



Project Title	Top Instagram Influencers Data (Cleaned) Dashboard
Tools	Tableau Desktop, SQL, Excel
Domain	Business Analyst
Project Difficulties level	intermediate

Dataset : Dataset is available in the given link. You can download it at your convenience.

[Click here to download data set](#)

### About Dataset

Instagram is an American photo and video sharing social networking service founded in 2010 by Kevin Systrom and Mike Krieger, and later acquired by Facebook Inc.. The app allows users to upload media that can be edited with filters and organized by hashtags and geographical tagging. Posts can be shared publicly or with preapproved followers. Users can browse other users' content by tag and location, view trending content, like photos, and follow other users to add their content to a personal feed.

Instagram network is very much used to influence people (the users followers) in a particular way for a specific issue - which can impact the order in some ways.

## About this file

In this file, basically there are 10 attributes. It has been ordered on basis of the rank which has been decided on basis of "followers".

rank: Rank of the Influencer on basis of number of followers they have

channel\_info: Username of the Instagrammer

influence score: Influence score of the users. It is calculated on basis of mentions, importance and popularity

posts: Number of posts they have made so far

followers: Number of followers of the user

avg\_likes: Average likes on instagrammer posts (total likes/ total posts)

60\_day\_eng\_rate: Last 60 days engagement rate of instagrammer as fraction of engagements they have done so far

new\_post\_avg\_like: Average likes they have on new posts

total Likes: Total likes the user has got on their posts. (in Billion)

country: Country or region of origin of the user.

## **Example: You can get the basic idea how you can create a project from here**

Creating a **Tableau project** on the "Top Instagram Influencers" dataset can be an engaging way to explore social media analytics, influencer impact, and marketing potential. Below is a comprehensive outline for a 5-year experienced analyst, covering essential steps, KPIs, calculated fields, exploratory data analysis (EDA), and visualizations.

### **Project Outline**

## Step 1: Import Data & Initial Cleaning

1. **Data Import:** Import the dataset into Tableau.
2. **Field Overview:** Verify the data types for each column:
  - `rank` (integer), `channel_info` (string), `influence_score` (float), `posts` (integer), `followers` (integer), `avg_likes` (float), `60_day_eng_rate` (float), `new_post_avg_like` (float), `total_likes` (integer), and `country` (string).
3. **Initial Cleaning:**
  - Remove or fill null values (if any).
  - Ensure columns like `influence_score` and `engagement_rate` are correctly formatted as percentages where needed.

## Step 2: Defining Key Performance Indicators (KPIs)

KPIs will measure influencer performance and engagement.

1. **Total Followers:** Sum of followers across all influencers.
2. **Average Engagement Rate:** The average `60_day_eng_rate`.
3. **Average Likes per Post:** Calculated by `total_likes / posts`.
4. **Influence Score Distribution:** Distribution analysis of `influence_score` across countries or ranks.
5. **Country with Highest Influencers:** A count of influencers by country to determine top regions.

## Step 3: Creating Calculated Fields

Calculated fields will enhance the data by creating new insights.

1. **Engagement Rate (ER) Calculation:**
  - Formula:  $([\text{avg\_likes}] / [\text{followers}]) * 100$

- This will standardize engagement rates to help compare across different influencer follower counts.

## 2. Growth Rate in New Post Likes:

- Formula: 
$$\frac{[\text{new\_post\_avg\_like}] - [\text{avg\_likes}]}{[\text{avg\_likes}] * 100}$$
- This field will indicate the growth or decline in average likes over time.

## 3. Like-to-Follower Ratio:

- Formula: 
$$\frac{[\text{total\_likes}]}{[\text{followers}]}$$
- Useful for identifying influencers whose total engagement surpasses their followers.

## Step 4: Exploratory Data Analysis (EDA)

EDA in Tableau will involve data aggregation, segmentation, and identifying outliers.

### 1. Distribution Analysis:

- Use histograms for `influence_score`, `followers`, and `avg_likes` to understand the spread and outliers.
- Create a scatter plot for `followers` vs. `avg_likes` to see if there's a correlation between follower size and average likes.

### 2. Top 10 Influencers:

- Rank influencers based on `influence_score` and show the top 10 for each country or globally.

### 3. Country-wise Analysis:

- Create a bar chart by `country` showing the total number of influencers and average engagement rate per country.

### 4. Trend Analysis for New Post Likes:

- Line chart showing the average `new_post_avg_like` over time to see engagement trends.

## Step 5: Data Visualization and Dashboard Design

Designing dashboards will provide a visually rich summary of influencer insights.

### 1. Dashboard 1: Overview of Influencer Performance

- **KPI Tiles:** Display the key metrics, including total followers, average engagement rate, and total likes.
- **Bar Chart:** Top 10 influencers based on `influence_score` with rank labels.
- **Map Visualization:** Geographic distribution of influencers by `country`, color-coded by the average engagement rate.

