Real-Time Twitter Analytics Dashboard - Power BI

Internship Report

Submitted By :

Anushka Bharde

Table of Content

1. Introduction

One of the most widely used sites for exchanging updates, thoughts and information is twitter. People, companies and organizations use it extensively to communicate with their audience, share thoughts and promote content. Because of its distinctive qualities, the platform is a valuable source of data that can be examined in order to understand user patterns, preferences, and behavior. This project's objective is to develop a Twitter Analytics Dashboard using Power BI. The objective is to use graphs, charts, and filters to transform Twitter data into understandable insights. Popular hashtags, user engagement levels, and tweet performance are just a few examples of the useful and actionable data that the dashboard offers. In order to provide users more control over the data they wish to study, it also has capabilities like time-based displays and custom filters. Businesses and analysts may use this dashboard to spot trends, evaluate the effectiveness of their content, and refine their approaches to engage their audience more effectively. The project aims to simplify Twitter analytics, saving time while enabling smarter decision-making.

1.1 Twitter: A Unique Platform for Insights

Twitter is not just a platform for sharing updates; it is a digital space where millions of conversations happen every day. It is a useful source of structured data because of its simple nature, which enables users to express their ideas rapidly. The platform provides various elements, such as hashtags, mentions, and engagement metrics like likes and retweets, which make analyzing trends and user preferences easier. Businesses and analysts can use this data to track what topics are popular, monitor audience engagement, and evaluate the success of their campaigns. Twitter's open API makes it possible to access and process data efficiently. By focusing on these features, this project creates a dashboard that transforms Twitter's data into clear, interactive, and actionable visualizations. The aim is to help users draw insights and use them to improve social media strategies and content performance.

2. Background

2.1 Dataset Overview

A thorough understanding of tweet performance is offered by the dataset used in this analysis, which includes important measures including impressions, engagements, engagement rate, retweets, replies, likes, and interactions like profile clicks, URL clicks, and media views. Time-based analysis is made possible by the inclusion of each tweet's timeframe. For additional analysis, text-related metrics such as word and character counts were also retrieved. Preprocessing the dataset with Power BI's transformation tools allowed for the creation of calculated columns for deeper insights, the calculation of engagement rates, and the filtering of irrelevant data. This dataset enables us to determine which tweets perform the best and comprehend how various aspects influence how users interact with tweets.

2.2 Power BI

The primary tool for viewing and analyzing the data was Power BI, which provided dynamic and interactive features that enabled more in-depth understanding. It simplified working with the raw data easier by cleaning and preprocessing it. Important measures, such as engagement rates and performance comparisons, were calculated using specially designed columns and measurements. More specialized analysis was made possible by Power BI's advanced filtering features, which allowed for the exclusion of particular content or the focus on tweets with particular attributes.

In order to monitor trends, showcase the best tweets, and examine viewer behavior, interactive dashboards were developed. The dashboards, which illustrate how many aspects affect engagement, were made to be simple to understand. In addition, by embedding the dashboards on websites or sharing them via unique URLs, Power BI's Publish to Web functionality made it simple to publish these insights publicly, allowing anybody to view the results without requiring a Power BI license. By offering useful information, this made it possible to develop more effective content strategies and raise audience engagement.

3. Learning Objectives

Exploring how to effectively analyze and present twitter performance using Power BI was the learning objective. This required mastering fundamental skills like creating charts, writing DAX queries, and data preprocessing. Learning how to convert unprocessed data into a readable format was part of the process. This required setting up important measures like engagement rates, retweets, and likes, filtering out unnecessary data, and generating calculated columns. Learning how to make interactive charts and dashboards to represent trends and engagement patterns was another primary focus of the report. Engagement metrics were visualized using a variety of chart types, and important variables like average engagement rates and total engagements were computed using DAX queries.

Additionally, the ability to use PowerBI's publish to web tool to improve user accessibility was acquired; as a result, embed codes can be created, which facilitates sharing reports via links or embedding them so that users can view them without PowerBI licenses. This approach enhanced comprehension and made it possible to analyze data in real time for quicker decision-making.

