Twitter Analytics Using Power BI

Penmetsa Navya
navyapenmetsa@gmail.com

Table of Contents

Introduction

During my internship, I was tasked with developing a comprehensive Twitter Analytics dashboard using Power BI. This project aimed to analyse and visualize Twitter data to derive actionable insights on user engagement. The dashboard was designed to meet specific business needs, such as identifying trends in engagement and understanding the impact of different types of content. The report highlights the objectives, tasks undertaken, skills developed, challenges faced.

Microsoft's Power BI is a common business intelligence tool used by business analysts and professionals. It is a data visualization and reporting platform that is used by businesses and professionals every day. While the platform is commonly used by business analysts, it is also designed to be easily accessible for those without any specialized data knowledge. Power BI stands for Power Business Intelligence and refers to a set of software tools and connectors that help in transforming data from multiple sources into actionable insights.

Background

Social media platforms, particularly Twitter, have become integral to brand communication and customer engagement strategies. Analysing Twitter data helps businesses understand user interactions, measure the effectiveness of content, and identify areas for improvement. The assignment focused on leveraging Power BI to transform raw Twitter data into meaningful visualizations, providing insights into tweet performance, audience engagement, and content effectiveness.

Learning Objectives

The primary learning objectives of this project were:

- **Proficiency in Power BI**: To develop advanced skills in Power BI, including data modelling, visualization creation, and dashboard design.
- **Understanding Social Media Metrics**: To gain a deeper understanding of key Twitter metrics such as engagement rate, retweets, likes, replies, and media interactions.
- **Data Analysis and Interpretation**: To learn how to analyse and interpret social media data to make informed decisions that can enhance social media strategies.

Activities and Tasks

Task 1: *Top 10 Tweets by Engagement*

Objective:

Identify the top 10 tweets by the sum of retweets and likes, filter out tweets posted on weekends, and show the user profile that posted each tweet.

Steps Taken:

1. Data Cleaning: Removed duplicates, handled missing values, and ensured consistency in data formats.
2. Data Transformation:
    a. Extracted the day of the week from the tweet timestamps.
    b. Applied a filter to exclude tweets posted on weekends.
    c. Created a new column to calculate total engagement (sum of retweets and likes).
3. Analysis: Ranked tweets by total engagement to identify the top 10.
4. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Bar Chart
    c. *Details*: Displayed the top 10 tweets. The bar chart included bars representing each tweet's total engagement, with tooltips showing detailed metrics (retweets and likes) when hovering over the bars.

Task 2: Average Engagement Rate Trend

Objective:

Create a line chart showing the trend of the average engagement rate over each month of the year, separating the lines for tweets with media content and those without.

Steps Taken:

1. Data Cleaning: Addressed missing values, standardized date formats, and verified data integrity.
2. Data Transformation:
    a. Calculated the engagement rate for each tweet.
    b. Extracted the month and year from tweet timestamps.
    c. Segregated tweets into two groups: with media content and without media content.
3. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Line Chart
    c. *Details*: Created a dual-line chart to show the monthly trends for tweets with and without media content. Each line represented the average engagement rate for the respective group. Custom colours and markers were used to differentiate between the two lines, and a legend was included for clarity. Annotations were added to highlight significant trends or changes.

Task 3: Comparison of Replies, Retweets, and Likes

Objective:

Develop a visualization that compares the number of replies, retweets, and likes for tweets that have received media engagements greater than the median value, including a filter for tweets posted in the last six months.

Steps Taken:

1. Data Cleaning: Cleared any data inconsistencies, handled missing values, and normalized data formats.
2. Data Transformation:
    a. Calculated the median value of media engagements.
    b. Filtered tweets to include only those with media engagements greater than the median.
    c. Applied an additional filter to include only tweets posted in the last six months.
3. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Stacked Bar Chart
    c. *Details*: Created a stacked bar chart to compare the number of replies, retweets, and likes for the selected tweets. Each bar represented a tweet, and the segments of the bar showed the distribution of replies, retweets, and likes. Custom colours were used to distinguish between the interaction types, and a filter was added to allow users to dynamically adjust the timeframe.

Task 4: Engagement Rate with App Opens

Objective:

Analyse tweets to show a comparison of the engagement rate for tweets with app opens versus tweets without app opens, including only tweets posted between 9 AM and 5 PM on weekdays.