### 2. Dashboard 2: Engagement and Influence Metrics

- **Scatter Plot:** `followers` vs. `avg_likes` with `influence_score` as a size dimension to understand the influence-power relationship.
- **Line Chart:** Time-based trends in `new_post_avg_like`, with filters to select specific countries or influencers.

### 3. Dashboard 3: Country-Specific Insights

- **Bar Chart:** Total influencers and average `influence_score` per country.
- **Bubble Chart:** Show influencers' `like-to-follower_ratio` by country for a quick view of regional influence.

### 4. Dashboard 4: Engagement Trends

- **Heatmap:** Display `60_day_eng_rate` by influencer rank and `country`, useful for spotting high-engagement influencers.
- **Trendline:** Plotting changes in engagement rates across ranks or countries over time.

## Additional Visualizations

- **Boxplot** for `influence_score` distribution per country to understand

variance.

- **Histogram** for `avg_likes` to see typical engagement patterns across influencers.

### Sample Code (Calculated Fields)

In Tableau, calculated fields help in making custom metrics:

tableau

code

```
// Engagement Rate
```

```
([avg_likes] / [followers]) * 100
```

```
// Growth Rate in New Post Likes
```

```
([new_post_avg_like] - [avg_likes]) / [avg_likes] * 100
```

```
// Like-to-Follower Ratio
```

```
[total_likes] / [followers]
```

### Final Report Outline

1. **Project Introduction:** Overview of the objective to analyze top influencers.
2. **Data Cleaning Summary:** Any modifications or assumptions applied to the dataset.
3. **KPI Results:** Present KPIs with visual explanations.
4. **Detailed Analysis:** Insights into influencers' engagement trends, country-specific data, and top-performing individuals.
5. **Conclusion:** Summary of findings and potential applications for marketing strategies.

## **Example: You can get the basic idea how you can create a project from here**

Creating a comprehensive, advanced Tableau project and dashboard for Instagram influencer analysis requires thorough planning, detailed visualizations, and key metrics. This project can span across multiple aspects of data exploration, business insights, and actionable recommendations. Below is a step-by-step guide, broken into several components with Tableau code, KPIs, calculated fields, exploratory data analysis (EDA), and dashboard creation. This framework will be suitable for an experienced user with at least five years of experience in Tableau.

---

### **Project Structure Overview**

This Tableau project will be structured into 5 major sections:

1. **Data Preprocessing & Setup**
2. **Exploratory Data Analysis (EDA)**
3. **Key Performance Indicators (KPIs)**
4. **Calculated Fields**
5. **Visualizations & Dashboards**

### **Data Preprocessing & Setup**

#### **1. Data Cleaning:**

- **Handle Missing Values:** Use Tableau's built-in data cleaning tool to address any null or missing values. For instance, replace null values with a median or zero.
- **Correct Data Types:** Ensure columns are appropriately typed (e.g., numeric for `followers`, `posts`, `total_likes`, `influence_score`; string for `channel_info`, `country`).

#### **2. Data Structure:**

- Load the data using Tableau's connection options (CSV, Excel, or SQL Server).
  - Rename columns if needed for clarity (e.g., `channel_info` might be `Influencer Name`).
- 

## Step 1: Exploratory Data Analysis (EDA)

EDA is the key to identifying data trends, distributions, and anomalies.

### 20 EDA Operations:

#### 1. Distribution of Influence Score:

- Histogram of `influence_score`.

Tableau Code:

tableau

code

```
Histogram([influence_score])
```

○

#### 2. Top Influencers by Influence Score:

- Bar Chart sorted by `influence_score`.

#### 3. Country Distribution:

- Map View of influencers by `country`, color-coded by `influence_score`.

#### 4. Followers vs. Engagement Rate:

- Scatter Plot of `followers` vs. `60_day_eng_rate`.

#### 5. Followers Distribution:

- Histogram of `followers`.



**6. Average Likes per Post:**

- **Bar Chart** showing `avg_likes` per `posts`.

**7. Influence Score Correlation with Followers:**

- **Scatter Plot** of `influence_score` vs. `followers`.

**8. Posts vs. Followers:**

- **Scatter Plot** comparing `posts` and `followers`.

**9. Total Likes by Country:**

- **Bar Chart** of `total_likes` by `country`.

**10. Average Likes Distribution:**

- **Histogram** of `avg_likes`.

**11. Influence Score vs. Average Likes:**

- **Scatter Plot** showing `influence_score` vs. `avg_likes`.

**12. Engagement Rate over Time (if time data exists):**

- **Line Chart** of `60_day_eng_rate` over time.

