WEBSITE USAGE ANALYSIS

The domain of the Project:

Power BI and SQL

Under the guidance of

Ms. Siddhika (Software Engineer)

By:

Ms. Anjali A S (B.Tech CSE Graduate)

Period of the project

May 2024 to August 2025

Declaration

The project titled "**Website Usage Analysis"** has been mentored by Ms. Siddhika, organised by SURE Trust, from May 2025 to August 2025, for the benefit of the educated unemployed rural youth for gaining hands-on experience in working on industry relevant projects that would take them closer to the prospective employer. I declare that to the best of my knowledge the members of the team mentioned below, have worked on it successfully and enhanced their practical knowledge in the domain.

Name

Ms. Anjali A S

Mentor
Software Engineer—HCL Tech

Prof. Radhakumari
Executive Director & Founder
SURE Trust

Table of contents

1. Executive summary
2. Introduction
3. Project Objectives
4. Methodology & Results
5. Social / Industry relevance of the project
6. Learning & Reflection
7. Future Scope & Conclusion

Executive Summary

This project, *Website Usage Analysis*, was undertaken to extract actionable insights from website usage data and demonstrate the application of SQL, Excel, and Power BI in analysing and visualizing key performance metrics. The primary objective was to understand user behaviour, traffic patterns, and engagement levels to identify opportunities for improving website performance.

The methodology combined data cleaning and preprocessing Power Query, followed by dashboard creation in Power BI for descriptive analytics. The dataset, sourced from Kaggle, was structured to highlight traffic volume, bounce rate, session duration, and conversion metrics. Interactive dashboards provided insights into country-wise traffic, browser performance, and temporal usage patterns across days and months.

Key findings reveal that India and the USA contribute the majority of website traffic, with Chrome emerging as the most dominant browser. Engagement levels peak on Mondays and Tuesdays, and June recorded the highest monthly activity. Browser analysis highlighted Opera with the highest bounce rate, while Safari and Edge demonstrated stronger retention.

The project underscores the importance of data visualization tools for informed decision-making, showing how raw metrics can be converted into clear, actionable insights.

Recommendations include:

- Integrating real-time analytics to enable dynamic monitoring of user engagement.
- Extending the analysis to cover demographic and device-level insights for deeper audience understanding.
- Applying predictive analytics to forecast future traffic and guide proactive website optimization strategies.

In summary, this project demonstrates how leveraging data analytics and visualization can enhance website performance evaluation and support strategic decisions for digital growth .

Background and Context

Websites today generate vast volumes of interaction data that are crucial for measuring performance, improving user experience, and guiding digital strategy. Website usage analysis provides the ability to interpret this data into meaningful insights on traffic patterns, engagement, and conversions. This project leverages data analytics and visualization techniques to transform raw website data into actionable business intelligence.

Problem Statement / Goals

Organizations often struggle to make sense of large datasets, resulting in missed opportunities for improving engagement and performance. Without clear visualizations, trends and bottlenecks remain hidden. The goal of this project was to analyse website usage data, identify key patterns in user behaviour, and provide decision-makers with clear, interactive dashboards that support data-driven improvements.

Scope and Limitations

The project focused on descriptive analytics using a static dataset sourced from Kaggle. Analysis was limited to key metrics such as page views, sessions, bounce rates, session duration, browser performance, and country-wise traffic. Limitations include the absence of real-time data, lack of demographic/device-level segmentation, and exclusion of predictive analytics or forecasting models.

Innovation Component

The innovation lies in converting unstructured website usage data into intuitive Power BI dashboards that highlight KPIs across time, browsers, and geographies. By simplifying complex datasets into clear visuals, the project makes it easier for stakeholders to monitor performance, identify improvement areas, and adopt a more evidence-driven approach to website optimization.

