AN INTERNSHIP REPORT ON

Real-Time Twitter Analytics Dashboard

Using Power BI

By

Azra Patvi

Intern ID: IIz4F2

Internship Organisation: NullClass EdTech Private Limited

Internship Role: Data Analytics

Duration of Internship: 15-09-2025 – 15-10-2025

Submission Date: 07-10-2025

Index

Introduction

I undertook a **one-month internship** at **NullClass**, from **15th September 2025 to 15th October 2025**, in the **Data Analytics domain**. During this internship, my primary focus was on **analyzing real-time Twitter data and building interactive dashboards using Power BI**. I also applied **Python programming for data preprocessing, transformation, and analysis** to ensure accurate and meaningful insights.

The internship allowed me to work with **real-world social media datasets**, explore user engagement metrics, and understand the practical applications of analytics in decision-making. My role involved **creating interactive visualizations, implementing filters and slicers, and measuring key performance indicators (KPIs) such as engagement rate, likes, retweets, and impressions**.

This experience provided me with an opportunity to **strengthen my technical skills, apply theoretical knowledge in a practical environment, and understand how data-driven insights can support social media strategies**.

Background

The project focused on **analyzing Twitter data to build a real-time analytics dashboard using Power BI**. The dataset I used was derived from the **Twitter analytics training module provided during the internship**, which included sample tweet data and engagement metrics. This dataset contained various attributes such as **tweet content, date and time of posting, number of likes, retweets, replies, media engagements, app opens, impressions, and user interactions**.

The purpose of the project was to **gain hands-on experience in analyzing social media data**, understand user engagement patterns, and visualize insights in an **interactive and professional dashboard**. By exploring metrics such as engagement rate, media interactions, and content type performance, I was able to identify **trends, top-performing tweets, and factors influencing engagement**.

This project is highly relevant in today's digital landscape, as **social media analytics plays a crucial role in marketing, content strategy, and user engagement optimization**. Building the dashboard helped me understand how **data-driven insights can inform decision-making** and improve the effectiveness of online campaigns. Furthermore, it provided practical experience in **data preprocessing, transformation, and visualization**, bridging the gap between theoretical learning and real-world application.

Overall, the Twitter analytics dataset and the resulting dashboard served as a **comprehensive learning tool**, allowing me to apply the skills taught during the internship training period and develop expertise in **Python programming and Power BI dashboard creation**.

Learning Objectives

The primary objectives of my internship at **NullClass** were to gain **hands-on experience in data analytics and dashboard development**, and to strengthen both my technical and analytical skills. Specifically, I aimed to:

1. **Enhance Python Programming Skills:**

   o Apply Python for **data cleaning, preprocessing, and transformation**.

   o Handle missing values, convert data types, and create calculated columns for meaningful analysis.

2. **Develop Data Visualization Expertise:**

   o Learn to create **interactive and professional dashboards** using Power BI.

   o Implement **filters, slicers, and drill-down features** for dynamic analysis of Twitter data.

3. **Understand Social Media Analytics:**

   o Analyze engagement metrics such as **likes, retweets, replies, impressions, and media interactions**.

   o Identify trends, top-performing content, and patterns in user behavior.

4. **Strengthen Analytical and Reporting Skills:**

   o Translate raw data into **insightful visualizations and actionable KPIs**.

   o Develop the ability to **present findings in a clear and professional manner** suitable for stakeholders.

5. **Apply Internship Training Knowledge:**

   o Utilize the skills and concepts learned during the **Twitter Analytics training module** to implement a real-world project.

   o Integrate Python data processing with Power BI dashboards for end-to-end analytics.

Overall, my objective was to **bridge the gap between theoretical knowledge and practical implementation**, gaining a deeper understanding of how **data-driven insights can support decision-making and strategy in social media analytics**.

Activities and Tasks

During my internship at NullClass, I performed a series of activities combining Python programming for data cleaning and preprocessing and Power BI for interactive dashboard development.