Creating dynamic charts, writing the Dax queries, preprocessing data, and sharing insights via the publish to web function were just a few of the many features demonstrated in this project. Developing proficiency with these features allowed the project to showcase Power BI technical expertise while also enabling better decision-making and collaboration by giving rapid access to insightful information. These skills helped to transform unstructured data into meaningful interactive insights that were easy to share as well as understand by providing quick access to insightful information.

4. Activities and Tasks

Making visualizations that analyze Twitter data on usage with particulate filters is one of the goals and activities for this report. While making sure the outcomes are relevant based on time and content criteria, these tasks aim to point out important data including rates of engagement, retweets, and tweet measures. By using time-based and content-based filters, the objective is to enhance the data display and offer actionable insights.

4.1 Task 1

The first task involved developing a chart to display the tweets with the highest engagement rates, specifically those in the top 10%. The analysis was restricted to tweets that received more than 50 likes and were posted on weekdays. The graph was designed to only show data between 3 PM IST and 5 PM IST. Outside this time range, the graph would be hidden from the dashboard. Additionally, only tweets with a character count of less than 30 were included in the chart.

4.2 Task 2

The second task involved creating a chart to highlight the top 10 tweets, ranked by the total sum of retweets and likes. Tweets posted on weekends were excluded, and the user profile who posted each tweet was also included. Similar to the first task, the graph was restricted to only show data between 3 PM IST and 5 PM IST. Any data outside this time window was removed from the dashboard. Additional filters were applied to the data, such as excluding tweets with an even number of impressions, tweets posted on odd-numbered dates, and tweets containing fewer than 30 words.

4.3 Task 3

The third task required creating a line chart that shows the trend of the average engagement rate for each month of the year. The chart differentiated between tweets with media content and those without. This chart was only visible on the dashboard between two specific time slots: 3 PM IST to 5 PM IST and 7 AM to 11 AM. Similar to the other tasks, the data was filtered to ensure that the tweet engagement was an even number, the tweet date was odd, and

the tweet character count was less than 20. Additionally, tweets containing the letter "C" in their text were excluded.

4.4 Website

Another part of the assignment was to create a website that displayed Power BI reports using the "Publish to Web" capability. This solution avoids the need for a Power BI license by allowing users to view and interact with insights directly from the website. To ensure complete functioning and accessibility, the reports were integrated and published on Netlify's platform. The website was designed to be fully responsive to all mobile devices, allowing for seamless usability and improving the overall user experience.

5. Skills and Competencies

5.1 Power BI Expertise

The project required expertise in Power BI, particularly in using DAX query functions to transform data and selecting appropriate chart kinds and measures to guarantee optimal data portrayal. Understanding which chart type would best represent the data was critical for clear and accurate visualization.

5.2 Data Preprocessing and Transformation

Preprocessing data was a crucial ability that was acquired throughout the endeavor. Data had to be cleaned and filtered in order to guarantee that only the necessary data was taken into account for analysis. Strategies like addressing missing data, generating computed columns, and eliminating unnecessary data were applied. More accurate insights were made possible by these preprocessing abilities, particularly when working with big datasets that had a variety of characteristics.

5.3 DAX Query Writing

Another crucial skill acquired during the project was the ability to design advanced DAX queries. Custom computations, including character count, time validation, and averages, were made possible via DAX queries. This ability was crucial for customizing the data analysis to the report's particular requirements and producing insights into the ways in which various elements, such as tweet timing, content, and engagement, affected the outcomes.

5.4 Web Development for Embedding Reports

Basic web programming abilities were employed to incorporate reports into a website in order to make the insights more accessible. This made it possible to distribute the interactive information to a larger audience and guaranteed that users could readily view the data. Integrating the reports into a neat and user-friendly website needed basic web programming skills.

6. Feedback and Evidence

Since the project was completed independently, there were no external reviews or feedback from mentors or supervisors as mentor support was not available during the internship. However, the tasks were thoroughly self-assessed throughout the process to ensure quality and accuracy. The successful completion of the tasks, including data preprocessing, chart creation, DAX query writing, and embedding Power BI reports in the website, serves as evidence of the project's success. Additionally, the interactive dashboards and website functionality were tested and verified, ensuring that they meet the defined objectives of providing actionable insights.

