**Data Analytics Internship Report - Twitter Analytics**

**Name: Aditya Moreshwar Diwate  
Internship Title: Twitter Data Analytics Dashboard**

**Introduction**

During the internship, I was assigned a series of data visualization and analytics tasks cantered around Twitter data. The primary goal was to explore user engagement through multiple dimensions such as impressions, clicks, media interactions, and tweet characteristics, while applying time-based and content-specific filters. This report summarizes the methodology and solutions implemented for each task.

**Background**

Social media analytics is crucial for understanding public sentiment and user interaction. This project used Twitter data (likes, retweets, impressions, engagement, etc.) to extract meaningful visual insights by applying advanced Power BI features like DAX, filters, and time-based conditions.

**Learning Objectives**

* Understand and clean raw social media data.
* Apply DAX formulas to build logic-driven visualizations.
* Design interactive and conditional dashboards.
* Filter and analyse data based on multiple criteria.

**Activities and Tasks**

* **Data cleaning**: Treated nulls, corrected data types, and formatted timestamps.
* Created 10 individual tasks/visuals with complex filtering conditions.
* Developed measures to show/hide visuals based on IST time windows dynamically.
* Implemented logic to exclude specific tweet content (e.g., words with certain letters).

**Task 1: Average Engagement Rate and Total Impressions**

**Objective:** Visualize the average engagement rate and total impressions for tweets posted in the first half of 2020.

* **Filters:** Tweets with ≥ 100 impressions, likes = 0, shown only between 3 PM IST to 5 PM IST.
* **Key Implementation:**
  + Time filter applied via DAX using HOUR() on tweet timestamp.
  + Conditional display flag for visualization on the dashboard.

**Task 2: Proportion of Total Clicks**

* **Objective:** Show the proportion of URL clicks, user profile clicks, and hashtag clicks with drill-down per tweet.
* **Filters:** Tweets with impressions > 500.
* **Visual:** Pie chart with drill-down hierarchy.
* **Key Implementation:**
  + Aggregation of click types per tweet.
  + Drill-down configured to view total and individual click types.

**Task 3: Top 10% Tweets by Engagement Rate (Weekdays, Likes > 50)**

* **Objective:** Display the top 10% of tweets by engagement rate.
* **Filters:** Likes > 50, tweets posted on weekdays, character count < 30, shown only from 3 PM to 5 PM IST.
* **Visual:** Bar chart highlighting top tweets.
* **Key Implementation:**
  + Calculated percentile filter using DAX PERCENTILEX.INC.
  + Weekday check using the WEEKDAY() function.

**Task 4: Media Engagement vs. Media Views**

* **Objective:** Analyse the relationship between media engagements and media views.
* **Filters:** Replies > 10, highlight tweets with engagement rate > 5%, tweet date odd, word count > 50, visible 6 PM to 11 PM IST.
* **Visual:** Scatter plot with highlighted points.
* **Key Implementation:**
  + Conditional formatting for highlights.
  + Word count and odd date filters.

**Task 5: Click Interaction by Tweet Category**

* **Objective:** Sum of clicks broken down by tweet category (media, links, hashtags).
* **Filters:** At least one interaction present, tweet date even, word count > 40, shown 3 PM to 5 PM IST.
* **Visual:** Clustered bar chart.
* **Key Implementation:**
  + Categorization of tweets.
  + Filtering on tweet attributes and time window.

**Task 6: Top 10 Weekday Tweets by Engagement**

* **Objective:** Identify the top 10 tweets by retweets + likes.
* **Filters:** Posted on weekdays, impressions even, tweet date odd, word count < 30, shown 3 PM to 5 PM IST.
* **Visual:** Table to show top 10 Tweets.
* **Key Implementation:**
  + Summation measure for retweets + likes.
  + Date and word count filters.

**Task 7: Weekly Media Views vs Engagements**

**Objective:** Show media views and engagement trends by weekday, highlight spikes.

* **Filters:** Tweet impressions even, tweet date odd, character count > 30, removal of words with 'H' from a tweet, visible 3 PM to 5 PM & 7 AM to 11 AM IST.
* **Visual:** Dual-axis line and bar chart.
* **Key Implementation:**
  + Last quarter date range calculation.
  + Text filtering for character exclusion.
  + Spike detection via conditional formatting.

**Task 8: Engagement by App Opens Status**

* **Objective:** Compare engagement rates between tweets with and without app opens.
* **Filters:** Tweet impressions even, tweets posted between 9 AM and 5 PM on weekdays, tweet date odd, char count > 30, removal of words with ‘D’ from a tweet, graph visible 12 PM-6 PM & 7 AM-11 AM IST.
* **Visual:** Column chart comparing two categories.
* **Key Implementation:**
  + Segmentation by app opens.
  + Time and character filtering.

**Task 9: Tweet Engagement Breakdown**

* **Objective:** Compare replies, retweets, and likes for tweets above median media engagement.
* **Filters:** Tweets from June-August 2020, tweet date odd, media views even, char count > 20, removal of words with 'S' from a tweet, graph visible 3 PM-5 PM & 7 AM-11 AM IST.
* **Visual:** Stacked bar chart.
* **Key Implementation:**
  + Median calculation measure.
  + Date range filtering.
  + Text and numeric attribute filters.

**Task 10: Monthly Engagement Rate Trend by Media Presence**

* **Objective:** Trend of average engagement rate monthly, separated by media content presence.
* **Filters:** Engagement even, tweet date odd, char count > 20, removal of words with ‘C’ from a tweet, graph visible 3 PM-5 PM & 7 AM-11 AM IST.
* **Visual:** Line chart with two lines.
* **Key Implementation:**
  + Grouping by month.
  + Separate measures for media presence.

**Skills and Competencies**

* Data preprocessing and transformation.
* DAX proficiency for custom calculations and conditional logic.
* Advanced visualization techniques in Power BI.
* Time-based and content-based data filtering.

**Feedback and Evidence**

* Each visualization was tested using slicers and time logic to ensure correct visibility.
* Conditional messages were displayed for blank visuals using DAX-based text boxes.

**Challenges and Solutions**

* **Challenge:** Handling blank visuals when filter conditions were unmet.  
  **Solution:** Used DAX and dynamic text boxes to display context-aware messages.
* **Challenge:** Aligning IST logic with Power BI's UTC-based system.  
  **Solution:** Used UTCNOW() with time offset logic.
* **Challenge:** Avoiding repetitive logic across tasks.  
  **Solution:** Created reusable calculated columns and measures.

**Outcomes and Impact**

* Successfully developed a robust Power BI dashboard with 10 interactive visualizations.
* Learned how to manipulate, clean, and analyse real-world social media data.
* Gained practical experience in building production-ready analytics solutions.

**Conclusion**

These tasks reinforced practical skills in data cleaning, filtering, DAX calculations, time-based visual controls, and dashboard creation using Power BI. Implementing conditional display based on time and tweet attributes helped me understand dynamic reporting and user experience considerations in business intelligence.