Twitter Data Analysis

**NullClass Internship Report**

This document certifies that I have successfully completed my **One-month internship**, including a structured training session. Throughout the internship, I worked on the project titled “**Real-Time Twitter Analysis**” using **Power BI** fulfilling both training and project tasks within the designated timeframe.

By

Saiteja Puppala

1. **Acknowledgements**

I would like to extend my heartfelt gratitude to my trainer, Miss Vaishnavi Kundeshwara, for her invaluable Training. Additionally, I express my sincere thanks to NullClass for entrusting me with this opportunity and providing me with this internship.

1. **Introduction**

**NullClass** is an organization that focuses on providing real-world experience through hands-on projects, mentorship, and internships. They aim to equip the next generation of tech talent with practical skills and industry-ready capabilities1.

* 1. **Power BI Training at NullClass**

NullClass offers comprehensive training in **Power BI**, which includes:

* **Hands-on Projects**: You'll work on real-world projects to apply your Power BI skills.
* **Mentorship**: Experienced mentors guide you through the learning process.
* **Certifications**: Opportunities to earn certifications that validate your skills.
  1. **Objectives of the internship**

The primary objective of this project was to leverage Power BI for comprehensive analysis of Twitter data. The goal was to extract actionable insights by visualizing various metrics such as tweet impressions, engagements, and interactions. This analysis aimed to identify patterns, trends, and anomalies in user behavior, thereby enabling data-driven decision-making for enhanced social media strategy.

1. **Roles and Responsibilities**

As a Power BI Developer, my daily tasks and responsibilities during the internship included:

1. **Data Cleaning and Preparation**:

* Imported and cleaned raw Twitter data to ensure accuracy and reliability.
* Utilized DAX (Data Analysis Expressions) to create calculated columns and measures for deeper analysis.
* Applied transformations to data using Power Query Editor for optimized visualization.

1. **Data Analysis and Feature Engineering**:

* Analyzed data to identify key metrics and performance indicators.
* Developed feature-rich datasets using DAX to enhance the analytical capabilities of the reports.

1. **Report Development**:

* Created interactive and dynamic dashboards using Power BI.
* Designed visualizations that provide clear and insightful data representations.
* Integrated various data sources to build comprehensive reports.

1. **Collaboration and Feedback**:

* Worked closely with team members to understand project requirements and deliverables.
* Attended regular meetings to discuss progress, challenges, and next steps.
* Incorporated feedback from mentors and stakeholders to improve report quality.

1. **Project Details**
   1. **Tittle:**

**Twitter Data Analysis Using Power BI**

* 1. **Methodologies:**

Data Collection

Data Cleaning and Preparation

Feature Engineering

Data Analysis

Data Visualization

Reporting and Insights

* 1. **Results**

Successfully cleaned and prepared a large dataset for analysis.

Created interactive dashboards that provided valuable insights into Twitter engagement metrics.

Identified patterns and trends in user behavior, helping to inform social media strategy.

Presented clear and actionable insights through well-designed visualizations.

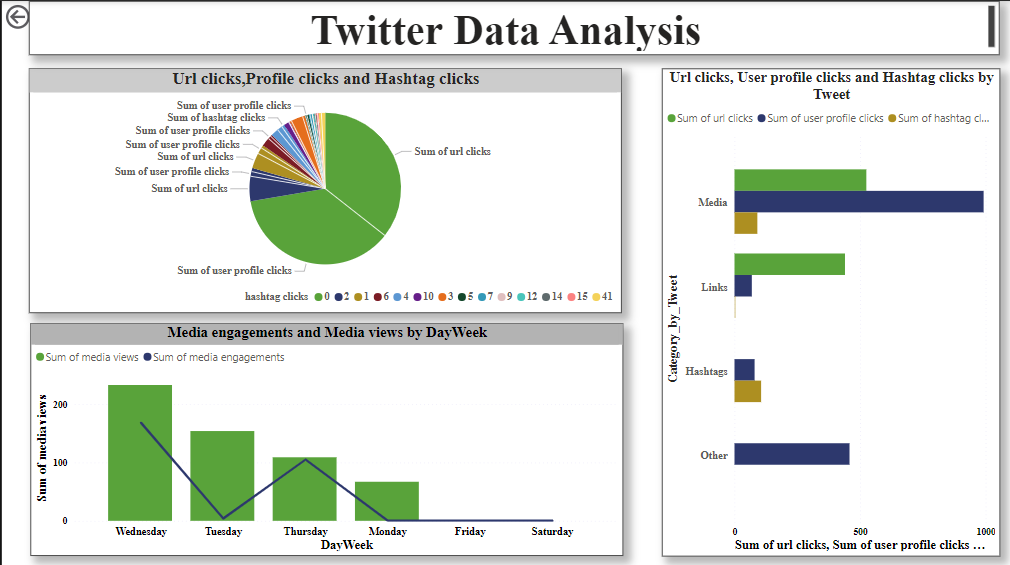
* + 1. **Challenges and Learning**
  1. **Data Cleaning and Preparation**:
* Ensuring data quality and consistency.
* Handling missing or incomplete data.
* Transforming data from various sources into a usable format.
  1. **Complex DAX Formulas**:
* Writing and optimizing complex DAX expressions.
* Debugging and troubleshooting errors in DAX formulas.
  1. **Data Analysis Skills**:
* Gaining proficiency in data analysis and visualization techniques.
* Developing a deeper understanding of data structures and relationships.
  1. **Advanced DAX Knowledge**:
* Mastering DAX to create powerful and flexible calculations.
* Learning best practices for writing efficient DAX expressions.
  1. **Performance Tuning**:
* Acquiring skills to optimize performance and manage large datasets.
* Implementing best practices for query optimization.