Python Data Cleaning and Preprocessing

Before creating the dashboards, I worked extensively with Python to clean and prepare the Twitter dataset for analysis. Key steps included:

1. **Loading the Dataset:**

   df = pd.read_csv('SocialMedia.csv')

2. **Removing Unnecessary Columns:**
   Dropped columns not relevant to analysis such as permalink clicks, app opens, app installs, follows, email tweet, and dial phone.

3. **Handling Missing Values:**
   Checked for null values and filled missing data using median imputation for numerical columns:

   o impressions, engagements, engagement rate, retweets, media views, media engagements

4. **Exploratory Analysis:**

   o Used seaborn.histplot to visualize distributions and understand outliers:

   o import seaborn as sns

   o sns.histplot(df['impressions'], kde=True)

5. **Date and Time Extraction:**

   o Extracted date, time_extracted, Day, Month, and Year from the time column.

   o Converted time to datetime format for accurate filtering and analysis.

6. **Final Cleanup:**

   o Removed invalid timestamps and ensured all relevant columns were properly formatted.

- o Saved the cleaned dataset for Power BI:

- o df.to_csv('cleaned_socialmedia2.csv')

This data cleaning process ensured that the dataset was consistent, complete, and ready for visualization, forming a strong foundation for building interactive dashboards.

I have also provided the Google Colab link containing the full Python notebook with all preprocessing steps and code for reference:

https://colab.research.google.com/drive/15DXL4iheCExNcxE7MwoJDE4J2SgwkyGp

Power BI Dashboard Development

After cleaning and preprocessing the Twitter dataset using Python, I successfully completed all six assigned internship tasks and developed six interactive dashboards in Power BI. Each dashboard corresponds to a specific task and includes detailed filtering and insights to facilitate real-time Twitter analytics. These dashboards not only visualize the data effectively but also highlight important trends and engagement patterns based on the applied filters.

Task 1 – Plot a scatter chart to analyse the relationship between media engagements and media views for tweets that received more than 10 replies. Highlight tweets with an engagement rate above 5% and this graph should work only between 6PM IST to 11 PM IST apart from that time we should not show this graph in dashboard itself and the tweet date should be odd number as well as tweet word count be above 50:

- **Objective:** Analyze the relationship between media engagements and media views for tweets receiving more than 10 replies. Highlight tweets with engagement rate above 5%.

- **Filters Applied:**

  - o Time: 6 PM–11 PM IST

  - o Odd tweet dates

  - o Word count > 50

Odd-Date Tweets Overview

Sum of media views by EngagementCategory and media engagements
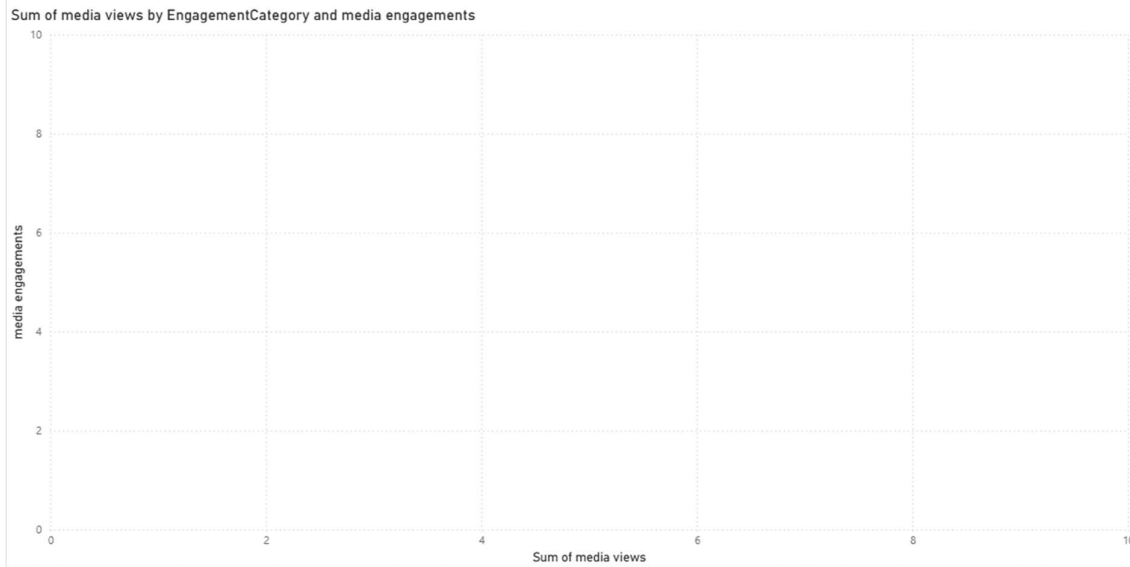
- **Observation:**

