Internship Report (Sakshi Thakare)

Data Analytics Internship - Twitter Analytics Dashboard

Introduction

This report covers the work I completed during my Data Analytics internship with Null Class, focusing on developing a Twitter Analytics Dashboard using Power BI, with a focus on analysing tweet data based on various engagement metrics. The internship aimed to enhance my skills in data visualization, analysis, and dashboard creation through hands-on tasks involving real-world social media datasets.

Background

The internship revolved around analysing Twitter data and presenting actionable insights through visual representations. This involved creating various Power BI dashboards for understanding tweet performance based on metrics such as clicks, retweets, likes, media engagements, and views. The tasks provided an opportunity to apply analytical techniques while using real-world data.

Learning Objectives

The main objectives of the internship were to:

1. Develop skills in data visualization and business intelligence using Power BI.
2. Gain hands-on experience in analysing and interpreting Twitter datasets.
3. Learn to create interactive and insightful dashboards that highlight key metrics.
4. Understand how to handle real-world data, including filtering, formatting, and validation.
5. To integrate data visualization into a responsive web project hosted on Vercel.
6. To learn the steps of making a dashboard mobile-responsive.

Activities and Tasks

During the internship, I completed the following tasks:

- **Task 1: Pie Chart for Clicks**:

1. Created a pie chart showing the proportion of total clicks (URL clicks, profile clicks, and hashtag clicks) for tweets with more than 500 impressions, with drill-down functionality.
2. The types of clicks considered include **URL clicks**, **user profile clicks**, and **hashtag clicks**.
3. A **drill-down feature** was implemented to provide more granular insights, allowing the user to click on a segment of the pie chart and drill down to view the specific types of clicks associated with each tweet.

Insights:

- **URL Clicks** accounted for the **largest proportion** of the total engagement, suggesting that users were more inclined to follow external links provided in tweets with high impressions.
- **Profile Clicks** made up a moderate portion, showing that users were interested in learning more about the tweet author, especially for tweets with high engagement.
- **Hashtag Clicks** had a relatively smaller proportion of engagement, indicating that hashtags, while useful for reach, did not generate as many direct interactions as URLs or profile visits.
- **Drill-Down Analysis**: Tweets with more than 500 impressions were primarily driven by **URL clicks**, while tweets with slightly lower impressions tended to have a more balanced distribution between profile and hashtag clicks.

Conclusion:

This visualization highlighted that for tweets with higher impressions, external content (URLs) played a pivotal role in driving engagement. The drill-down feature provided valuable insights into how different click types contribute to overall engagement for specific tweets.

- **Task 2: Top 10 Tweets Chart**:

1. Built a chart to identify the top 10 tweets based on the sum of retweets and likes, applying filters for time of day, even tweet impressions, odd tweet dates, and tweet word count under 30.
2. The dataset was filtered to exclude tweets posted on weekends, with additional constraints:
    - The tweet must have been posted between **3 PM to 6 PM**.
    - The tweet must have an **even impression count**, an **odd tweet date**, and a **word count below 30**.

Insights:

- **Engagement Peaks during Weekdays**: The analysis showed that tweets posted between **3 PM and 6 PM** on weekdays had the highest levels of engagement in terms of retweets and likes. This confirms that this time window is optimal for gaining user attention.
- **Tweets with Lower Word Count**: Tweets with fewer words (below 30) consistently performed better in terms of engagement, confirming that **concise content** often captures attention more effectively.
- **Odd Date Influence**: Interestingly, tweets posted on **odd-numbered dates** showed better engagement. While this is likely a coincidental trend, it suggests that tweet timing and scheduling may subtly influence engagement.
- **High Retweets & Likes Correlation**: Tweets with the highest impressions also had a strong correlation between the number of retweets and likes. This indicates that higher impressions generally drive engagement in both metrics.

Conclusion:

This task provided insights into the **optimal timing** and content length for generating high engagement on Twitter. The filtered data showed that weekday afternoons, short tweets, and even impression numbers were key factors contributing to better performance.

- **Task 3: Dual-Axis Chart**:

1. Developed a dual-axis chart displaying media views and media engagements by day of the week for the last quarter, highlighting days with significant spikes in media interactions. The same filtering criteria as Task 2 were applied, meaning that only tweets posted between **3 PM to 6 PM** with even impressions, odd dates, and word counts below 30 were considered.
2. The focus was on identifying days with significant spikes in media interaction.

Insights:

- **Significant Spikes on Tuesdays and Thursdays**: The analysis revealed that **Tuesdays** and **Thursdays** saw significant spikes in both media views and media engagements. This aligns with broader social media trends, where midweek days often result in higher user engagement.
- **Higher Engagement on Shorter Tweets**: Tweets with lower word counts continued to outperform longer tweets in terms of both media views and engagements, reinforcing the value of concise communication.
- **Media Views vs. Engagements Correlation**: There was a clear positive correlation between the number of **media views** and **engagements**, indicating that tweets with higher visual media (such as images or videos) tend to engage users more effectively. The more media views a tweet received, the higher the overall engagement.
- **Quarterly Trends**: The analysis of the last quarter showed a consistent pattern where engagement dropped on weekends and increased during the middle of the week, particularly on Thursdays. These findings suggest that organizations could maximize reach by focusing on high-quality media content during midweek posts.

Conclusion:

This dual-axis chart provided deeper insights into the importance of **visual media** and how **timing** affects media interaction on Twitter. The clear spikes in engagement on certain days emphasize the strategic timing of tweets for maximizing media reach.

Skills and Competencies

Through this internship, I developed and refined the following skills:

1. **Power BI**: Learned to create complex dashboards with multiple visualizations and filters.
2. **Data Analysis**: Gained experience in handling and analyzing large datasets.
3. **Data Validation**: Ensured data accuracy by applying filters and validating the filtered dataset.
4. **Problem Solving**: Addressed challenges such as data inconsistencies and the application of custom filters.

5. **Web Development**: Learned to deploy projects on Vercel and ensure mobile responsiveness for real-time web applications.

Challenges and Solutions

1. **Filtering Complex Data**: One of the key challenges was applying multiple filters (time, impressions, tweet dates, word count) for specific tasks. By using DAX formulas and troubleshooting errors, I was able to solve these challenges and ensure the correct data was displayed.
2. **Data Validation**: Ensuring the dataset met all criteria for accurate analysis required careful attention. I addressed this by cross-checking filters and conditions before building each visualization.
3. **Responsiveness**: Ensuring that the dashboard was fully responsive on all devices required careful CSS adjustments and testing on different screen sizes.
4. **Image Hosting Issues**: There were initial issues with image visibility on mobile, which were resolved by ensuring the correct image paths and deploying the images alongside the HTML files.

Outcomes and Impact

The tasks I completed helped create insightful dashboards that could be used for analyzing Twitter engagement metrics effectively. These visualizations offer valuable insights into tweet performance, which can aid in data-driven decision-making for social media strategies. Personally, the internship helped me strengthen my data analysis and visualization skills and provided a deeper understanding of working with social media data.

Conclusion

The internship at Null Class provided a valuable learning experience in data analytics and visualization using Power BI. I completed key tasks related to Twitter data analysis, overcoming challenges related to data filtering and validation. The skills gained will be instrumental in my future career in data analytics and business intelligence.