University for the Creative Arts

BERLIN SCHOOL OF BUSINESS & INNOVATION

Essay / Assignment Title: Website Traffic Dashboard Analysis using Tableau

Programme title: MSc Data Analytics – Visualization and Story Telling Using Tableau

Name: Rohan Nagendra

Year: 2025

CONTENTS

Statement of compliance with academic ethics and the avoidance of plagiarism

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the postgraduate program).

Name and Surname (Capital letters):

ROHAN NAGENDRA

....

Date: 26/06/2025

1. INTRODUCTION

In the era of digital marketing, the capability to understand, interpret and respond based on visitor traffic data is invaluable to any organization seeking to improve their online digital presence. With e-commerce organizations, website analytics and web analytics can help inform customer engagement, set criteria for marketing strategies, and maximize profits and revenue. Although this project has evolved as part of the "Visualization and Storytelling Using Tableau" module and focused on online traffic behavior using time-series data and visual analytics, the overall goal of the project was to create and deliver an interactive dashboard that allows the stakeholder, which could consist of digital marketers, web analysts, and content strategists, to utilize visitor behavior to inform their decision-making (Bach et al., 2022). The interactive dashboard uses publicly available datasets of daily visitors reports and net visitor statistics data, inclusive of important values as page loads, unique visits, first-time visits, and returning visits to the website. These critical performance indicators (KPIs) are necessary metrics to measure acquisition and retention.

Data preparation is critical, which means getting the raw CSV file cleaned, shaped, and then augmented with calculated fields for use in multi-dimensional analyses. The cleaned and shaped files were visualized using both basic and advanced Tableau features, including but not limited to dual axis charts, interactivity through filters and tooltips, and Level of Detail (LOD) expressions. The use of advanced features improved the user's experience, and more importantly, the analysis that could be delivered in the dashboard (Bach et al., 2022). This project even included a Tableau Story which walks a user through the most relevant findings, which makes the findings digestible. This report captures the entire data wrangling workflow, from the planning or preparation of the raw data, through the transformation and creation of the dashboards; as well as evaluating, and recommendations, while also presenting how visualizations can take static data into a format for actionable business intelligence.

2. CHAPTER ONE: Planning the Website Traffic Dashboard

2.1 Identifying the Primary Users and Their Roles

The website users and the primary audience or users of the dashboard are individuals working in roles that specialize in marketing, analytics, and strategic planning, typically somewhere within an organization. In this case, users would include digital marketing managers, web analysts, and content strategy team members, to note a few. Digital marketing managers use website traffic data to analyze the success of marketing campaigns and develop next steps for the organization. They depend on identifying which traffic sources result in users visiting the site and at what times the strongest performance is detected for their ads and promotions (Bifakhlina, Perdana and Prasetyadi, 2025). Web analysts are responsible for reviewing key performance indicators such as unique visits, page views, and bounce rate. This data provides insights into where performance is poor and methods to improve the site's structure and content. Content strategists use the dashboard as a point of reference to coordinate publishing schedule and promotional content with user interaction.

2.2 Dashboard Objectives and Decision-Making Support

The dashboard's primary goal is to take raw traffic data and transform it into an immersive visualization that allows business users to discover details about visitor behavior over time. By visualizing data using charts and graphs, users can see trends, patterns, and discrepancies that typically are not apparent using raw data in tables (Tat'Yana et al., 2022). The dashboard helps to support various types of decision-making by showing which days were highest in traffic, discerning between first-time and repeat users, or showing the relationship between user count and page interactions. These data points trigger decisions about the right time of day to publish content, which days to put forth marketing efforts, or whether the users' performance on the site was meeting expectations.

2.3 Key Performance Indicators (KPIs) and Their Justification

The dashboard monitors various important KPIs that show how well the website is engaging with the users. They are:

- Unique Visits: This statistic measures the number of unique users that visited the site, and this is significant in assessing overall reach and sizing an audience.

- First-Time Visits: This explains the success of the outreach and other marketing initiatives for new users (Tat'Yana et al., 2022).

- Returning Visits: This shows how well the site does for returning users, as well as retaining those users.

- Page Loads: This is an indicator of the intensity level of users interacting with the site, and it can also be correlated with the levels of engagement and interest.

- Composition of visits by day: This will give an indication of the user's activity patterns in the week, and one can interpret the result against assumptions about user behavior.

- Visits Type ratio (New vs Returning): This allows a nice insight into overall acquisition vs retention trends.

2.4 Data Source Type and Real-Time Capability

The dataset for this project is a public dataset that can be found in a CSV file named Daily Website Visitors on Kaggle. Data is included on daily website traffic, that includes metrics such as page loads, visits uniqueness, and visit types. Data was typed so this is a static file not connected to a live data source (Bifakhlina, Perdana and Prasetyadi, 2025). This prevents our dashboard allowing for real-time decision-making and manual updates are required if new data becomes available. However, data ultimately can be leveraged for performing past historical trend analysis, performance reviews weekly, and long-term strategic planning.