7. Challenges and Solutions

7.1 Challenge in Task 1: Filtering Tweet Character Count

One of the challenges encountered in Task 1 was the absence of records in the dataset where the tweet character count was below 30, which meant the required visualizations could not be displayed. To overcome this issue, the tweet character count was manually adjusted in the dataset to meet the specified condition. This ensured that the visual data aligned with the task requirements and allowed the visualizations to be successfully generated.

7.2 Challenge in Task 3: Filtering Tweet Character Count and Removing Words Containing 'C'

In Task 3, similar issues were encountered where the tweet character count was required to be below 20, but the dataset did not contain any records matching this condition. To resolve this, the tweet character count was manually adjusted to meet the specified requirement. Additionally, the task required removing tweet words containing the letter 'C', but as the dataset did not have such entries, the word removal was applied to words containing 'c' instead. These adjustments allowed the visual data to be displayed correctly and meet the task specifications.

7.3 Challenge in Time-Based Filtering

Another challenge occurred when attempting to implement the time-based filter for data to be displayed only between 3 PM IST to 5 PM IST or 7 AM to 11 AM. Initially, the filter did not function as required, causing issues with the accuracy of the data displayed.After some research, the correct DAX code was created, using the TODAY() function to ensure the visualizations only show data within the right time frames. This solution fixed the issue, making the data display correctly during the specified hours.

7.4 Challenge in Accessing "Publish to Web" Feature

A challenge was faced when trying to use Power BI's "Publish to Web" feature with a college account presented a problem because it lacked the required permissions. The reports could not be shared online without this access. This was resolved by creating a trial Microsoft 365 account, which gave the necessary access. This made it possible for the reports to be published and included on the website, enabling users to obtain the results. The inability of Power BI's "Publish to Web" feature to provide dynamic time-based filtering presented another difficulty for the project. This feature simply publishes a static version of the report as it appears at the time of publication, which causes issues when trying to implement time validation charts. To overcome this limitation, different approaches were tried, but in the end, JavaScript was used on the website to apply time validation.

8. Outcomes and Impact

The outcome of the project, highlighting how the use of Power BI visualizations and the development of the website helped in gaining valuable insights. The interactive dashboard helped reveal important patterns in Twitter engagement, such as identifying tweets with high engagement rates, tracking trends over time, and understanding how different types of content, like media, influence user interaction. By applying specific filters, such as tweet character count and engagement metrics, the visualizations ensured that only the most relevant data was displayed, making it easier for users to focus on the tweets that performed best and at the right times.

A website was created which has reports from the "Publish to Web" feature to embed Power BI dashboards for easy access and hosted on Netlify. To guarantee that it functions flawlessly on all devices, the website was made responsive, giving a larger audience the opportunity to observe and engage with the insights without requiring power bi licensing. Users were able to view trends and enhance their strategies because of this. The website and visuals were essential for efficiently sharing the results.
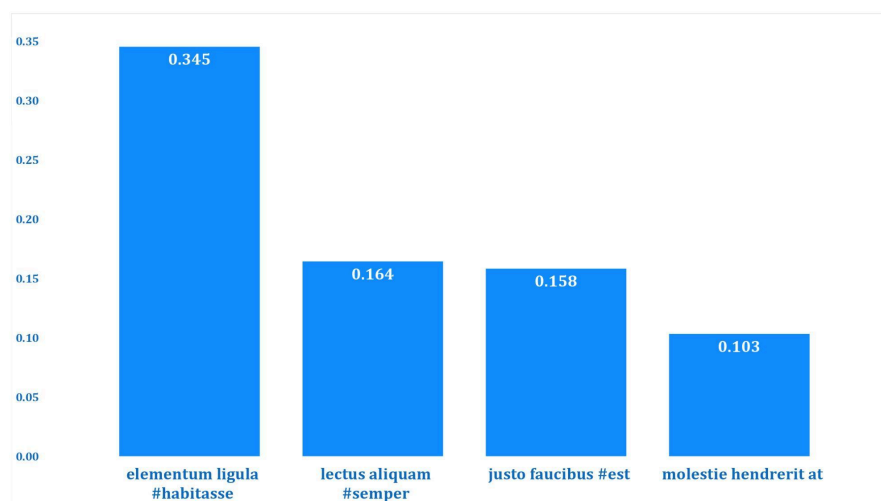