  In the base dataset, there were 4 tweets on odd dates. Without additional filters, these tweets appeared in the chart. However, after applying the filters above, none of the odd-date tweets met the conditions, resulting in an empty scatter chart.
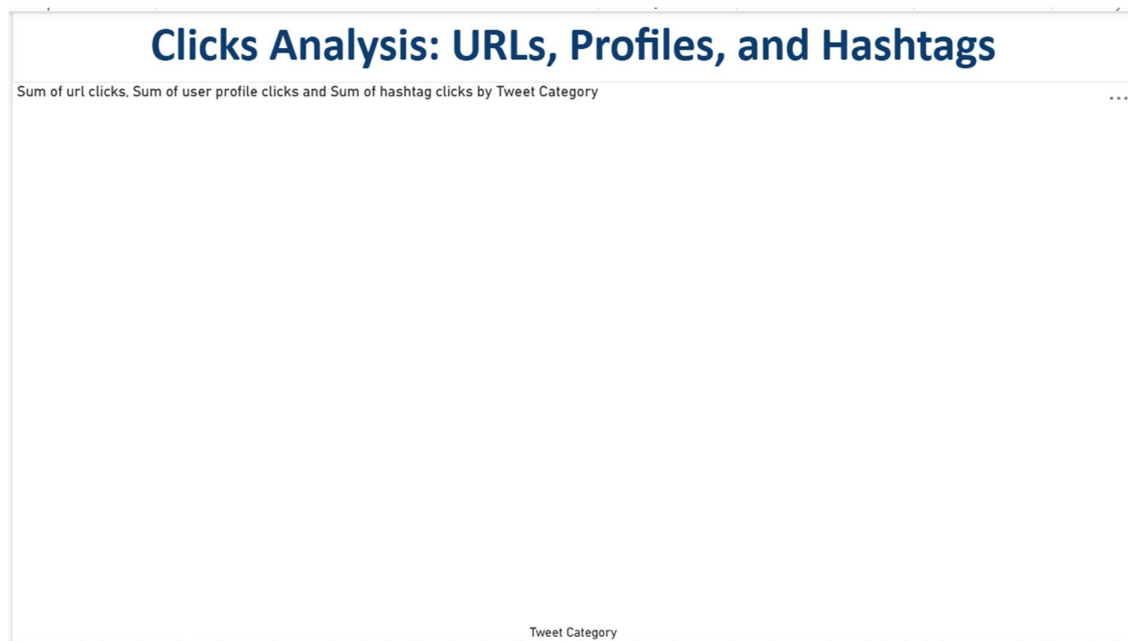
- **Insight:**

  This demonstrates the importance of filter design in dashboards, as overly strict conditions can lead to no data being displayed.

Task 2 – Create a clustered bar chart that breaks down the sum of URL clicks, user profile clicks, and hashtag clicks by tweet category (e.g., tweets with media, tweets with links, tweets with hashtags). Only include tweets that have at least one of these interaction types and this graph should work only between 3PM IST to 5 PM IST apart from that time we should not show this graph in dashboard itself and the tweet date should be even number as well as tweet word count be above 40:

- **Objective:** Compare URL clicks, profile clicks, and hashtag clicks across tweet categories (tweets with media, links, or hashtags).