**13. New Post Engagement Rate:**

- **Bar Chart** of `new_post_avg_like` by `channel_info`.

**14. Top Countries by Engagement Rate:**

- **Map** visualizing `60_day_eng_rate` by `country`.

**15. Posts and Engagement by Country:**

- **Bar Chart** of `posts` and `60_day_eng_rate` by `country`.

**16. Engagement Trend:**

- **Line Chart** of average `60_day_eng_rate`.

**17. Influence Score vs. New Posts:**

- **Scatter Plot** of `influence_score` and `new_post_avg_like`.

**18. Total Likes Distribution:**

- **Histogram** of `total_likes`.

**19. Top Influencers by Engagement:**

- **Bar Chart** showing top influencers by **60\_day\_eng\_rate**.

20. **Post Frequency by Country:**

- **Bar Chart** showing the number of posts by **country**.
- 

## Step 2: Key Performance Indicators (KPIs)

Key performance indicators (KPIs) help summarize the key metrics from the dataset. These KPIs are essential for high-level insights.

### 20 KPIs:

1. **Total Influencers:**

- **COUNT(channel\_info)** in Tableau.

2. **Average Influence Score:**

- **AVG([influence\_score])**

3. **Total Followers:**

- **SUM([followers])**

4. **Average Engagement Rate:**

- **AVG([60\_day\_eng\_rate])**

5. **Top 10 Influencers by Influence Score:**

- **TOP N Filter** on **influence\_score**.

6. **Average Likes per Post:**

- **AVG([avg\_likes])**

7. **Average Posts per Influencer:**

- **AVG([posts])**

8. **Influence Score by Country:**

- **AVG([influence\_score])** grouped by **country**.

9. **Total Likes:**

- **SUM([total\_likes])**

10. **Top Countries by Engagement Rate:**
    - **AVG([60\_day\_eng\_rate])** by **country**.
  11. **Average Followers per Influencer:**
    - **AVG([followers])**
  12. **Engagement Rate vs. Followers:**
    - **AVG([60\_day\_eng\_rate]) / AVG([followers])**
  13. **Influence Score to Post Ratio:**
    - **[influence\_score] / [posts]**
  14. **Total Posts Across Influencers:**
    - **SUM([posts])**
  15. **Average Likes per Country:**
    - **AVG([avg\_likes])** by **country**.
  16. **Engagement Rate by Rank:**
    - **AVG([60\_day\_eng\_rate])** by **rank**.
  17. **Post Frequency:**
    - **COUNT([posts])**
  18. **Growth in Followers Over Time:**
    - **[followers]** over time if time is available.
  19. **Total Likes by Rank:**
    - **SUM([total\_likes])** by **rank**.
  20. **Influence Growth:**
    - Difference between **new\_post\_avg\_like** and **avg\_likes**.
- 

### Step 3: Calculated Fields

Calculated fields provide new insights derived from the data.

#### 20 Calculated Fields:

### **Engagement Rate per Post:**

tableau

code

$$[\text{Engagement Rate per Post}] = [\text{avg\_likes}] / [\text{followers}]$$

1.

### **Post to Follower Ratio:**

tableau

code

$$[\text{Post/Follower Ratio}] = [\text{posts}] / [\text{followers}]$$

2.

### **Growth in Engagement:**

tableau

code

$$[\text{Engagement Growth}] = [\text{new\_post\_avg\_like}] - [\text{avg\_likes}]$$

3.

### **Influencer Engagement Potential:**

tableau

code

$$[\text{Engagement Potential}] = [\text{followers}] * [60\_day\_eng\_rate]$$

4.

### **Engagement Score:**

tableau

code

$$[\text{Engagement Score}] = [\text{influence\_score}] * [60\_day\_eng\_rate]$$

5.

### **Likes per Post Growth:**

tableau

code

```
[Likes per Post Growth] = [new_post_avg_like] - [avg_likes]
```

6.

### **Influence vs Engagement:**

tableau

code

```
[Influence vs Engagement] = [influence_score] /  
[60_day_eng_rate]
```

7.

### **Followers Growth:**

tableau

code

```
[Followers Growth] = [followers] - LOOKUP([followers], -1)
```

8.

### **Total Post Engagement:**

tableau

code

```
[Total Post Engagement] = [posts] * [avg_likes]
```

9.

### **Influence Growth Rate:**

tableau

code

```
[Influence Growth Rate] = ([new_post_avg_like] - [avg_likes]) /  
[avg_likes]
```

10.

### **Post Efficiency:**

tableau

code

```
[Post Efficiency] = [total_likes] / [posts]
```

11.

### **Influence Score per Post:**

tableau

code

```
[Influence Score per Post] = [influence_score] / [posts]
```

12.

