, we seek to enhance overall user satisfaction and



DATA ANALYTICS

Most sought after skills in coming years
- World Economic Forum

engagement. Through data visualization, advanced analysis, and actionable recommendations, we aim to empower website owners to make informed decisions and continuously improve their digital platforms for the long term.