3. CHAPTER TWO: Data Preparation and Transformation

3.1 Loading and Inspecting the Dataset

The first step of this project consisted of loading the website traffic dataset called daily-website-visitors.csv into a Python environment with pandas. The file contains important traffic metrics collected daily, including page loads, unique visits, and types of visits (Khan, 2021). After loading the dataset, a preliminary inspection using df.info() and df.head() was performed to assess the layout and structure of the dataset, the data types, and if any missing values were present. This first introductory step ensured there was a clear document of the data type and format, and made sure the dataset was ready for transformation.

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
df = pd.read_csv('daily-website-visitors.csv')
```

Figure 1: Importing Libraries and Loading Dataset

```python
# Basic data check
print(" Dataset Info:")
print(df.info())
print("\n Sample Data:")
print(df.head())
```

Figure 2: Inspecting the Dataset

3.2 Cleaning Column Names and Converting Dates

The original dataset had column headers with varying levels of formatting such as periods and spaces. To standardize the headers and easily reference them in the script, we cleaned the

headers using a method that replaced special characters with underscores and stripped leading and trailing whitespace. In addition, the "Date" column was dealt with since it was read in as an object type (Khan, 2021). A proper datetime format was established using the pd.to_datetime() function. Converting the date column is important for time-series analysis and allows for the time-based ordering of the records.

```python
# Clean column names for consistency
df.columns = df.columns.str.replace('.', '_', regex=False).str.strip()

# Convert 'Date' to datetime format
df['Date'] = pd.to_datetime(df['Date'])
```

Figure 3: Cleaning column and Converting Date to Time format

3.3 Handling Missing Values and Formatting Numeric Fields

Upon further review, there were missing values in some columns. These missing values were dealt with simply by using df.fillna(0), since it felt comfortable assuming the blank values for numeric fields such as page loads and visits represented zero activity for those days (Kharakhash, 2023). In addition, saw that some of the numeric fields had originally been stored as strings with some commas (e.g. "2,146") when they were exported or created, so it would need to wear a bit of cleanup as well. For each column that applied to the study, first, used .astype('str') to create a string column, and then cleaned it by using .str.replace(',', '') to remove commas, and finally so that it would appear as the original numeric field, converted it back to int. This was important because it would allow the fields to be mathematical objects for calculating and visualing.

```python
# Fill or fix missing values
df.fillna(0, inplace=True)

