

# Internship Project Report: Twitter Dashboard Analysis

---

## 1. Introduction

This report provides a comprehensive overview of the data analysis and visualization project completed during the Nullclass Data Analyst Internship. The project's core focus was to analyze a dataset of Twitter tweets to develop a real-time analytics dashboard using Power BI. The following sections detail the project's background, objectives, tasks, and outcomes, demonstrating the skills and competencies acquired throughout the process.

## 2. Background

The project was based on a provided dataset of Twitter data, including metrics such as impressions, engagements, retweets, and various click types. The primary business objective was to build a dynamic dashboard that could present these metrics in a meaningful way, helping stakeholders understand tweet performance, identify trends, and analyze user interaction. The project was designed to simulate a real-world data analytics scenario, requiring data cleaning, transformation, and advanced visualization techniques.

## 3. Learning Objectives

The key learning objectives for this project were to:

- Master Data Preparation: Learn to clean, transform, and prepare raw data from a .xlsx file for analysis using Python and Power BI's Power Query Editor.
- Develop Advanced Visualization Skills: Create a variety of chart types, including clustered bar charts, line charts, and dual-axis charts, to effectively present complex data.
- Apply DAX for Dynamic Analysis: Utilize Data Analysis Expressions (DAX) to create custom measures for conditional visibility, median calculations, and other advanced filtering logic.
- Enhance Problem-Solving Abilities: Identify and resolve common data-related issues, such as file encoding errors, missing columns, and data type mismatches.
- Understand End-to-End Project Workflow: Experience the full lifecycle of a data project, from initial data cleaning to final dashboard deployment.

## 4. Activities and Tasks

The project involved the successful completion of ten distinct tasks, each focusing on a different aspect of data analysis. Key activities included:

1. Creating a dual-axis chart to analyze average engagement rate and total impressions over a specific date range.
2. Developing a chart to display the top 10% of tweets by engagement rate, with filters for

likes and word count.

3. Building a pie chart with a drill-down feature to show the proportion of total clicks (URL, user profile, hashtag).
4. Plotting a scatter chart to analyze the relationship between media views and engagements, including conditional highlighting.
5. Designing a clustered bar chart to break down clicks by tweet category (e.g., tweets with media, links, etc.).
6. Identifying the top 10 tweets based on a custom metric of retweets and likes.
7. Creating a dual-axis chart to show media views and engagements by day of the week, with advanced filtering rules.
8. Building a line chart to trend average engagement rates over time, separating tweets with and without media.
9. Comparing engagement rates for tweets with and without app opens.
10. Developing a visualization to compare replies, retweets, and likes for high-performing tweets (above the median media engagement).

## 5. Skills and Competencies

This project honed both technical and soft skills essential for a data analyst:

- Data Manipulation: Proficiency in Python with pandas for data cleaning, type conversion, and feature engineering.
- Data Visualization: Expertise in Power BI for creating a wide range of interactive and aesthetically pleasing charts.
- DAX (Data Analysis Expressions): Competency in writing complex DAX measures for dynamic filtering and custom metrics.
- Problem-Solving: Ability to debug and resolve unexpected errors, such as file-reading issues (KeyError, UnicodeDecodeError) and software dependency problems (openpyxl).
- Attention to Detail: Meticulous approach to applying precise filter conditions and ensuring data integrity.

## 6. Feedback and Evidence

Feedback from mentors highlighted the structured approach to solving complex filtering problems and the professional presentation of visuals. Evidence of work includes the GitHub repository link containing the Power BI (.pbix) file, Python scripts for preprocessing, and exported PDF reports.

## 7. Challenges and Solutions

Several challenges were encountered during the project:

- File Reading Errors: The Python script initially failed to read the source data due to file format and encoding issues. This was resolved by using 'pd.read\_excel' instead of 'read\_csv', specifying the correct encoding, and a robust parsing engine.

- Missing Python Dependencies: The 'openpyxl' library was required to read the .xlsx file but was not installed. Installing it via 'pip install openpyxl' resolved the issue.
- DAX Measure vs. Calculated Column: A metric for Total Likes & Retweets was initially implemented as a calculated column, resulting in static values. Changing it to a DAX measure made the calculation dynamic and context-aware.

## **8. Outcomes and Impact**

The project successfully delivered a comprehensive and fully functional Power BI dashboard. Visualizations provided actionable insights into tweet performance, helping identify the most engaging tweet types and the relationship between media and user interaction. The advanced filtering rules showcased an ability to create tailored, real-time analytics solutions.

## **9. Conclusion**

This internship project was a valuable learning experience, allowing the application of end-to-end data analytics processes. Completing all ten tasks demonstrated the ability to prepare data, create dynamic dashboards, and solve technical challenges independently. These skills will be directly applicable in future data-driven roles.