Project Objectives

Objectives and Goals

- Systematically analyse raw website logs and aggregated usage tables to reveal who the users are, what content they consume, and how they behave. This includes calculating and profiling core metrics (page views, sessions, unique users, bounce rate, average session duration, conversion events), identifying top-performing pages and under-performing pages, and producing clear summaries that a non-technical stakeholder can act on.
- Build Power BI dashboards with an executive overview (KPI cards), trend analysis (time-series charts), segmentation controls (slicers for date range, country, browser, traffic source), and drill-through pages for detailed investigation (page-level metrics, user cohorts). Dashboards will be organized so executives see high-level health metrics at a glance while analysts can dig into causes with a few clicks.
- Implement a documented data-preparation workflow (Power Query / Excel / SQL steps) that: removes duplicates, normalizes timestamps (timezone handling), parses URLs and referrers, derives session IDs and bounce/conversion flags, handles missing values and outliers, and produces validated output files (CSV / Excel / data model). The workflow will be repeatable so future data loads require minimal manual effort.

Expected Outcomes and Deliverables

- A single .pbix containing multiple, well-labeled pages:
  Power BI dashboards providing visual insights into traffic distribution, user engagement, and browser performance.
  - Executive Overview (KPIs, top-line trend)
  - Traffic Trends (daily / weekly / monthly trends)
  - Geo & Source Analysis (country map, referral channels)
  - Browser & Device Performance (bounce, session length by browser/device)
  - Content / Page-level Analysis (top pages, exit rates)
  - Drill-through / Diagnostics (page details, filters for date/country/browser)

- Identification of peak activity periods (by country, browser, day, and month) to guide optimization strategies.
- A comprehensive project report documenting methodology, findings, and recommendations for future improvements.

Methodology and Results

1. Methods / Technology Used

The project employed a descriptive analytics approach to transform raw website usage data into actionable insights. The workflow followed these stages:

- **Data Cleaning & Preprocessing** – Using Power Query to handle missing values, remove duplicates, and format data consistently.
- **Data Modeling** – Structuring the dataset in Power BI to support relational models and efficient filtering.
- **Visualization & Analysis** – Designing interactive dashboards to present KPIs and usage patterns in a decision-friendly format.
- **Insight Extraction** – Identifying key traffic drivers, peak engagement periods, and browser performance issues through visual exploration.

2. Tools / Software Used

- **Power BI** – The primary visualization tool for building interactive dashboards and generating insights.
- **Kaggle** – Open-source platform from which the raw dataset was sourced.

---

3. Data Collection Approach

- The dataset was sourced from **Kaggle**, containing website usage metrics such as page views, sessions, bounce rates, session durations, and conversions.
- Raw data underwent a **quality check** to detect and address inconsistencies, missing values, and duplicate records.
- The cleaned dataset was imported into Power BI for dashboard creation and further analysis.

---

4. Project Architecture

The project followed a structured **four-layer architecture**:

1. **Data Collection** – Website usage dataset obtained from Kaggle.
2. **Data Preparation** – Cleaning and preprocessing performed in Excel and Power Query (duplicate removal, missing data treatment, normalization).
3. **Data Modeling** – Structuring and transforming data in Power BI, applying measures (DAX formulas) for KPIs such as bounce rate and session duration.
4. **Visualization & Insights** – Designing dashboards in Power BI with multiple views: overview, traffic trends, and browser/country analysis.

This architecture ensured a seamless flow from raw data to polished business intelligence reports.

5. Final Project Working Screenshots & Explanation

Figure 1: Website Overview Dashboard

- Displays total page views (8.37M), bounce rate (34.35%), average session duration (1m 12s), and session-level conversion analysis.
- Highlights top contributing countries: India, USA, Nigeria, and Germany.
- Shows browser distribution, confirming Chrome dominance in traffic.