### **Likes per Follower:**

tableau

code

```
[Likes per Follower] = [avg_likes] / [followers]
```

13.

### **New Post Engagement per Follower:**

tableau

code

```
[New Post Engagement per Follower] = [new_post_avg_like] /
```

[followers]

14.

### **Engagement per Rank:**

tableau

code

[Engagement per Rank] = [60\_day\_eng\_rate] / [rank]

15.

### **Average Engagement per Country:**

tableau

code

[Average Engagement by Country] = AVG([60\_day\_eng\_rate])

16.

### **Likes per Rank:**

tableau

code

[Likes per Rank] = [total\_likes] / [rank]

17.

### **Influence per Country:**

tableau

code

[Influence per Country] = [influence\_score] / [country]

18.

### **Followers per Post:**

tableau

code

```
[Followers per Post] = [followers] / [posts]
```

19.

### Engagement per Rank:

tableau

code

```
[Engagement per Rank] = [60To develop a 100-page comprehensive  
Tableau project with dashboards and analysis for Instagram  
Influencers using the provided dataset, here's a detailed  
step-by-step framework that covers everything from Exploratory  
Data Analysis (EDA) to Key Performance Indicators (KPIs),  
Calculated Fields, and Dashboard Development. This guide is  
tailored for an experienced Tableau user with 5+ years of  
experience.
```

20.

---

## Project Overview

The project aims to analyze the top Instagram influencers' data with the following columns:

- **rank**
- **channel\_info** (Influencer's name)
- **influence\_score**
- **posts**



- **followers**
- **avg\_likes**
- **60\_day\_eng\_rate**
- **new\_post\_avg\_like**
- **total\_likes**
- **country**

The deliverable will include visualizations, KPIs, calculated fields, and multiple dashboards with deep insights into influencer performance, engagement, and geographical trends.

---

## **Step 1: Data Preprocessing & Setup**

### **1.1. Data Cleaning and Validation**

- **Handle Missing Values:** Identify and handle null values. You could either replace them with average values or remove rows if critical.
- **Verify Data Types:** Ensure that **followers**, **posts**, **total\_likes**, and **influence\_score** are numeric, and **channel\_info** and **country** are string types.

### **1.2. Data Import into Tableau**

- Connect Tableau to your data source (CSV, Excel, or SQL database).
- 

## **Step 2: Exploratory Data Analysis (EDA)**

In this step, perform various operations to gain insights into the data.

### **20 EDA Operations**

## 1. Distribution of **Influence Score**:

- **Histogram** to analyze the spread.
- Tableau: `Histogram([influence_score])`

## 2. Top 10 Influencers by **Influence Score**:

- **Bar Chart** showing the top influencers based on influence score.

## 3. Geographical Distribution by Country:

- **Map** chart using `country` and color code by `influence_score`.

## 4. Followers vs. Engagement Rate:

- **Scatter Plot** comparing `followers` vs. `60_day_eng_rate`.

## 5. Followers Distribution:

- **Histogram** for followers to understand influencer size distribution.

## 6. Total Likes per Influencer:

- **Bar Chart** showing total likes by influencers.

## 7. Engagement Rate Distribution:

- **Box Plot** for `60_day_eng_rate` to see the spread and outliers.

## 8. Correlation between **Influence Score** and **Followers**:

- **Scatter Plot** to show how influence correlates with follower count.

## 9. Top Countries by **Total Likes**:

- **Bar Chart** of `total_likes` by `country`.

## 10. New Post Engagement:

- **Bar Chart** of `new_post_avg_like` by influencer.

## 11. Average Likes Distribution:

- **Histogram** to understand the average likes distribution.

## 12. Engagement Rate Trends:

- **Line Chart** showing the trend of `60_day_eng_rate`.

## 13. Total Likes and Followers Comparison:

- **Scatter Plot** showing `total_likes` vs. `followers`.

14. **Influencers with Most Posts:**
    - **Bar Chart** for top influencers based on the number of posts.
  15. **Post Engagement:**
    - **Bar Chart** for average engagement per post using `avg_likes`.
  16. **Country-wise Engagement:**
    - **Map View** showing engagement levels across countries.
  17. **Influence Score vs. Posts:**
    - **Scatter Plot** comparing `influence_score` and `posts`.
  18. **Engagement Rate Distribution by Country:**
    - **Box Plot** of `60_day_eng_rate` by `country`.
  19. **Total Likes vs. Rank:**
    - **Scatter Plot** of `total_likes` vs. `rank` to see if higher-ranked influencers get more likes.
  20. **Posts and Engagement:**
    - **Line Chart** comparing `posts` and `60_day_eng_rate`.
- 

### Step 3: Key Performance Indicators (KPIs)

KPIs provide quick snapshots of influencer performance.