1. **Appendices**

Following the completion of my training, I have developed these visualizations. They provide valuable insights and are instrumental in analyzing the data effectively.

**Appendix-1**

****

**Appendix-2**

1. **Conclusion**
   1. **Insights from Appendix-1**

**Tweet Activity by Weekday**:

The **tweet count** is highest on Thursday and gradually decreases towards Sunday, indicating higher engagement or content posting earlier in the week.

**Media Engagements vs. Media Views**:

Both media engagements and views peak on Thursday, then steadily decline towards Sunday. This suggests that Thursday is the most active day for media interactions, with higher user engagement and content consumption.

**Top Performing Tweets**:

The **horizontal bar chart** highlights the tweets with the highest URL clicks, with the tweet starting with "in tempus sit ..." leading in performance. This information can help identify the content types that drive the most clicks.

**Click Distribution**:

The **pie chart** shows that user profile clicks constitute the majority (39.86%) of the total clicks, followed by hashtag clicks and URL clicks. This indicates that users are more interested in viewing profiles than clicking on URLs or hashtags.

**Retweets and Likes**:

The **gauges** indicate substantial user engagement, with 402 retweets and 8,187 likes, showcasing the popularity and reach of the tweets.

**Impressions and Engagement Rates**:

The **impressions** metric stands at 773.28, while the **engagement rate** is 0.07. This suggests that while a significant number of users see the tweets, the proportion engaging with the content is moderate.

**Media Views**:

The **media views** metric shows a substantial count of 89K, indicating high visual content consumption by users.

**Impressions by Week**:

The **line graph** shows that impressions peak on Friday and then decline towards Sunday, highlighting the importance of posting content towards the end of the week for higher visibility.

* 1. **Insights from Appendix-2**

### URL Clicks, Profile Clicks, and Hashtag Clicks (Pie Chart):

* **URL Clicks**: There are 41 URL clicks, representing the majority of the interactions.
* **User Profile Clicks**: There are three categories for user profile clicks with 15, 10, and 4 clicks, respectively.
* **Hashtag Clicks**: Hashtag clicks have 12 interactions.
* **Other Categories**: Additional categories include 6 URL clicks and 2 URL clicks.

This pie chart illustrates the distribution of various types of clicks, emphasizing that URL clicks constitute a significant portion of user interactions.

### URL Clicks, User Profile Clicks, and Hashtag Clicks by Tweet (Bar Chart):

* **Categories**: This chart breaks down the clicks by tweet category, showing the total clicks for URL, user profile, and Hashtag clicks.
* **Media**: The Media category leads with the highest number of user profile clicks.
* **Other, Links, and Hashtags** : Follow in descending order of user profile clicks, indicating different levels of user engagement with various tweet types.

This bar chart provides a clear comparison of engagement types across different tweet categories.

### Media Engagements and Media Views by Day of the Week (Bar and Line Chart):

* **Media Views (Green Bars)**: Peak on Wednesday, followed by Tuesday, Thursday, Monday, and Saturday.
* **Media Engagements (Blue Line)**: Also highest on Wednesday, indicating that this is the day with the most significant media interaction.

The dual representation (bar and line) effectively showcases the correlation between media views and engagements by the day of the week, highlighting Wednesday as the day of maximum activity.