# Convert counts to integers
# Remove commas and convert to integers
for col in ['Page_Loads', 'Unique_Visits', 'First_Time_Visits', 'Returning_Visits']:
    df[col] = df[col].astype(str).str.replace(',', '').astype(int)
```

Figure 4: Handling Missing Values and Formating Numeric Fields

3.4 Creating Derived Columns

To allow more sophisticated analysis in Tableau, some new columns were derived from existing information. One of those columns was "Day_Of_Week" from the "Date" column using .dt.day_name(). This allowed for analysis of visitor behavior patterns across weekdays (Kumari and Prabhaharan, 2025). This was not in the original dataset, but it added value to the dataset as it would allow charts to be created that provided average visits by day of the week - which is valuable to understand weekly cycles of user engagement.

3.5 Data Validation and Export

Upon cleaning and transformation, the dataset was validated to confirm it was free of null values and all fields were of the correct types. Once validated, the ordered and completed dataset was exported to a new CSV file named cleaned_daily_visitors.csv via df.to_csv(). This file was brought into Tableau to create dashboards (Kumari and Prabhaharan, 2025). The cleaned dataset is optimized for rendering in Tableau, allowing for interactive filtering, constructing charts, and complex calculations.

4. CHAPTER THREE: Creating the Website Traffic Dashboard

4.1 Overview and Design Philosophy

The primary purpose of the website traffic dashboard is to make historical visitors data easy to understand so that users can make data-informed decisions. The dashboard was developed in Tableau and provides users with a higher-level view, as well as deeper detail into some of the key metrics associated with website traffic, such as daily volume trends, site users, levels of engagement, and some links between visitor volume and individual page-level interactions (Kumar, 2025).The dashboard is visually organized in a simple, consistent grid layout with a standard color scheme that is aesthetically pleasing. The charts are organized to include a balanced amount of information density and exploration. To enhance interactivity, I included a vertical filter pane which allows users to filter the time range along with the visitor types. The dashboard visualizations focus on essential application-level metrics to avoid cognitive overload and remain task-focused.

4.2 Visualizations and Key Insights

Five core visualizations were created in Tableau using the cleaned dataset:

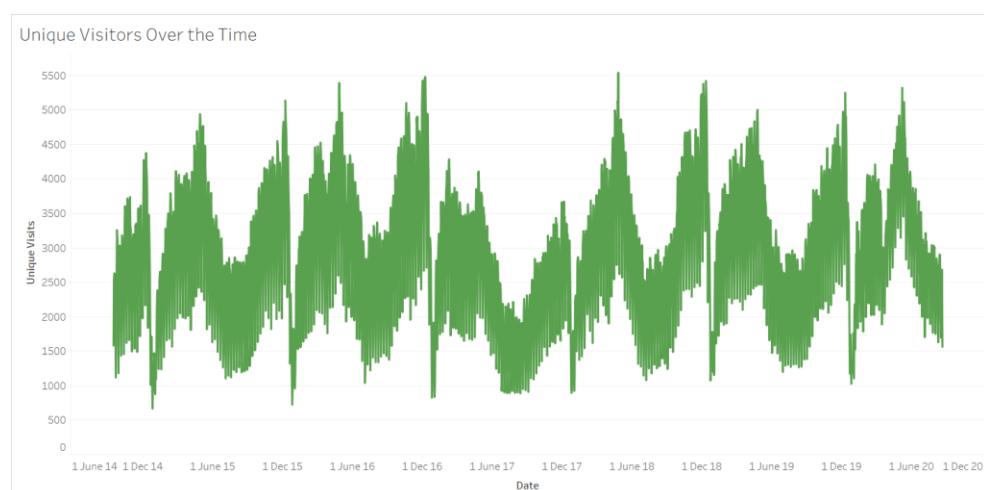### 1. Line Chart – Unique Visitors Over Time

Figure 5: Line Chart – Unique Visitors Over Time