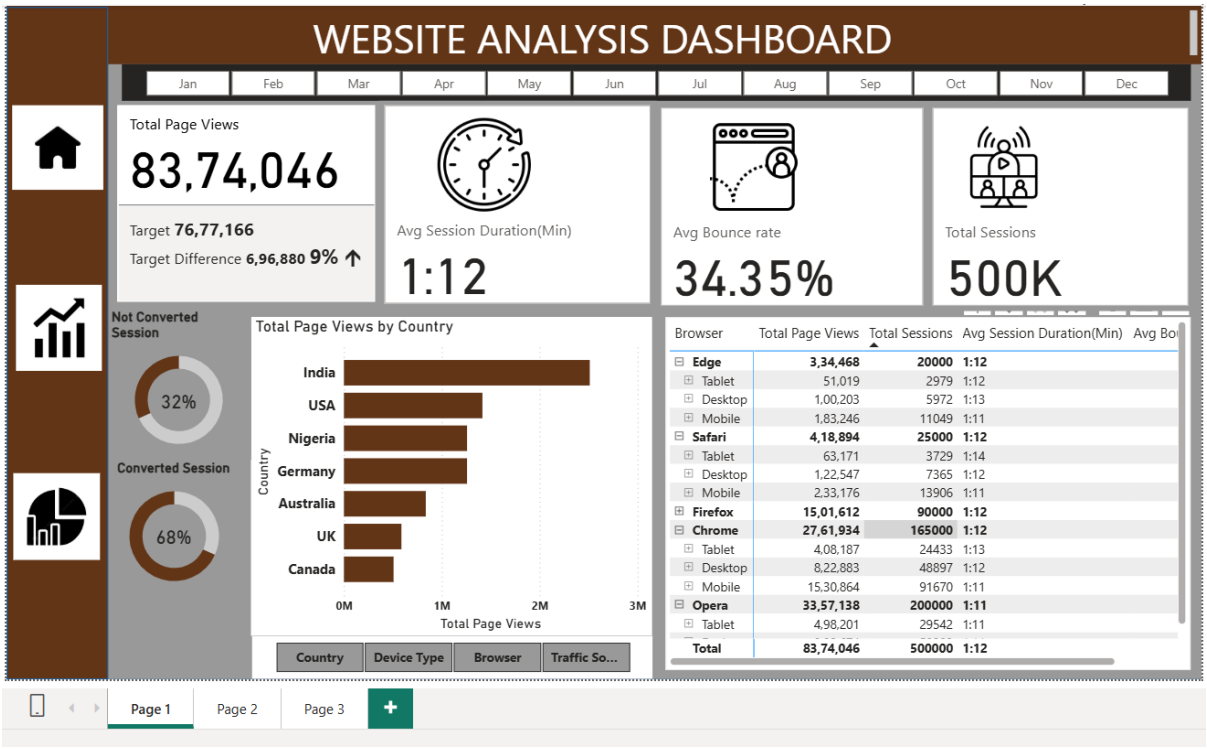


Figure 2: Traffic Trends by Date and Weekday

- Analyzes daily and weekly engagement trends.
- Mondays record the highest activity (1.22M page views), suggesting strong user engagement at the start of the week.