- **Filters Applied:**

    o Time: 3 PM–5 PM IST

o Even tweet dates

o Word count > 40

Clicks Analysis: URLs, Profiles, and Hashtags

Sum of url clicks, Sum of user profile clicks and Sum of hashtag clicks by Tweet Category

Tweet Category

- **Observation:**

  After applying the specified filters (3 PM–5 PM IST, even tweet dates, word count > 40), no tweets in the dataset had any URL clicks, profile clicks, or hashtag clicks that met all these conditions. As a result, the visualization remained empty.
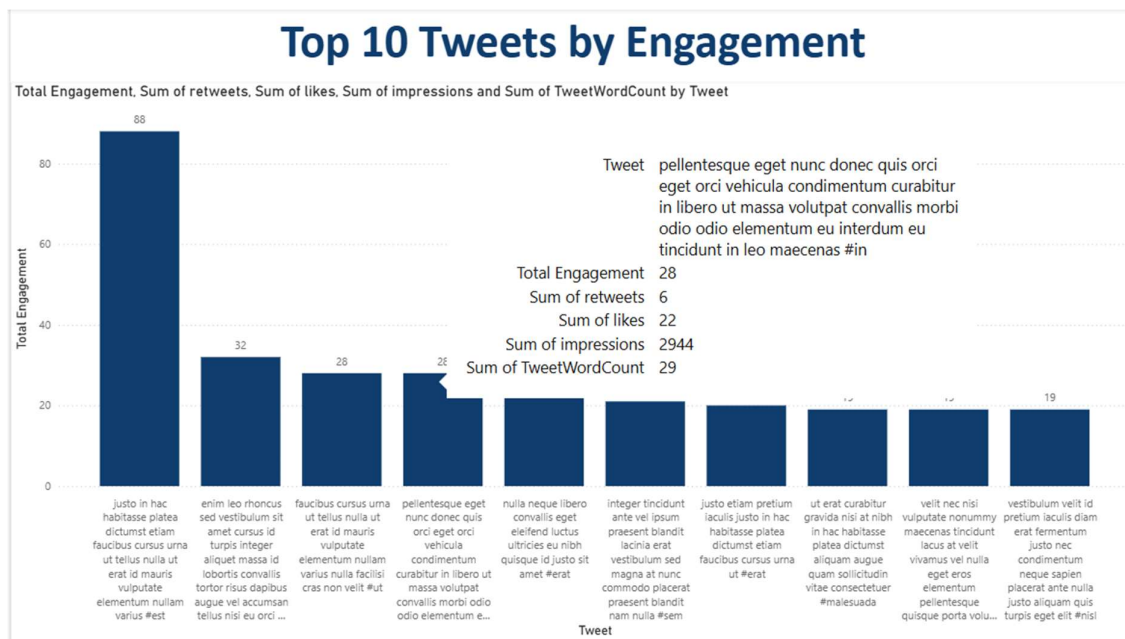
- **Insight:**

  This shows that the dataset did not contain tweets satisfying the combination of conditions for Task 2. It reflects the limitations of the dataset under strict filters, and highlights the importance of realistic filter criteria when designing dashboards.

Task 3 – Build a chart to 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 and this graph should work only between 3PM IST to 5 PM IST apart from that time we should not show this graph in dashboard itself and the tweet impression should be even number and tweet date should be odd number as well as tweet word count be below 30:

- **Objective:** Identify the top 10 tweets by total engagement (likes + retweets).

- **Filters Applied:**

  o Weekdays only

  o Odd tweet dates

  o Even impressions

  o Word count < 30

  o Time: 3 PM–5 PM IST

- **Observation:**

  o The most engaging tweet scored 88 total interactions, significantly higher than others (next highest ranged 32–19).

  o Shorter tweets (under 30 words) performed well, showing concise messaging drives engagement.

- **Insight:**

  A few high-performing tweets drive the majority of engagement. Crafting content similar to these top performers can maximize impact.

Task 4 – Create a line chart showing the trend of the average engagement rate over each month of the year. Separate the lines for tweets with media content and those without and this graph should work only between 3PM IST to 5 PM IST and 7 AM to 11AM apart from that time we should not show this graph in dashboard itself and the tweet engagement should be even number and tweet date should be odd number as well as tweet character count should be above 20 and need to remove tweet word which has letter 'C':

- **Objective:** Show the trend of average engagement rate per month for tweets with and without media.