Steps Taken:

1. Data Cleaning: Standardized data formats, handled missing values, and ensured accuracy in the timestamps.
2. Data Transformation:
    a. Extracted the hour from tweet timestamps to filter tweets posted between 9 AM and 5 PM.
    b. Filtered tweets to include only those posted on weekdays.

    c. Segregated tweets into two groups: with app opens and without app opens.
    d. Calculated the average engagement rate for each group.
3. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Clustered Column Chart
    c. *Details*: Created a clustered column chart to display the engagement rates for tweets with and without app opens. Each column represented the sum of engagement rate for each group. Custom colours were used to differentiate between the groups, and tooltips provided additional details about the engagement metrics. Filters were added to allow users to adjust the time frame and days of the week.

Task 5: Media Views and Engagements

Objective:

Create a dual-axis chart that shows the number of media views and media engagements by the day of the week for the last quarter, highlighting days with significant spikes in media interactions.

Steps Taken:

1. Data Cleaning: Verified data accuracy, handled missing values, and ensured consistency in data formats.
2. Data Transformation:
    a. Extracted the day of the week from the timestamps.
    b. Aggregated the data to calculate the average number of media views and engagements for each day of the week.
3. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Line and Clustered Column Chart
    c. *Details*: Created a dual-axis chart to display media views and engagements by the day of the week. One axis represented media views, and the other represented media engagements. Custom colours and markers were used to highlight significant spikes in media interactions.

Task 6: Proportion of Total Clicks

Objective:

Build a pie chart that represents the proportion of total clicks (URL clicks, user profile clicks, and hashtag clicks) for tweets with more than 500 impressions, including a drill-down to view the specific types of clicks for each tweet.

Steps Taken:

1. Data Cleaning: Removed duplicates, handled missing values, and ensured consistency in data formats.
2. Data Transformation:
    a. Filtered tweets to include only those with more than 500 impressions.
    b. Calculated the total clicks for each category (URL, user profile, hashtag).
3. Visualization:
    a. *Tool Used*: Power BI
    b. *Visual Type*: Pie Chart
    c. *Details*: Created a pie chart to display the proportion of each type of click. Filters were added to show tweets with more than 500 impressions. Custom colours and labels were used to enhance clarity, and tooltips provided additional details on the click metrics.

Skills and Competencies

This project allowed me to develop and enhance the following skills and competencies:

- **Data Analysis**: Improved ability to analyse complex data sets and extract relevant insights.
- **Power BI Proficiency**: Advanced skills in using Power BI for data visualization, including creating interactive dashboards.
- **Problem-Solving**: Developed problem-solving skills by addressing data-related challenges and refining visualizations to better meet the project objectives.
- **Attention to Detail**: Strengthened attention to detail, particularly in filtering data and ensuring accurate representation of trends and metrics.

Challenges and Solutions

During the project, I encountered several challenges, including:

- **Data Quality Issues**: Faced challenges with incomplete or incorrect data entries. To ensure better results, implemented rigorous data cleaning procedures and validation checks.
- **Data Filtering Issues**: Initially, filtering tweets by specific criteria was challenging. I resolved this by carefully refining the data model and using advanced filtering options in Power BI.
- **Visual Overlap**: Some visualizations were too complex and led to overlapping information. I addressed this by simplifying the visuals and using dual-axis charts where appropriate.

- **Various Calculations**: Various calculations were needed for different tweet categories which required a deep understanding of the underlying data. I overcame this by working closely with the data set and iteratively refining the calculations.

Outcome and Impact

The Twitter Analytics dashboard provided significant value by offering clear insights into user engagement. The ability to identify top-performing tweets, understand trends in engagement, and compare different types of content will allow businesses to make data-driven decisions. The dashboard also served as a tool for optimizing future content strategies, ensuring higher engagement rates and better alignment with audience preferences. This also contributes to the business's understanding of its Twitter audience and helped in making data-driven decisions to enhance engagement and reach.

Conclusion

This internship project was a valuable learning experience that enhanced my skills in data analysis, Power BI, and social media analytics. The Twitter Analytics dashboard not only met the project objectives but also provided the organization with a powerful tool to drive social media strategy. The challenges faced during the project helped me grow as a problem-solver and deepened my understanding of data visualization best practices.