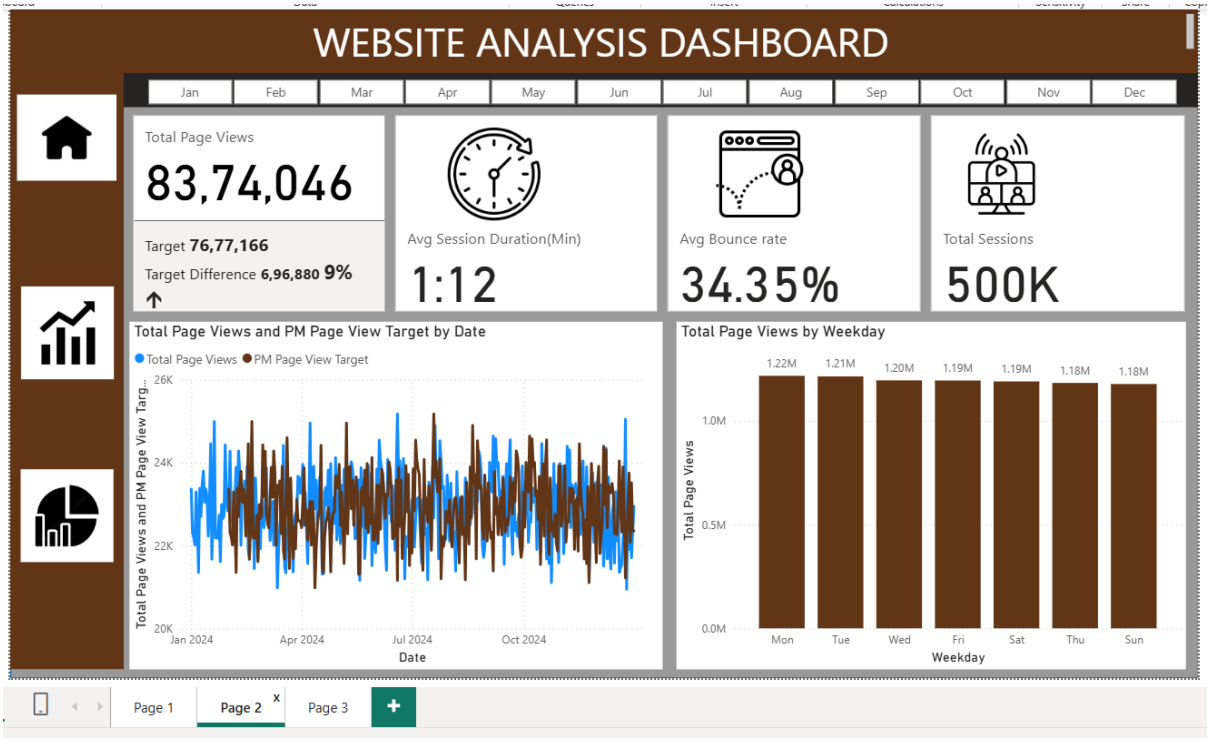
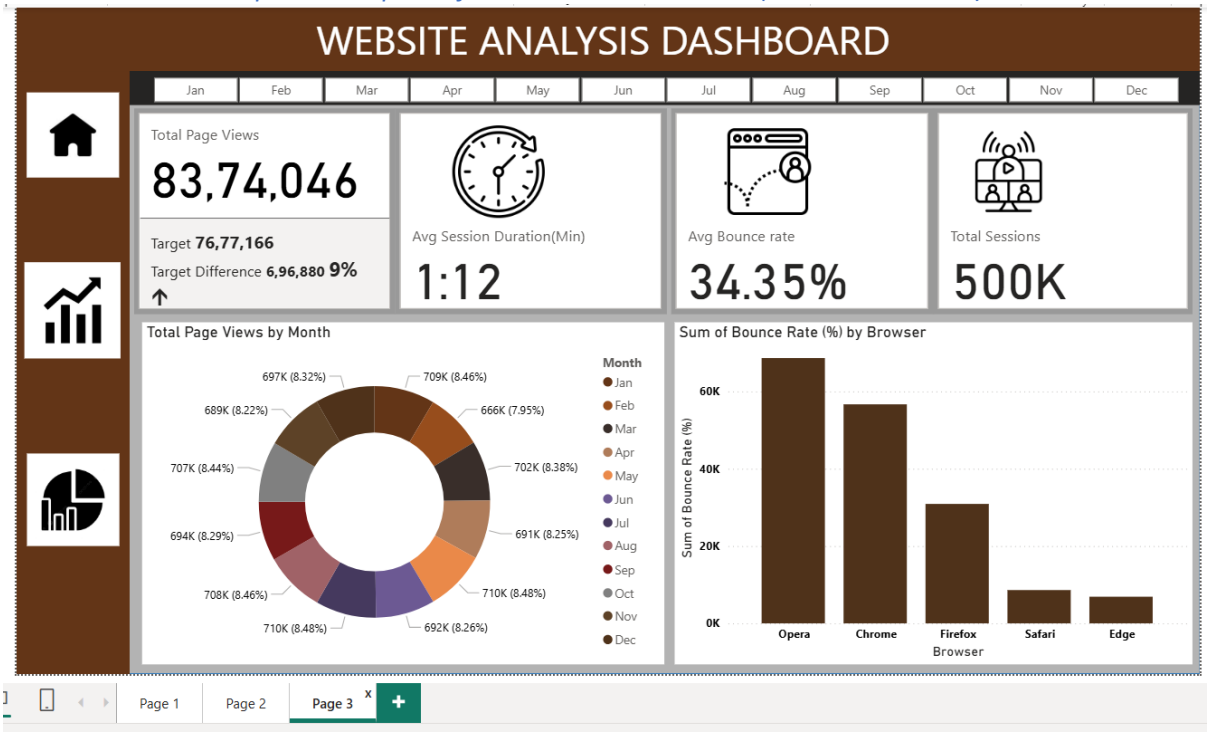


Figure 3: Monthly Trends and Bounce Rate by Browser

- Summarizes monthly activity levels, with June recording the highest engagement (710K views).
- Browser analysis revealed Opera with the highest bounce rate, while Safari and Edge showed better performance in retaining users.