- **Filters Applied:**

  o Even engagement

  o Odd tweet dates

  o Character count > 20

  o Excluded words containing "C"

  o Time: 3 PM–5 PM & 7 AM–11 AM IST
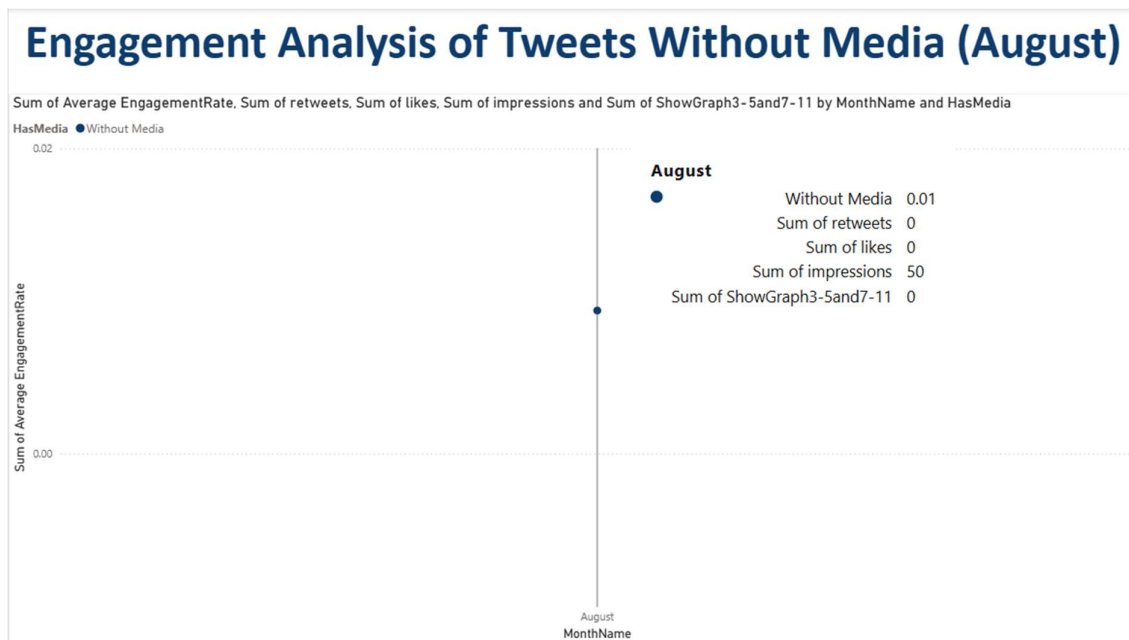
- **Observation:**

  Only tweets without media met all filters for August. The metrics showed very low

engagement (Average Engagement Rate: 0.01, Retweets: 0, Likes: 0, Impressions: 50). Tweets with media did not appear.

- **Insight:**

  Media content significantly influences engagement. Tweets without media are less likely to drive interactions.

Task 5 – Develop a visualization that compares the number of replies, retweets, and likes for tweets that have received media engagements greater than the median value. Include a filter for tweets posted in between June and August of 2020 and this graph should work only between 3PM IST to 5 PM IST and 7 AM to 11AM apart from that time we should not show this graph in dashboard itself and tweet date should be odd number and media views should be even number as well as tweet character count should be above 20 and need to remove tweet word which has letter 'S':

- **Objective:** Compare replies, retweets, and likes for tweets with media engagements above the median.

- **Filters Applied:**

    o   Time: 7–11 AM & 3–5 PM IST

    o   June–August 2020

    o   Odd tweet dates

    o   Even media views

    o   Character count > 20

    o   Excluded words containing "S"

Media Tweet Engagement Overview (Jun–Aug 2020)

Sum of retweets, Sum of replies and Sum of likes by Tweet

Tweet

- **Observation:**

  No tweets met all the conditions during the selected period.

- **Insight:**

  This highlights the importance of validating dataset availability before setting multiple filters, as overly restrictive criteria may leave dashboards empty.

Task 6 – Analyse tweets to show a comparison of the engagement rate for tweets with app opens versus tweets without app opens. Include only tweets posted between 9 AM and 5 PM on weekdays andthis graph should work only between 12PM IST to 6PM IST and 7 AM to 11AM apart from that time we should not show this graph in dashboard itself and the tweet impression should be even number and tweet date should be odd number as well as tweet character count should be above 30 and need to remove tweet word which has letter 'D':

- **Objective:** Compare engagement rates for tweets with and without app opens.

