Instagram Influencers Data Analysis Report

Name: Isha Koolwal

Date: 15/08/2025

Tools: ML, Python, SQL, Excel

1. Introduction

This project involves the analysis of Instagram influencer metrics including followers, engagement rates, likes, and influence scores. The aim is to identify top influencers, understand engagement patterns, and build predictive models to estimate influence scores. The dataset covers a range of influencers across multiple countries.

2. Dataset Overview

The dataset contains information for top Instagram influencers with the following columns:
- Rank: Position based on number of followers
- Channel: Instagram username or handle
- Influence Score: Calculated based on mentions, importance, and popularity
- Posts: Number of posts made
- Followers: Number of followers
- Average Likes: Mean likes per post
- 60-Day Engagement Rate: Engagement rate in the past 60 days
- New Post Average Like: Average likes for recent posts
- Total Likes: Cumulative likes across all posts
- Country: Influencer's country of origin

3. Correlation Matrix

The correlation matrix identifies relationships between numeric metrics such as followers, likes, engagement rates, and influence scores.
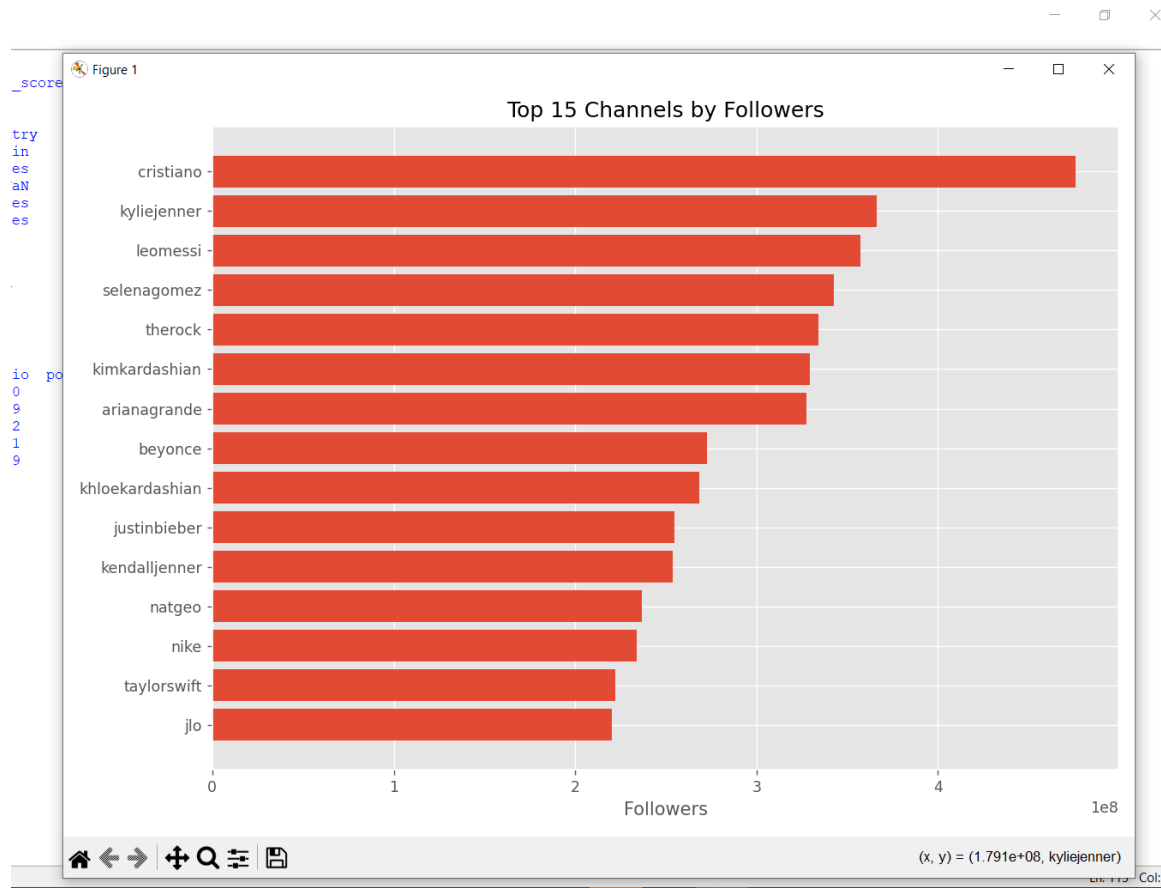
Correlation Matrix
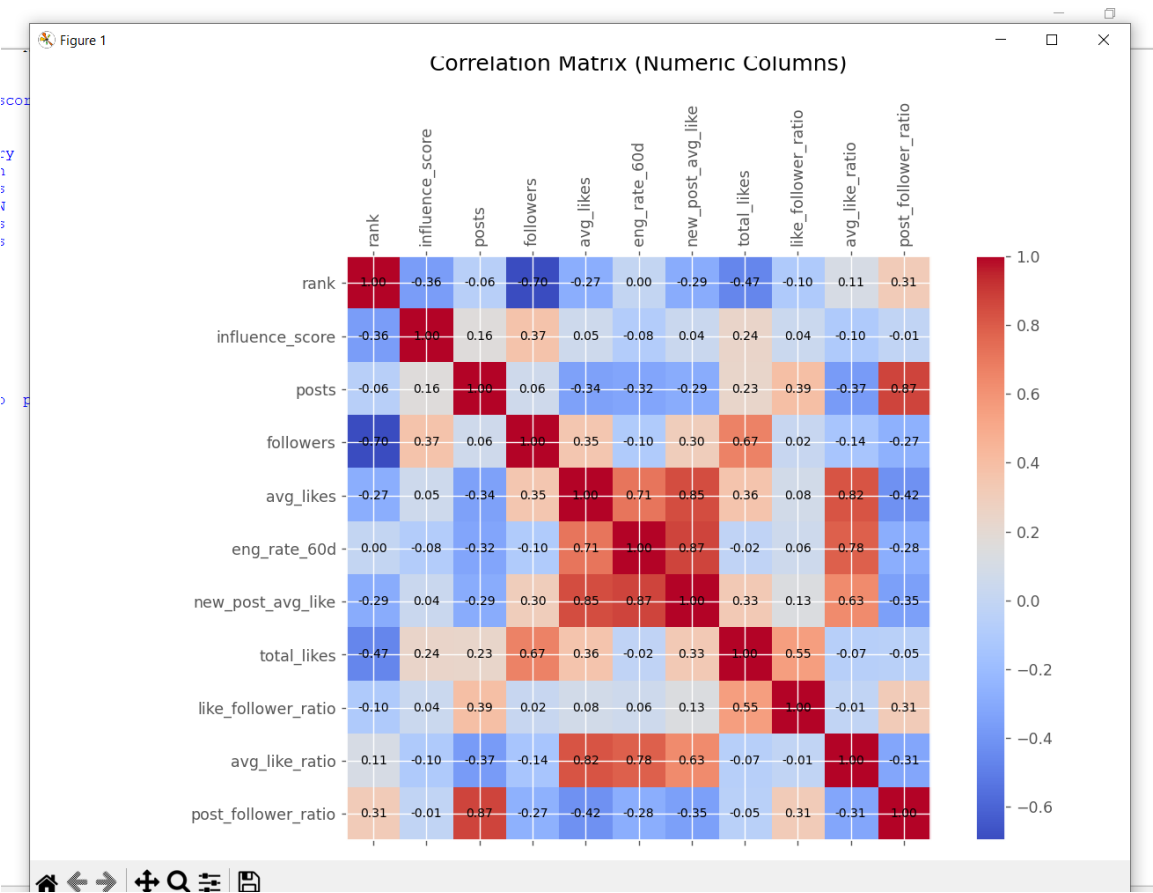
4. Top Influencers by Followers