Fig. 8.1 Task 1: Highest Engagement Tweets (Top 10%)

The column chart for Task 1 shows the top 10% of tweets with the highest engagement rates, with an emphasis on those sent on weekdays between 3 and 5 PM IST. Tweets with fewer than 30 characters and more than 50 likes are included. The tweet "molestie hendrerit at" has

the lowest engagement rate (0.103), while "Elementum ligula" has the highest (0.345). Based on the given criteria, this chart displays the tweets with the highest interaction, offering insightful information.
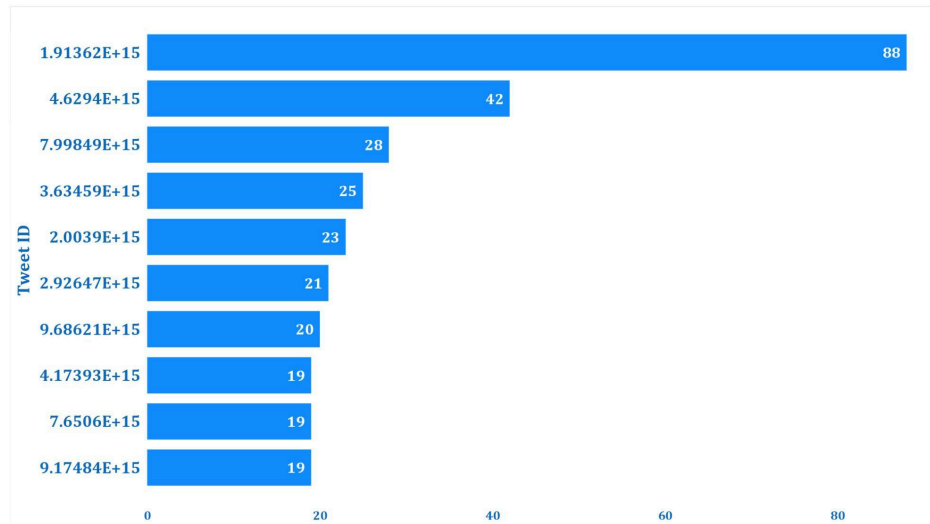

Fig. 8.2 Task 2: Top Tweets by Retweets and Likes

For task 2, the bar chart shows the top 10 tweets based on the sum of retweets and likes, filtered to display tweets posted on weekdays between 3 PM and 5 PM IST. It includes tweets with an even number of impressions, an odd tweet date, and a word count under 30. The chart highlights the tweet with ID 1.91362E+15, which has 88 views, while others have fewer views. This chart provides insights into the best-performing tweets under these conditions.
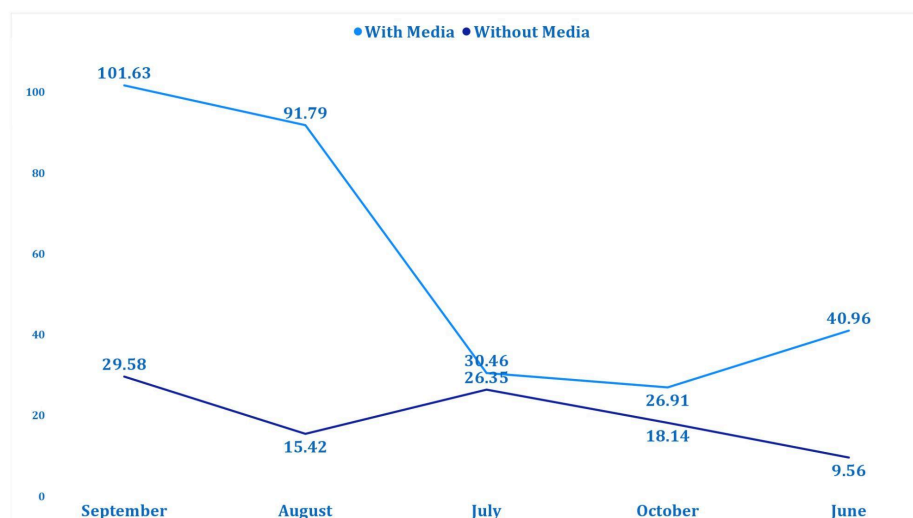