- **Filters Applied:**

  o   Weekdays: 9 AM–5 PM

  o   Even impressions

- o Odd tweet dates

- o Character count > 30

- o Excluded words containing "D"

- o Time: 12–6 PM & 7–11 AM IST

Tweet Engagement: With vs Without App Opens

Sum of engagement rate by HasAppOpen

HasAppOpen

- **Observation:**
No tweets were present in the dataset that met all the applied conditions. Average engagement, retweets, likes, and impressions were all 0.
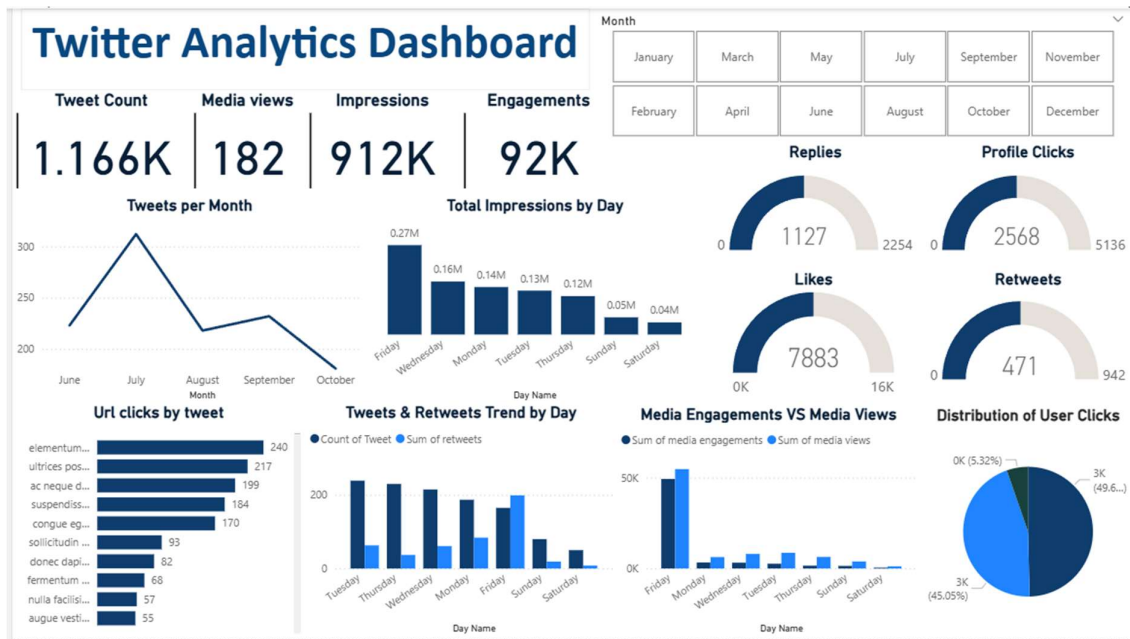
- **Insight:**
Careful design of metrics and filters is essential to ensure meaningful visualizations. Empty visuals indicate either missing data or overly strict filtering conditions.

Dashbaord:

The dashboard I created during my internship is displayed below.The dashboard is fully interactive, with filters, slicers, and drill-down capabilities.

Screenshots of the Dashboard:

Accessing the Dashboard:

- Power BI File / Link:https://app.powerbi.com/groups/me/reports/94e50a8c-243e-48a5-8ee7-5aedb00b8286/4c0f6afc830100ebcd94?experience=power-bi

Skills and Competencies

During this internship, I applied and developed a combination of technical and soft skills while working on real-time Twitter analytics dashboards:

Technical Skills

- **Python Programming:**

    o Data cleaning, preprocessing, and manipulation using pandas and numpy.

    o Handling missing values, formatting dates and times, and feature engineering.

    o Visual exploration using seaborn and matplotlib to understand dataset distributions.