#### 20 KPIs

1. **Total Influencers:** Count of `channel_info`.
  - `COUNT([channel_info])`
2. **Average Influence Score:** Average of `influence_score`.
  - `AVG([influence_score])`
3. **Total Followers:** Sum of `followers`.
  - `SUM([followers])`

4. **Average Engagement Rate:** Average of `60_day_eng_rate`.
  - `AVG([60_day_eng_rate])`
5. **Average Likes per Post:** Average of `avg_likes`.
  - `AVG([avg_likes])`
6. **Top 10 Influencers by Influence Score:** Filtered to show top 10 influencers.
  - Use a **Rank Filter** on `influence_score`.
7. **Total Likes:** Sum of `total_likes`.
  - `SUM([total_likes])`
8. **Top Countries by Influence:** Countries with the highest sum of `influence_score`.
  - `SUM([influence_score])` by `country`.
9. **Engagement Rate by Country:** Average `60_day_eng_rate` by `country`.
  - `AVG([60_day_eng_rate])` by `country`.
10. **Post Frequency:** Average posts per influencer.
  - `AVG([posts])`
11. **Likes per Follower:** `avg_likes` divided by `followers`.
  - `[avg_likes] / [followers]`
12. **Influence Score by Rank:** Average `influence_score` by `rank`.
  - `AVG([influence_score])` by `rank`
13. **Top Countries by Followers:** Countries with the most total followers.
  - `SUM([followers])` by `country`.
14. **Average Posts per Influencer:** Average of `posts`.
  - `AVG([posts])`
15. **Engagement Growth:** Change in engagement between two periods.
  - `([new_post_avg_like] - [avg_likes])`
16. **Top Influencers by Engagement:** Top influencers based on engagement

rate.

- TOP N filter by 60\_day\_eng\_rate.

17. **Engagement per Rank:** Divide 60\_day\_eng\_rate by rank.

- $[60\_day\_eng\_rate] / [rank]$

18. **Post Efficiency:** Likes per post divided by the number of posts.

- $[total\_likes] / [posts]$

19. **Engagement by Post:** Average engagement per post for each influencer.

- $AVG([new\_post\_avg\_like]) / [posts]$

20. **Followers per Post:** Number of followers per post.

- $[followers] / [posts]$

---

## Step 4: Calculated Fields

Calculated fields allow for more customized analysis and derive new insights.

### 20 Calculated Fields

#### Engagement Rate per Post:

tableau

code

$[Engagement\ Rate\ per\ Post] = [avg\_likes] / [followers]$

1.

#### Followers Growth:

tableau

code

$[Followers\ Growth] = [followers] - LOOKUP([followers], -1)$

2.

### **Post to Follower Ratio:**

tableau

code

$$[\text{Post/Follower Ratio}] = [\text{posts}] / [\text{followers}]$$

3.

### **Likes per Rank:**

tableau

code

$$[\text{Likes per Rank}] = [\text{total\_likes}] / [\text{rank}]$$

4.

### **Engagement Score:**

tableau

code

$$[\text{Engagement Score}] = [\text{influence\_score}] * [\text{60\_day\_eng\_rate}]$$

5.

### **Engagement Growth:**

tableau

code

$$[\text{Engagement Growth}] = [\text{new\_post\_avg\_like}] - [\text{avg\_likes}]$$

6.

### **Influence per Post:**

tableau

code

$$[\text{Influence per Post}] = [\text{influence\_score}] / [\text{posts}]$$

7.

### **Influence vs Engagement:**

tableau

code

```
[Influence vs Engagement] = [influence_score] /  
[60_day_eng_rate]
```

8.

### **Post Efficiency:**

tableau

code

```
[Post Efficiency] = [total_likes] / [posts]
```

9.

### **Growth in Likes per Post:**

tableau

code

```
[Growth in Likes per Post] = [new_post_avg_like] - [avg_likes]
```

10.

### **Engagement Rate per Country:**

tableau

code

```
[Engagement Rate per Country] = AVG([60_day_eng_rate])
```

11.

### **Likes per Follower:**

tableau

code

```
[Likes per Follower] = [avg_likes] / [followers]
```

12.

### New Post Engagement:

tableau

code

```
[New Post Engagement] = [new_post_avg_like] - [avg_likes]
```

13.

### Influence Growth Rate:

tableau

code

```
[Influence Growth Rate] = ([new_post_avg_like] - [avg_likes]) /  
[avg_likes]
```

14.

### Engagement vs Followers:

tableau

code

```
[Engagement vs Followers] = [60_day_eng_rate] / [followers]
```

15.