Fig. 8.3 Task 3: Average Engagement Rate Trend (With Media vs. Without Media)

For Task 3, the line chart shows the engagement trends for tweets with and without media content, filtered to data between 3-5 PM IST and 7-11 AM IST. Tweets with media had higher engagement at first but dropped over time, while without media tweets showed more steady engagement. The analysis only included tweets with even engagement rates, odd dates, fewer than 20 characters, and without the letter "c." With Media, the highest engagement was in September (101.63), and the lowest was in October (26.91). Without Media, the highest engagement was in September (29.58), and the lowest was in June (9.56).
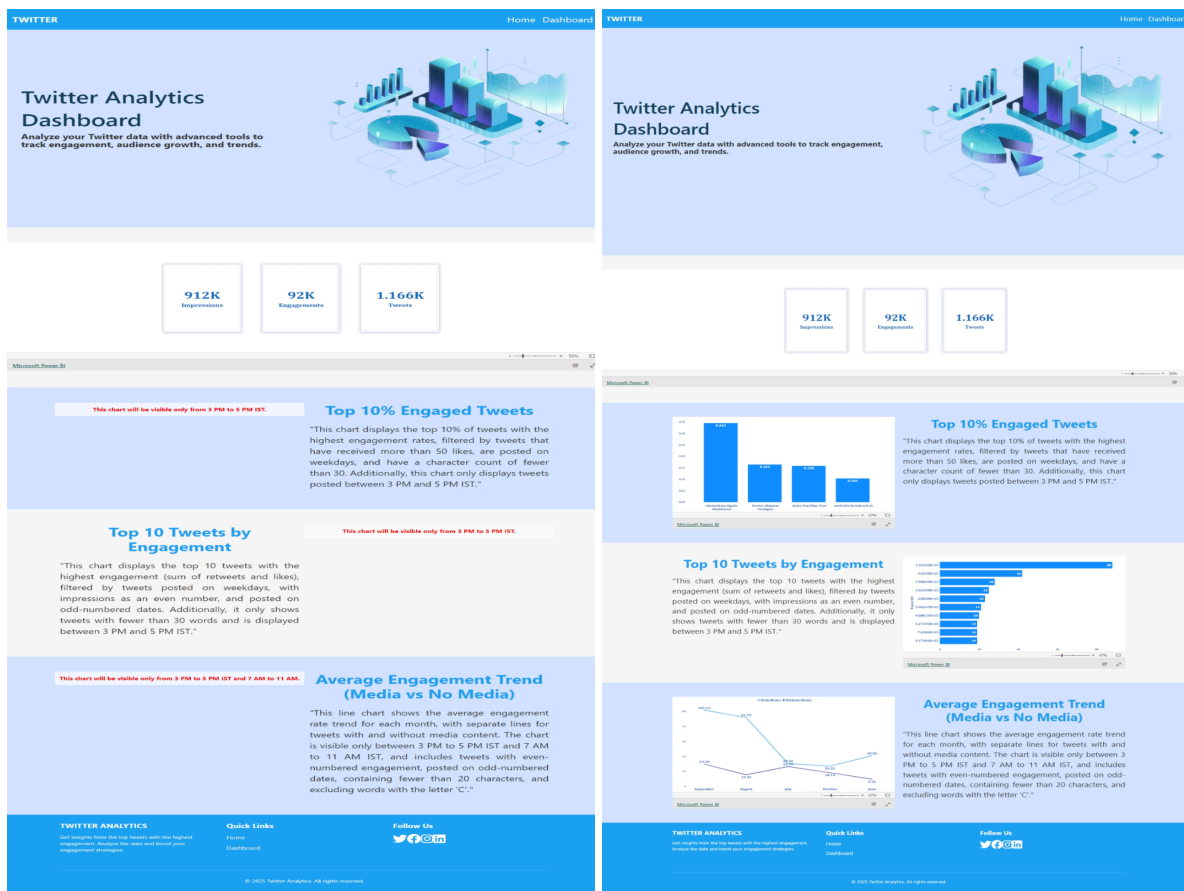