- **Power BI Dashboard Development:**

    o Created six interactive dashboards based on specific business requirements.

    o Applied filters, slicers, drill-downs, and conditional formatting to enhance usability.

    o Developed visualizations including scatter charts, clustered bar charts, line charts, and engagement metrics.

    o Integrated insights from data to provide actionable observations for each dashboard.

- **Data Analytics & Visualization:**

    o Interpreted Twitter engagement metrics (likes, retweets, replies, profile clicks, URL clicks, hashtag clicks).

    o Converted raw data into meaningful visual insights.

    o Applied real-time filters to make dashboards interactive and context-aware.

Soft Skills

- **Problem-Solving:** Identified issues in datasets and dashboards, like empty visualizations due to strict filters, and implemented practical solutions.

- **Analytical Thinking:** Extracted actionable insights from data patterns and user behavior metrics.

- **Time Management:** Completed all six tasks efficiently within the internship duration (15-09-2025 to 15-10-2025).

- **Professional Communication**: Documented processes, shared Power BI dashboards, and explained findings clearly in written reports.

Mentor Support and Guidance:

During my internship, I received excellent support and guidance from my mentors. Simran Ma'am played a pivotal role in assisting me throughout the internship. Whenever I faced any issues while working on the tasks, I could message her, and she would help me resolve them efficiently.

In cases where she was not able to provide direct guidance, she advised me to reach out to the training team for assistance. She also ensured I felt supported by saying, *"You are under my responsibility; I am always there to guide you."* This encouragement gave me confidence to work independently while knowing that mentorship was always available.

Challenges and Solutions

During my internship, I encountered a few challenges while working on the Power BI dashboards. One major difficulty was that some of the visualizations did not display any data. I was initially concerned that my dataset might be insufficient to fulfill the filtering criteria, and I worried that the dashboards would not meet the task requirements.

Additionally, I faced challenges while working with DAX (Data Analysis Expressions), as I was not fully familiar with it initially. Some calculations and measures were not working as expected, which slowed down my progress.

However, Simran Ma'am guided me and assured me that this was not a problem. She explained that the dataset provided during the training was limited, and the absence of data in some visuals was expected. Her guidance helped me focus on completing the tasks correctly.

Solution:

- Followed the instructions precisely and applied all filters as required.

- Used the dataset provided during training to complete the tasks.

- Conducted independent research using YouTube, Google, ChatGPT, Gemini, and other resources to understand and resolve issues with DAX calculations.

- Learned to focus on building dashboards with correct logic and filter application, even if the dataset did not produce visible results.

This experience taught me patience, problem-solving skills, and the importance of self-learning when facing technical challenges.

Conclusion

This internship provided me with valuable practical exposure to real-world data analytics and Power BI dashboard development. I enhanced my Python programming skills through data cleaning, preprocessing, and manipulation, and learned to work with datasets effectively.

Through building six interactive Power BI dashboards, I gained hands-on experience in data visualization, filter design, and DAX calculations. Despite facing challenges such as empty visuals due to dataset limitations and initial difficulties with DAX, I learned to problem-solve independently by researching solutions and applying them successfully.

The guidance and support from Simran Ma'am were instrumental in navigating technical challenges and understanding task requirements. Her mentorship taught me the importance of perseverance, self-learning, and structured problem-solving.

Overall, this internship strengthened my analytical, programming, and visualization skills, provided insights into real-time data analytics, and improved my confidence in developing professional dashboards and handling complex datasets.

References

1. **Python Documentation:** For implementing data cleaning, preprocessing, and model building.
   https://docs.python.org/3/

2. **Pandas & NumPy Libraries:** For data manipulation and statistical computation in Python.
   https://pandas.pydata.org/
   https://numpy.org/

3. **Scikit-learn Library:** For building and evaluating the heart disease prediction model.
   https://scikit-learn.org/

4. **Power BI Documentation:** For creating interactive dashboards and data visualizations.
   https://learn.microsoft.com/power-bi/

5. **Matplotlib & Seaborn:** For visual data exploration in Python.
   https://matplotlib.org/
   https://seaborn.pydata.org/

6. **NullClass Internship Resources:** Guidance and mentorship for project completion and documentation.