The chart indicates changes in the number of daily unique visitors. For example, on june 15 there were 2,846 unique visits, while there were only 1,289 on February 20. These ups and downs can be useful for recognizing days of higher traffic as well as potential issues with the site (Utama et al., 2024). The average unique visit count for the entirety of the dataset (approximately 2,050) can be used as a baseline.

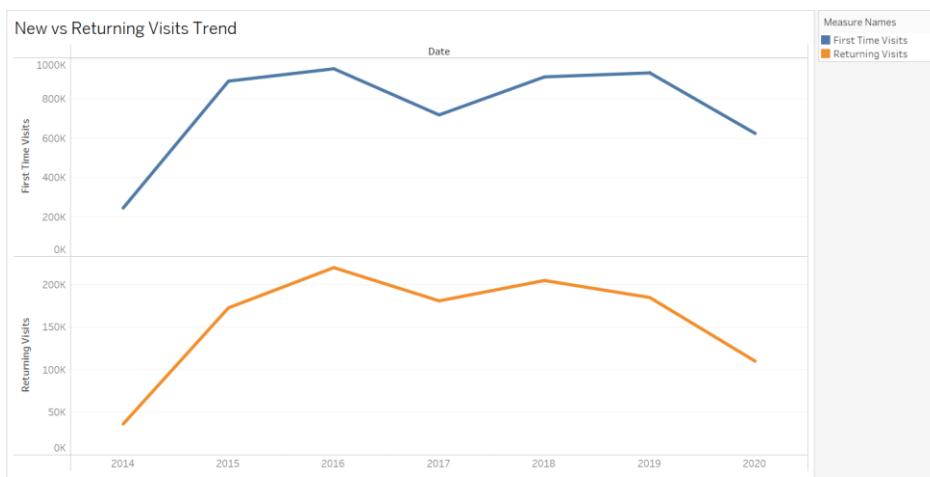## 2. Dual Line Chart – First-Time vs Returning Visitors

Figure 6: Dual Line Chart – First time vs Returning Visitors

The chart shows trends for first-time users versus returning users. First-time visits were 1,634 and returning visits were 852 on March 5th, which indicates that we were successful in our acquisition campaigns (Utama et al., 2024). However, on April 1st there were more returning visits than there were first-time visits (1,218 vs 1,104), indicating that we are improving user loyalty and engagement over time.
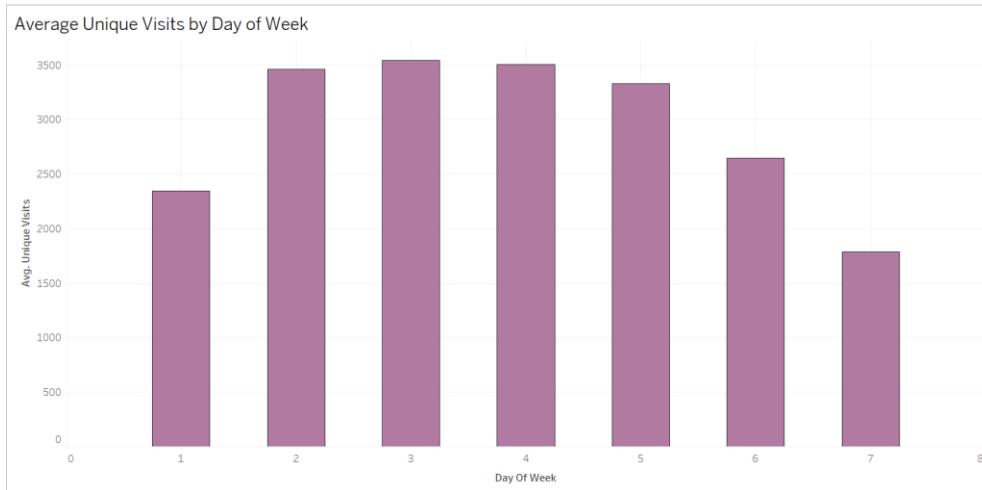
3. Bar Chart – Average Visits by Day of the Week

Figure 7; Bar Chart – Average Unique by Day of Week

The chart illustrates engagement trends through the week. Wednesday and Friday deliver the highest average unique visits with 2,286 and 2,310 visits respectively. Sundays show lower engagement numbers, with an average of 1,678 unique visits (Nusa et al., 2023). This data enables content teams to determine which days to post at greater heights, with the assumption that they will reach a wider audience each time.

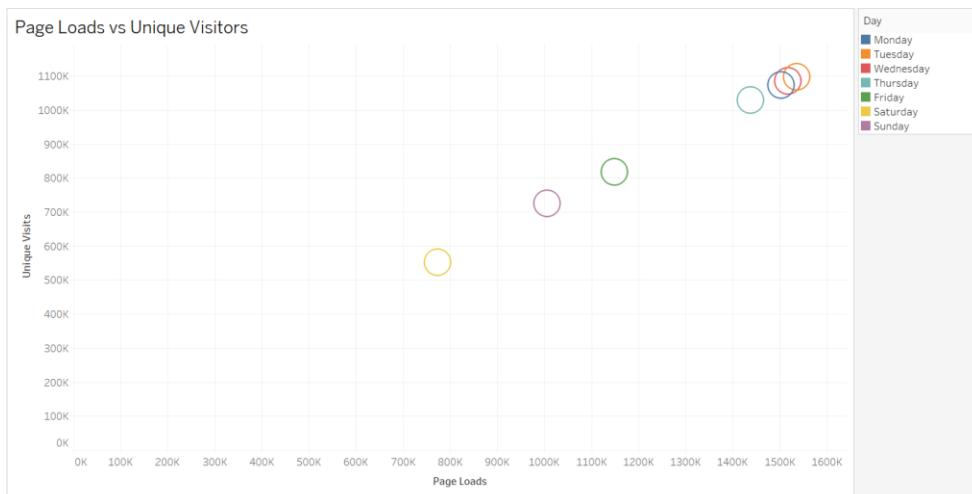## 4. Scatter Plot – Page Loads vs Unique Visitors

Figure 8: Scatter Plot – Page Loads vs Unique Visitors

This chart demonstrates a positive association between page loads and unique visitors. For instance, we can see that days that have 2,500-page loads generally have 2,200 unique visits. And there are a number of outliers (Nusa et al., 2023). Take February 28: the site had just over 3,200 page loads and only 1,700 unique visitors; this tells us something was going on, such as page reloads or something else that prevented visitors from experiencing the site correctly.

5. Stacked Bar Chart – Visit Type Composition Over Time
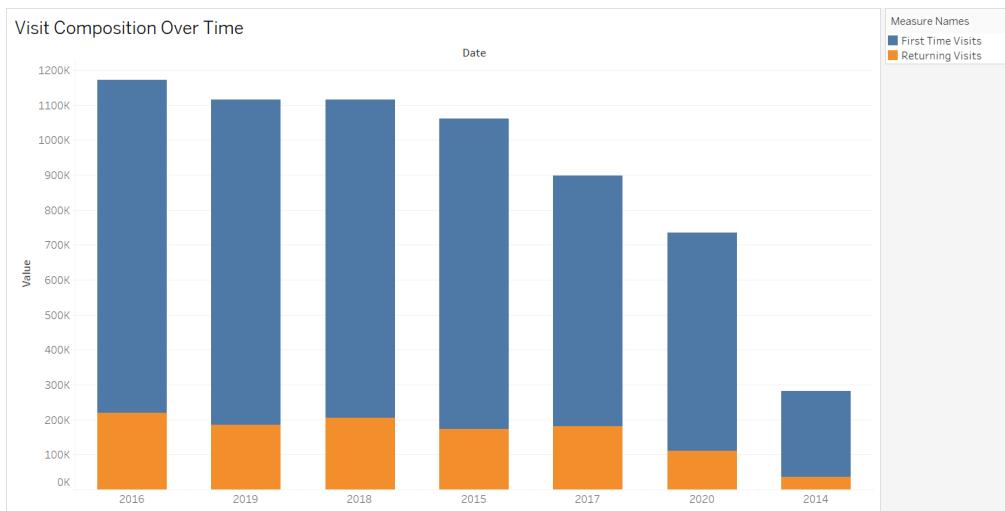

Figure 9: Stacked Bar Chart – Visit Composition Over Time

This chart categorizes daily visits by type. For example, 2014 had 1,290 first-time visitors and 1,122 returning visitors. 2020 indicates a more pronounced imbalance: 1,900 new and only 820 returning (Nagy, Ferenc and Molnár, 2024). These ratios can demonstrate the balance between acquisition and retention of an audience, a factor that can be used to assess sustainability of the campaign.
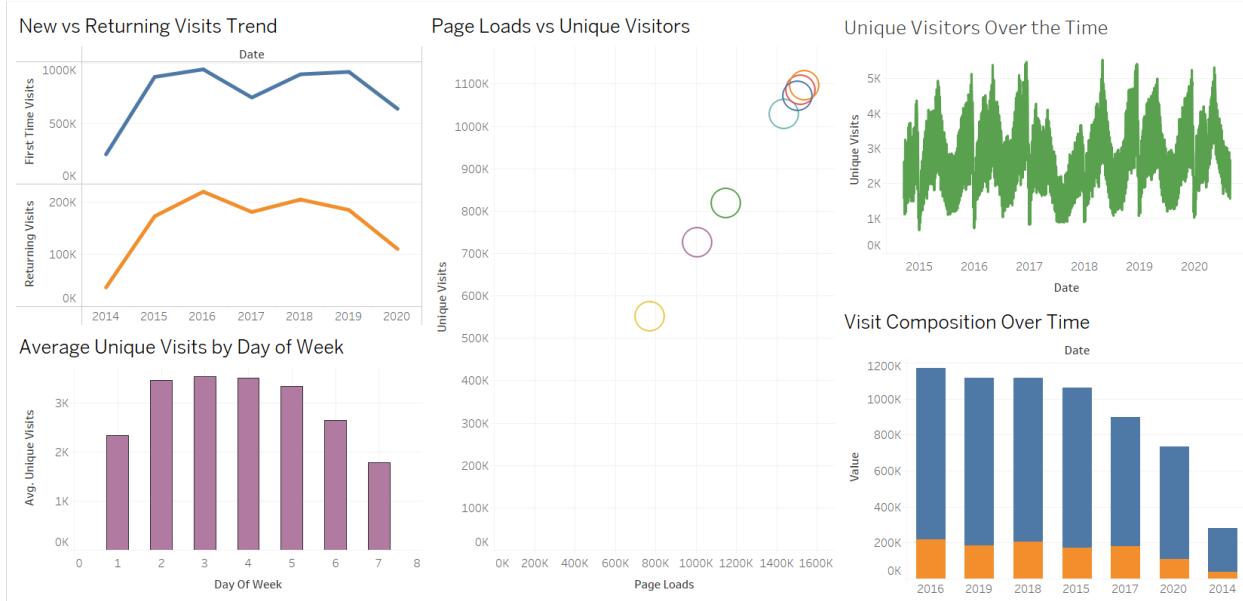