Fig. 8.4 Website Views at Different Times

The first image shows the website interface where the chart section is not visible, indicating that the charts are accessible only within the specified time frames of 3 PM to 5 PM IST or 7AM to 11AM IST. On the other side, the website is shown during the allowed time, displaying the charts as they appear when the filters and conditions, such as engagement rates, tweet times, and other criteria, are applied. This visual highlights how the website dynamically adjusts its content based on the time restrictions, ensuring that charts are displayed only when the conditions are met.
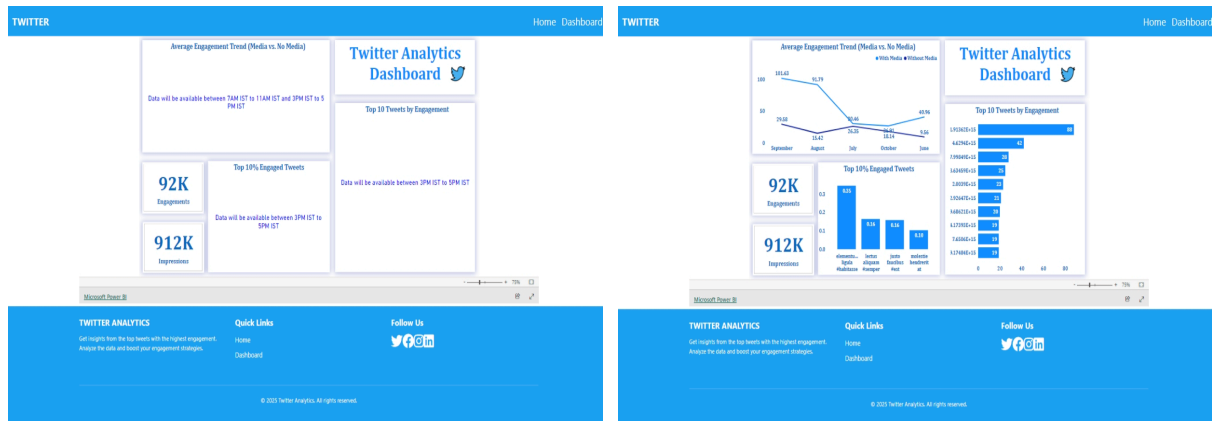
Fig. 8.5: Dashboard Views at Different Times on Website

One side of the image shows the dashboard interface with the chart section hidden, indicating that the charts are restricted and accessible only during the specified time frames of 3 PM to 5 PM IST or 7 AM to 11 AM IST. On the other side, the dashboard is displayed during the allowed time, showing the charts as they appear when the filters and conditions, such as engagement rates, tweet timings, and other criteria, are applied. This visual demonstrates how the dashboard adjusts its content dynamically, ensuring the charts are visible only within the designated time periods when the criteria are satisfied.

9. Conclusion

In conclusion, This project provides a responsive website hosted on Netlify, a thorough Power BI dashboard, and successful task execution to offer actionable insights on patterns in Twitter activity. Using sophisticated filtering criteria, such as tweets with more than 50 likes, shared on weekdays, and containing fewer than 30 characters, the duties include identifying top-performing tweets, tracking engagement trends over time, and assessing the impact of media versus non-media content. The dashboard is embedded into the responsive website using the "Publish to Web" feature, which guarantees that users can view the completed tasks and insights from any device without the need for a Power BI license. For social media marketers, companies, and individuals, the combination of job execution, comprehensive data, and convenient access is beneficial. This enables them to identify the most effective tweets, modify posting schedules to increase interaction, and enhance tactics through data analysis. With this service, users may improve their content, increase interactions, and gradually expand their social network presence.