16. **Average Engagement per Influencer:** ``tableCreating a comprehensive and advanced **Tableau Project** and **Dashboard** for Instagram Influencers data requires thoughtful planning and a detailed implementation of visualizations, KPIs, Calculated Fields, and Dashboards. Here's a step-by-step guide for a **100-page** project that provides insights from the data, using key columns such



as `rank`, `channel_info`, `influence_score`, `posts`, `followers`, `avg_likes`, `60_day_eng_rate`, `new_post_avg_like`, `total_likes`, and `country`.

---

## Overview

This project will cover:

1. **Exploratory Data Analysis (EDA):** Uncover data patterns, distributions, and relationships.
  2. **KPIs:** Critical metrics to gauge influencer performance.
  3. **Calculated Fields:** Derive useful insights.
  4. **Visualizations:** Help tell the story of the data.
  5. **Dashboards:** Bring all insights together into a cohesive view.
- 

## Step 1: Data Setup & Preprocessing

1. **Import Data:** Start by importing your dataset into Tableau (CSV, Excel, or SQL).
  2. **Handle Missing Data:** Cleanse the data by replacing `null` values with appropriate defaults (e.g., 0, median values).
  3. **Data Types Check:** Ensure that `followers`, `posts`, `total_likes`, and `influence_score` are numerical; `country` and `channel_info` are categorical (string).
- 

## Step 2: Exploratory Data Analysis (EDA)

EDA helps identify key patterns and understand the distribution of data points.

## 20 EDA Operations

1. **Distribution of Influence Score:** Use a histogram to examine the spread of `influence_score`.
2. **Top Influencers by Influence Score:** Create a bar chart that ranks influencers based on `influence_score`.
3. **Followers Distribution:** Use a histogram to visualize how followers are distributed across influencers.
4. **Influence Score vs. Followers:** Scatter plot to see if `influence_score` correlates with `followers`.
5. **Posts vs. Followers:** Scatter plot to analyze the relationship between posts and followers.
6. **Average Likes per Post:** Bar chart showing the average likes per post for each influencer.
7. **Country Distribution:** Map visualization to visualize the `country` of each influencer.
8. **Engagement Rate by Country:** Bar chart displaying the `60_day_eng_rate` by `country`.
9. **Total Likes vs. Followers:** Scatter plot for `total_likes` and `followers`.
10. **Engagement Growth:** Line graph showing how `60_day_eng_rate` changes over time (if time data is available).
11. **New Post Engagement:** Bar chart showing `new_post_avg_like` for each influencer.
12. **Likes per Follower:** Calculated field for `avg_likes / followers`, visualized as a bar chart.

13. **Top Influencers by Engagement Rate:** Sort by `60_day_eng_rate` to identify the most engaging influencers.
  14. **Influence Score by Country:** Average influence score grouped by `country`.
  15. **Posts and Engagement:** Scatter plot of `posts` vs. `60_day_eng_rate`.
  16. **Engagement vs. Rank:** Bar chart comparing `60_day_eng_rate` with `rank`.
  17. **Engagement vs. New Post Likes:** Scatter plot comparing `60_day_eng_rate` with `new_post_avg_like`.
  18. **Total Likes Distribution:** Histogram showing the distribution of `total_likes`.
  19. **Top 10 Countries by Followers:** Bar chart of `followers` aggregated by `country`.
  20. **Likes per Post Distribution:** Distribution plot for `avg_likes`.
- 

### Step 3: Key Performance Indicators (KPIs)

KPIs provide high-level summaries of influencer performance.

#### 20 KPIs

1. **Total Influencers:** `COUNT([channel_info])`
2. **Average Influence Score:** `AVG([influence_score])`
3. **Total Followers:** `SUM([followers])`
4. **Average Engagement Rate:** `AVG([60_day_eng_rate])`
5. **Top 10 Influencers by Influence Score:** `TOP N` filter on `influence_score`.
6. **Total Likes:** `SUM([total_likes])`

7. **Followers per Post:**  $\text{AVG}([\text{followers}]) / [\text{posts}]$
  8. **Post Frequency:**  $\text{AVG}([\text{posts}])$
  9. **Influence Score per Rank:**  $\text{AVG}([\text{influence\_score}])$  by rank.
  10. **Engagement per Follower:**  $\text{AVG}([60\_day\_eng\_rate]) / [\text{followers}]$
  11. **Likes per Post:**  $\text{AVG}([\text{avg\_likes}])$
  12. **Top Countries by Engagement:**  $\text{AVG}([60\_day\_eng\_rate])$  by country.
  13. **Influence Score by Country:**  $\text{AVG}([\text{influence\_score}])$  by country.
  14. **Likes per Follower:**  $\text{AVG}([\text{avg\_likes}]) / [\text{followers}]$
  15. **Engagement Growth:**  $([\text{new\_post\_avg\_like}] - [\text{avg\_likes}])$
  16. **Top Influencers by Engagement:** TOP N filter on 60\_day\_eng\_rate.
  17. **Total Posts per Country:**  $\text{SUM}([\text{posts}])$  by country.
  18. **Engagement per Rank:**  $\text{AVG}([60\_day\_eng\_rate])$  by rank.
  19. **Followers Growth:**  $([\text{followers}] - \text{LOOKUP}([\text{followers}], -1))$
  20. **Average Likes per Influencer:**  $\text{AVG}([\text{avg\_likes}])$  by channel\_info.
- 

#### Step 4: Calculated Fields

Calculated fields provide additional insights that can help with decision-making.