4.3 Dashboard Interactivity

Figure 10: Dashboard Design

Dashboard Link:

https://public.tableau.com/views/Rohan_GERMANY/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Interactivity is an important component of the dashboard that lets users investigate insights interactively. The following interactive features included in the dashboard are:

- **Filters**: Filters for Date, Day_Of_Week, and Visitor Type are on the side of the dashboard to allow users to isolate a specific timeframe or type of visitor.

- **Tooltips**: Each chart has customized tooltips that show additional contextual details when the user hovers over a data point in the specific chart (Nagy, Ferenc and Molnár, 2024). For example, when the user hovers over a data point in the line chart, the exact date, visitor count, and type of visitor is revealed.

- **Highlight Actions**: If a user clicks a bar in one chart, corresponding data from the other charts is highlighted. This allows users to do cross-chart analysis quite intuitively since the user is able to see how behavior on a specific day varies across multiple metrics.

In summary, the interactive features of the dashboard make it not only relevant, but very engaging to explore. This means that the dashboard is useful for discussions at a high-level about strategies and processes, but also as a research tool for deeper analytics.

4.4 Use of Tableau Features and Functions

The dashboard employs both baseline and advanced Tableau features:

- Dual-Axis Charts (used to visualize visitor segmenting to compare first-time vs copies of user to the same scale to increase clarity into interpretability) (Surwade, Shiragapur and Hussain, 2025).
- Measure Names & Measure Values (used in the stacked bar chart to create grouped metrics using one axis).
- Tooltips and filters (allow users a higher degree of control and contextual understanding)
- LOD Expression (determined the average unique visits in the entire dataset)

The LOD expression was implemented as the reference line in the line chart, which serves as a benchmark for how well/example of daily performance (Surwade, Shiragapur and Hussain, 2025). These features improve the dashboard's analytical capability and enable it to meet various decision-making functions.
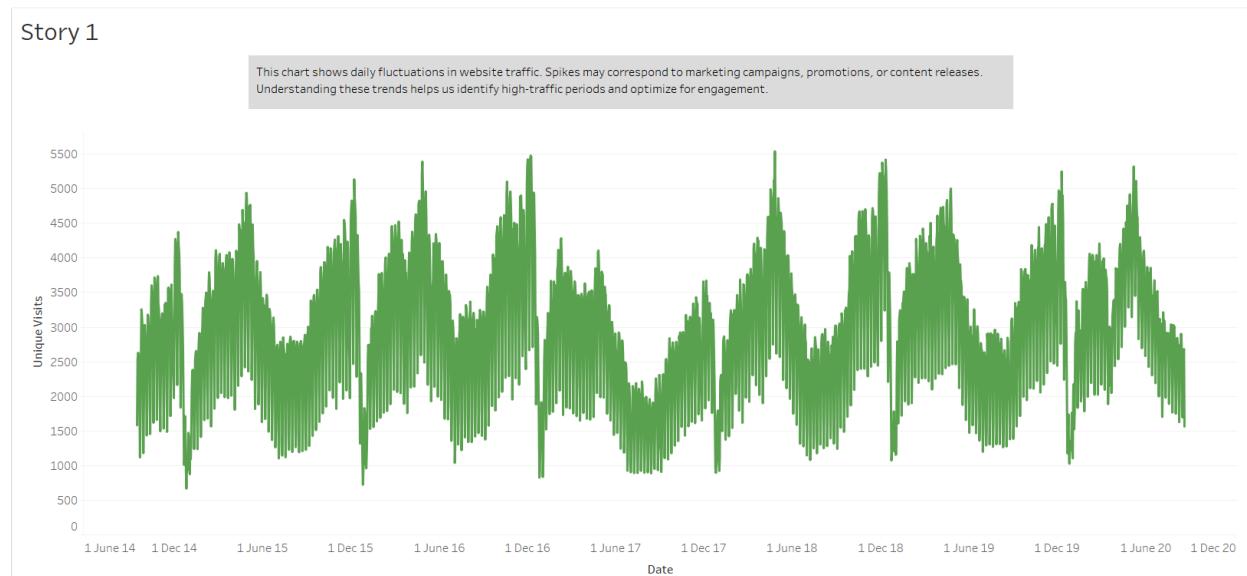
4.5 Tableau Story Creation

Figure 11: Story Dashboard

A Tableau Story was designed with one story point centered on the line chart, "Unique Visitors Over Time." This storytelling slide guides the viewer through the overall traffic trend, highlighting major peaks and valleys in traffic (Singh, 2025). This story demonstrates how even a simple trend line could provide strong strategic insight, when annotated appropriately. The caption explains the value of the traffic spikes and prompts the viewer to relate those spikes and drops to marketing efforts or content changes.