These dashboards provided an intuitive and interactive platform for identifying usage patterns and making data-driven recommendations.

Learning and Reflection

Throughout the project, several valuable technical and professional learnings were achieved:

- **Technical Skills**
  - Gained hands-on experience in working with Kaggle datasets, including understanding dataset structure and preparing it for visualization.
  - Acquired strong proficiency in Power BI, including building data models, creating relationships, and designing interactive dashboards.
  - Learned to use DAX (Data Analysis Expressions) to calculate important performance metrics such as bounce rate, conversion rate, and session duration.
  - Strengthened the ability to convert raw data into clear visual insights, making patterns like traffic peaks, browser performance, and country-level engagement easily understandable.
- **Analytical & Problem-Solving Skills**
  - Improved the ability to identify patterns and trends in web traffic data (e.g., peak activity days, country contributions, and browser effectiveness).
  - Learned how to translate analytical observations into practical recommendations for website performance improvement.
- **Management & Soft Skills**
  - Enhanced time management by systematically planning and executing tasks within the project timeframe (May 2024 – August 2025).
  - Strengthened communication and presentation skills by preparing dashboards and documenting insights in a professional report format.
  - Gained exposure to working in a structured project environment, simulating industry-level expectations for reporting and deliverables.

The overall experience of working on this project was highly rewarding:

- The project provided a real-world application of Power BI, demonstrating how data visualization tools can simplify large datasets and generate actionable business insights.
- Working directly with a Kaggle dataset created exposure to open-source data platforms and reinforced the importance of clean, well-structured data in analytics.
- Designing dashboards in Power BI was both a technical and creative exercise, combining data modeling with visual storytelling.
- Challenges such as choosing the right visuals, defining the most impactful KPIs, and ensuring dashboards met stakeholder needs contributed significantly to personal growth and confidence.

Conclusion and Future Scope

The primary objective of this project was to analyze website usage data obtained from Kaggle and transform it into actionable insights through Power BI dashboards. The project successfully achieved the following:

- Built interactive dashboards that visualized critical metrics such as page views, sessions, bounce rates, session durations, conversions, and traffic distribution.
- Identified country-wise traffic contributions, highlighting India and the USA as the top sources of engagement.
- Analyzed browser performance, confirming Chrome as the dominant browser while highlighting Opera's higher bounce rate and Safari/Edge's stronger retention.
- Explored time-based patterns, revealing that Mondays had the highest engagement and June recorded the peak monthly activity.
- Demonstrated the ability to convert raw data into clear, data-driven insights that can inform website performance optimization strategies.

Through these achievements, the project not only met its stated objectives but also showcased the practical value of Power BI as a tool for business intelligence and decision-making.

Future Scope

While the project successfully delivered descriptive analytics and interactive dashboards, several opportunities exist to expand and strengthen its impact:

- Integration of Real-Time Data – Incorporating APIs or live data sources would enable continuous monitoring of traffic and engagement metrics.
- Demographic and Device-Level Insights – Adding user demographics and device usage data would provide a more comprehensive understanding of audience behaviour.
- Predictive Analytics – Applying forecasting models to predict traffic patterns and user engagement trends could help in proactive decision-making.
- Enhanced KPI Tracking – Expanding dashboards to include conversion funnels, click-through analysis, and revenue-related KPIs would improve strategic evaluation.
- Scalability for Large Datasets – Optimizing the data model and implementing incremental refresh in Power BI to handle larger and more complex datasets efficiently.
- Cross-Platform Analysis – Extending the scope to include social media and marketing analytics, linking traffic sources with user engagement outcomes.