#### 20 Calculated Fields

##### Engagement Rate per Post:

tableau

code

$[\text{Engagement Rate per Post}] = [\text{avg\_likes}] / [\text{followers}]$

1.

### **Post Efficiency:**

tableau

code

```
[Post Efficiency] = [total_likes] / [posts]
```

2.

### **Engagement Growth:**

tableau

code

```
[Engagement Growth] = [new_post_avg_like] - [avg_likes]
```

3.

### **Influence vs Engagement:**

tableau

code

```
[Influence vs Engagement] = [influence_score] /  
[60_day_eng_rate]
```

4.

### **Likes per Follower:**

tableau

code

```
[Likes per Follower] = [avg_likes] / [followers]
```

5.

### **Followers Growth:**

tableau

code

```
[Followers Growth] = [followers] - LOOKUP([followers], -1)
```

6.

### **New Post Engagement:**

tableau

code

```
[New Post Engagement] = [new_post_avg_like] - [avg_likes]
```

7.

### **Engagement Score:**

tableau

code

```
[Engagement Score] = [influence_score] * [60_day_eng_rate]
```

8.

### **Influence Growth Rate:**

tableau

code

```
[Influence Growth Rate] = ([new_post_avg_like] - [avg_likes]) /  
[avg_likes]
```

9.

### **Posts per Follower:**

tableau

code

```
[Posts per Follower] = [posts] / [followers]
```

---

## Step 5: Visualizations and Dashboards

Now, bring together all your KPIs and calculated fields into **interactive dashboards**.

### 10 Dashboards

1. **Overview Dashboard:** A high-level overview of key KPIs, such as Total Influencers, Followers, and Total Likes.
2. **Engagement Dashboard:** Focus on engagement metrics, including engagement rate and likes per post.
3. **Geography Dashboard:** A map view showing influencer distribution by country, with country-specific engagement rates.
4. **Influencer Comparison:** Bar charts comparing top influencers by influence score, followers, and engagement rate.
5. **Post Performance Dashboard:** Metrics focused on posts—frequency, likes per post, and total likes.
6. **Influence Growth:** A dashboard showing influence growth trends using calculated fields like **Influence Growth Rate**.
7. **Engagement Analysis:** A deep dive into engagement metrics, such as **Engagement per Post** and **Likes per Follower**.
8. **Country-wise Performance:** A breakdown of performance by country, showing likes, followers, and engagement rates by country.
9. **Followers vs. Engagement:** Scatter plots and trend lines showing the relationship between followers and engagement rate.
10. **Top Influencers by Engagement:** A dashboard ranking influencers by their engagement rate, influence score, and posts.

[Reference link](#)



You can get practice and experience for SQL Project from here

## SQL Project: Analysis of Top Instagram Influencers Data

This project is designed for someone with approximately five years of experience in data analysis and SQL, aiming to dive into exploratory data analysis (EDA), data transformations, and visualizations for insight extraction on Instagram influencers.

### Dataset Schema

- **rank**: Rank of the influencer
- **channel\_info**: Name or handle of the influencer
- **influence\_score**: Score reflecting influencer's overall impact
- **posts**: Number of posts by the influencer
- **followers**: Total follower count
- **avg\_likes**: Average number of likes per post
- **60\_day\_eng\_rate**: Engagement rate over the past 60 days
- **new\_post\_avg\_like**: Average likes on recent posts
- **total\_likes**: Total likes across all posts
- **country**: Country of the influencer

---

### Step 1: Database Setup and Data Insertion

First, create the database and import the data. We will use SQL for data manipulation, and the project may also involve Python for visualization.

Sql code

```
-- Create the database  
CREATE DATABASE InstagramInfluencers;  
USE InstagramInfluencers;
```

```
-- Create the table
CREATE TABLE influencers (
    rank INT PRIMARY KEY,
    channel_info VARCHAR(100),
    influence_score DECIMAL(5, 2),
    posts INT,
    followers BIGINT,
    avg_likes INT,
    sixty_day_eng_rate DECIMAL(4, 2),
    new_post_avg_like INT,
    total_likes BIGINT,
    country VARCHAR(50)
);
```

### **Insert Data**

Load the data into this table from a CSV file or directly insert sample data if testing manually.