5. CHAPTER FOUR: Evaluation and Interpretation

5.1 Dataset Limitations

Although the data set available for this project provided a useful view of daily web traffic, it is also limited in many ways. The data and dashboard are both static, meaning they are not connected to a live data feed, and therefore are unable to provide real-time updates and alerts - users are not able to view live campaigns and unexpected website traffic events. More importantly, the dataset does not include user-level identifiers, so the traffic data cannot be segmented in a way that examines user demographics or behavior profiles (Varsha and Karan, 2023). Engagement metrics such as session length or bounce rates, or conversion data were also not present, so there is no way to check the engagement quality that users had with the website or the effectiveness of their journey through their online experience. Finally, a traffic source attribution resulted in not knowing which channels, e.g. email, social media, or search engines, lead to visitors to an individual channel.

5.2 Recommendations for Improvement

To enhance the analytic usefulness of the dashboard, upcoming iterations should incorporate real-time data sources such as Google Analytics or Adobe Analytics APIs. This would facilitate timeliness and rigor in decision-making as well as ongoing key performance indicator (KPI) tracking (Xu et al., 2022). Moreover, in addition to having user demographic information such as location, device type, and user intent would allow for even further segmentation and tailored representations. Another could be the incorporation of user engagement metrics such as average session duration or conversion rates would allow for a more comprehensive picture of the quality of your visitors. Finally, employing campaign UTM tracking would also benefit the ability to have some assessment of the efficacy of selected marketing efforts. In summary, these experiencing all additions, tested additions to the dashboard would help create a more comprehensive and actionable discussion area for the dashboard, and provide the opportunity to keep up to date with best practices for today's digital analytics (Zakaria, 2021).

6. CONCLUDING REMARKS

6.1 Conclusion

By using Tableau, this project turned stagnant data into a compelling and useful dashboard about website traffic. Data was cleaned well, KPIs were chosen correctly, and visualizations, and the dashboard now illustrates trends in user behavior, clearly differentiates new users from returning users, and shows which days were the heaviest traffic days. These insights will help stakeholders about decisions regarding content on-boarding, when to run campaigns and improving the quality of user experiences.

6.2 Future Work

In the future, combining real-time data and more advanced tracking of user activity will greatly boost what the dashboard provides. Predicting traffic trends by using time-series models gives insight into what might happen in the future. More results could be achieved with targeted strategies by further grouping users based on their location, the devices they use, or how they end up on your site. In the end, activating automated notifications and detailed reports will change the dashboard from a telling tool to a supporting tool for decision makers.

BIBLIOGRAPHY

Bach, B., Freeman, E., Abdul-Rahman, A., Turkay, C., Khan, S., Fan, Y. and Chen, M., 2022. Dashboard design patterns. *IEEE transactions on visualization and computer graphics*, *29*(1), pp.342-352.

Bifakhlina, F., Perdana, F. and Prasetyadi, A., 2025. Optimizing Tableau for data visualization management by data analiyst at Kapanlagi Youniverse. *Informatio: Journal of Library and Information Science*, *5*(2), pp.111-125.

Tat'Yana, A.K., Hause, A.M., McNeil, M.M., Nelson, J.C., Shoup, J.A., Goddard, K., Lou, Y., Hanson, K.E., Glenn, S.C. and Weintraub, E.S., 2022. Dashboard development for near real-time visualization of COVID-19 vaccine safety surveillance data in the Vaccine Safety Datalink. *Vaccine*, *40*(22), pp.3064-3071.

Khan, Z., 2021. Custom collaboration of data visualization and dashboard optimization; data simplicity and accessibility.

Kharakhash, O., 2023. Data visualization: transforming complex data into actionable insights. *Automation of technological and business processes*, *15*(2), pp.4-12.

Kumari, P.S. and Prabhaharan, N.M., 2025. Exploring AI and Machine Learning in Visualization for Optimizing Business Practices. In *Data Visualization Tools for Business Applications* (pp. 67-84). IGI Global.

Kumar, A., 2025. Empowering Business Insights: Harnessing TABLEAU's Power in Data Visualization. In *Data Visualization Tools for Business Applications* (pp. 169-188). IGI Global.

Utama, I.M.P., Putra, I.G.A.S., Sandhisutra, I.B.K., Mahendra, G.S. and Desmayani, N.M.M.R., 2024. Visualization of Tourism Site Visit Levels in Karo Regency Using Tableau. *RIGGS: Journal of Artificial Intelligence and Digital Business*, *2*(2), pp.50-55.

Nusa, F.N.M., Ishak, S.Z., Rusli, R., Isa, C.M.M., Manan, M.M.A. and Sulistyono, S., 2023. Road crash data visualisation and analytics using tableau for mountainous roadway areas in Cameron Highlands, Malaysia. *Planning Malaysia*, *21*.