Sql code

– Sample insert

```
INSERT INTO influencers (rank, channel_info, influence_score,
posts, followers, avg_likes, sixty_day_eng_rate,
new_post_avg_like, total_likes, country)
VALUES
```

```
(1, 'influencer1', 85.50, 500, 1000000, 20000, 4.50, 21000,
10000000, 'USA'),
(2, 'influencer2', 80.25, 450, 950000, 18000, 4.20, 18500,
8500000, 'UK');
```

---

## Step 2: Exploratory Data Analysis (EDA)

### 1. Distribution of Followers

Sql code

```
-- Find the distribution of followers
SELECT
    country,
    AVG(followers) AS avg_followers,
    MAX(followers) AS max_followers,
    MIN(followers) AS min_followers
FROM influencers
GROUP BY country
ORDER BY avg_followers DESC;
```

This query helps to understand which countries have the most followed influencers.

### 2. Top Influencers by Influence Score

Sql code

```
-- Retrieve top influencers based on influence_score
SELECT
```

```
rank,  
channel_info,  
influence_score,  
followers,  
avg_likes  
FROM influencers  
ORDER BY influence_score DESC  
LIMIT 10;
```

This identifies the highest-ranking influencers by influence score, giving insights into their audience reach and engagement.

### 3. Engagement Rate Analysis

Sql code

```
-- Calculate engagement rate statistics  
SELECT  
    AVG(sixty_day_eng_rate) AS avg_eng_rate,  
    MAX(sixty_day_eng_rate) AS max_eng_rate,  
    MIN(sixty_day_eng_rate) AS min_eng_rate  
FROM influencers;
```

This query provides average, maximum, and minimum engagement rates to assess overall engagement trends among top influencers.

### 4. Influencers with High Engagement but Low Followers

Sql code

-- Identify influencers with high engagement but relatively low followers

```
SELECT
    channel_info,
    followers,
    sixty_day_eng_rate
FROM influencers
WHERE followers < 500000
AND sixty_day_eng_rate > 5
ORDER BY sixty_day_eng_rate DESC;
```

This finds influencers who, despite a lower follower count, have high engagement rates, making them potential high-return micro-influencers.

---

### Step 3: Visualization and Analysis with Python

After querying, we can use Python with libraries like `matplotlib` and `seaborn` to visualize the data.

#### Example Visualization: Distribution of Followers by Country

Python code

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load data into pandas DataFrame
```

```
data = pd.read_sql_query("SELECT country, AVG(followers) AS  
avg_followers FROM influencers GROUP BY country", connection)  
  
# Plot the data  
plt.figure(figsize=(12, 6))  
sns.barplot(data=data, x="country", y="avg_followers")  
plt.xticks(rotation=45)  
plt.title("Average Followers by Country")  
plt.xlabel("Country")  
plt.ylabel("Average Followers")  
plt.show()
```

### **Visualization: Engagement vs Follower Count**

Python code

```
# Query for engagement vs followers  
data = pd.read_sql_query("SELECT followers, sixty_day_eng_rate  
FROM influencers", connection)  
  
# Scatter plot  
plt.figure(figsize=(10, 6))  
sns.scatterplot(data=data, x="followers",  
y="sixty_day_eng_rate", hue="country")  
plt.title("Engagement Rate vs Followers")  
plt.xlabel("Followers")  
plt.ylabel("60-Day Engagement Rate")
```

```
plt.show()
```

---

## Step 4: Advanced Analysis

### 1. Growth Potential (New Post Like Analysis)

Sql code

```
-- Find influencers whose new post average likes have increased  
by 10% over average likes  
SELECT  
    channel_info,  
    avg_likes,  
    new_post_avg_like  
FROM influencers  
WHERE new_post_avg_like > avg_likes * 1.1  
ORDER BY new_post_avg_like DESC;
```

This finds influencers experiencing growth in engagement, useful for targeting those with increasing influence.

### 2. Country-wise Average Influence Score

Sql code

```
-- Find average influence score by country  
SELECT  
    country,  
    AVG(influence_score) AS avg_influence_score
```

```
FROM influencers  
GROUP BY country  
ORDER BY avg_influence_score DESC;
```

---

## **Final Report Summary**

### **1. Influencer Rankings and Distribution:**

- Analyzed top influencers, their reach, and engagement patterns.

### **2. Country-based Trends:**

- Identified which countries have the most active influencer bases and higher follower counts.

### **3. Micro-Influencers with High Engagement:**

- Discovered lower-followed influencers with significant engagement.

### **4. Growth Indicators:**

- Showcased influencers with rising engagement rates on recent posts.