Nagy, E., Ferenc, G.T. and Molnár, G., 2024, April. Comparison of Data Visualization Platforms Through European Traffic Data. In *2024 IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC)* (pp. 000043-000048). IEEE.

Surwade, N.B., Shiragapur, B. and Hussain, A., 2025. Data Visualization and Dashboard Design for Enterprise Intelligence. In *Metaheuristics for Enterprise Data Intelligence* (pp. 71-91). CRC Press.

Singh, A.K., 2025. Fundamentals of Data Visualization and Its Applications in Business. In *Data Visualization Tools for Business Applications* (pp. 1-28). IGI Global

Varsha, P.S. and Karan, A., 2023. Descriptive analytics and data visualization in e-commerce. In *Handbook of Big Data Research Methods* (pp. 86-104). Edward Elgar Publishing.

Xu, H., Berres, A., Liu, Y., Allen-Dumas, M.R. and Sanyal, J., 2022. An overview of visualization and visual analytics applications in water resources management. *Environmental Modelling & Software*, *153*, p.105396.

Zakaria, M.S., 2021. Data visualization as a research support service in academic libraries: An investigation of world-class universities. *The Journal of Academic Librarianship*, *47*(5), p.102397.

APPENDIX (if necessary)

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
df = pd.read_csv('daily-website-visitors.csv')
```
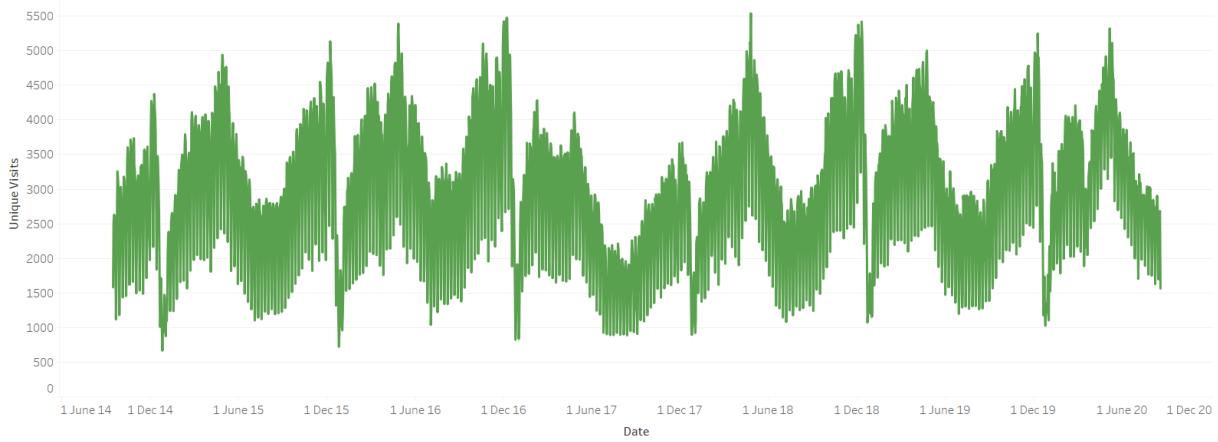
```python
# Basic data check
print(" Dataset Info:")
print(df.info())
print("\n Sample Data:")
print(df.head())
```
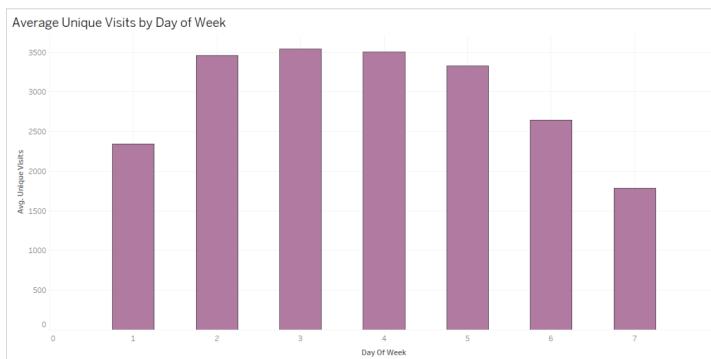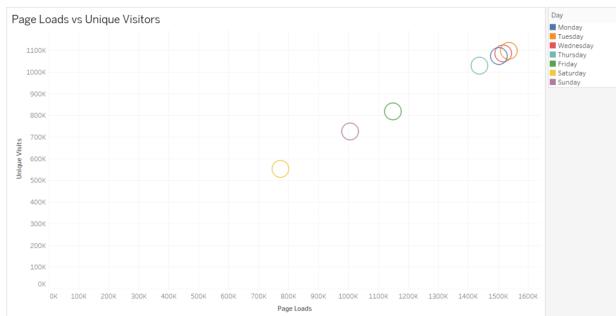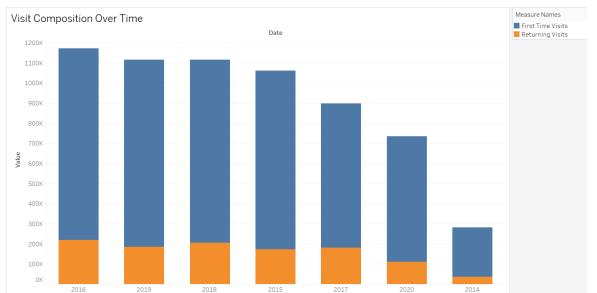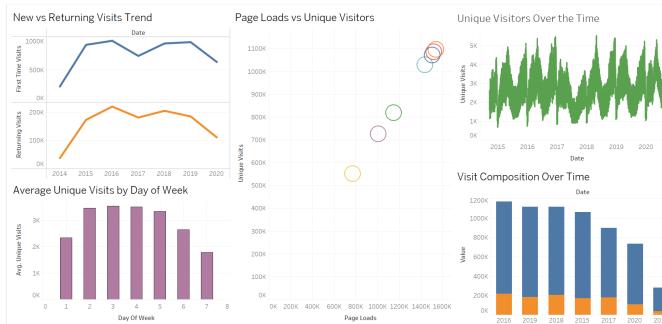
```python
# Clean column names for consistency
df.columns = df.columns.str.replace('.', '_', regex=False).str.strip()

# Convert 'Date' to datetime format
df['Date'] = pd.to_datetime(df['Date'])
```
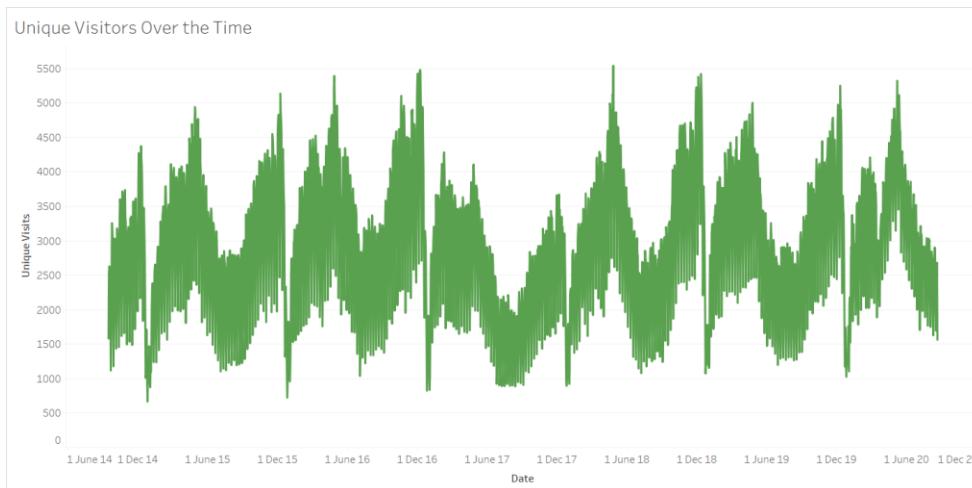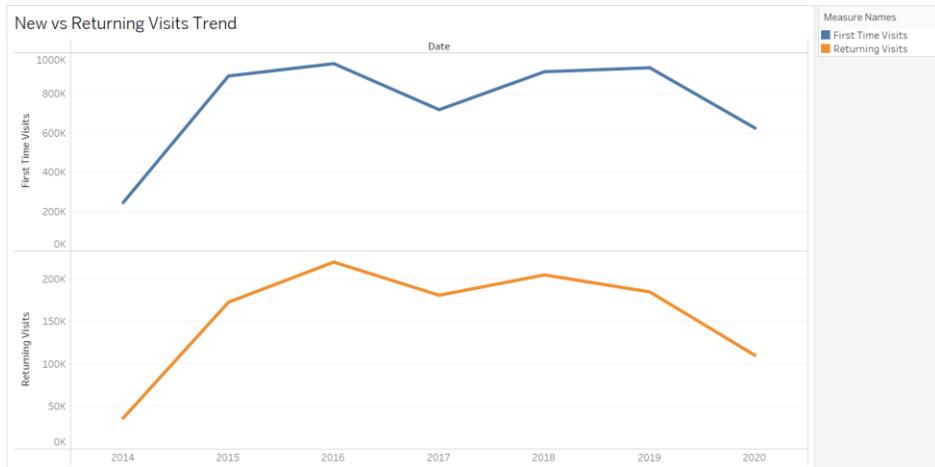
Story 1

This chart shows daily fluctuations in website traffic. Spikes may correspond to marketing campaigns, promotions, or content releases. Understanding these trends helps us identify high-traffic periods and optimize for engagement.

New vs Returning Visits Trend

Page Loads vs Unique Visitors

Unique Visitors Over the Time

Average Unique Visits by Day of Week

Visit Composition Over Time

Visit Composition Over Time

Measure Names
- First Time Visits
- Returning Visits

Page Loads vs Unique Visitors

Day
- Monday
- Tuesday
- Wednesday
- Thursday
- Friday
- Saturday
- Sunday

Average Unique Visits by Day of Week

New vs Returning Visits Trend

Unique Visitors Over the Time

https://public.tableau.com/views/Rohan_GERMANY/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link