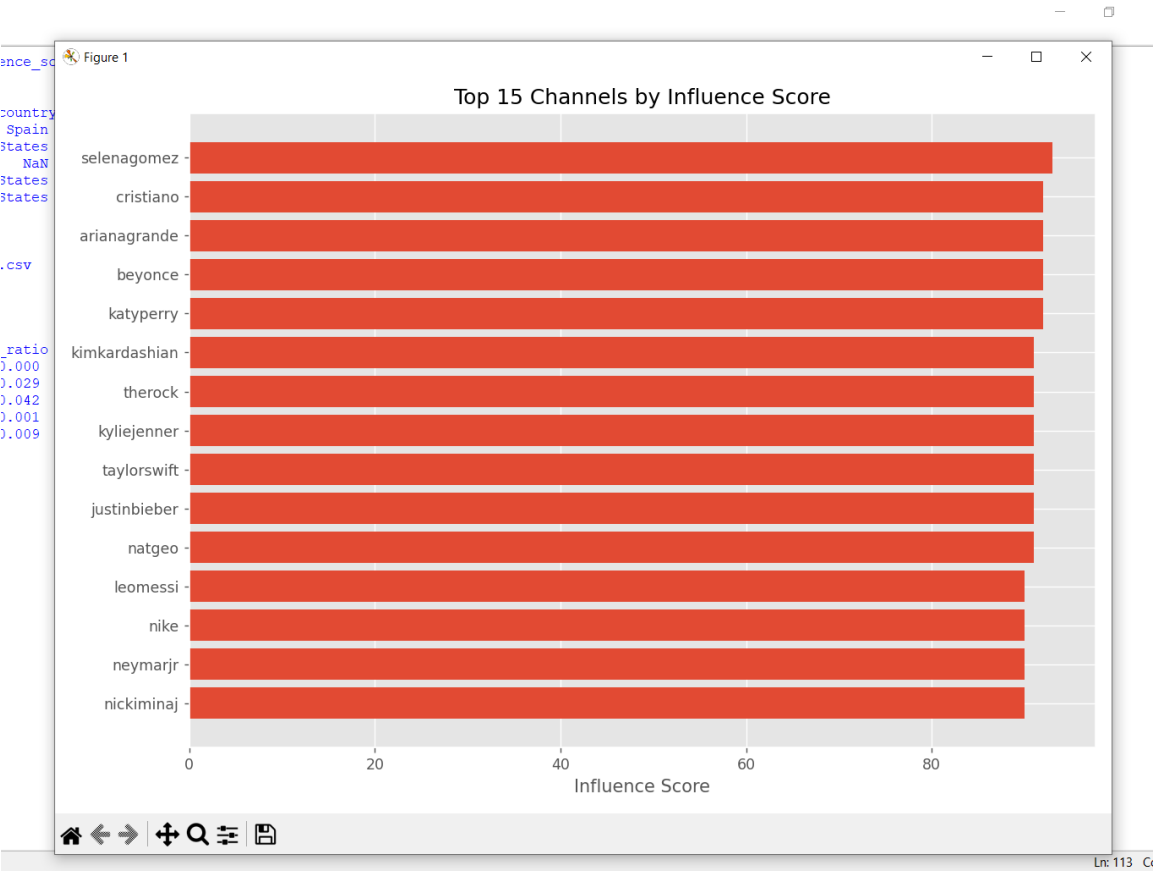This section presents the top 15 influencers ranked by their follower count.

Top 15 Channels by Followers

5. Top Influencers by Influence Score

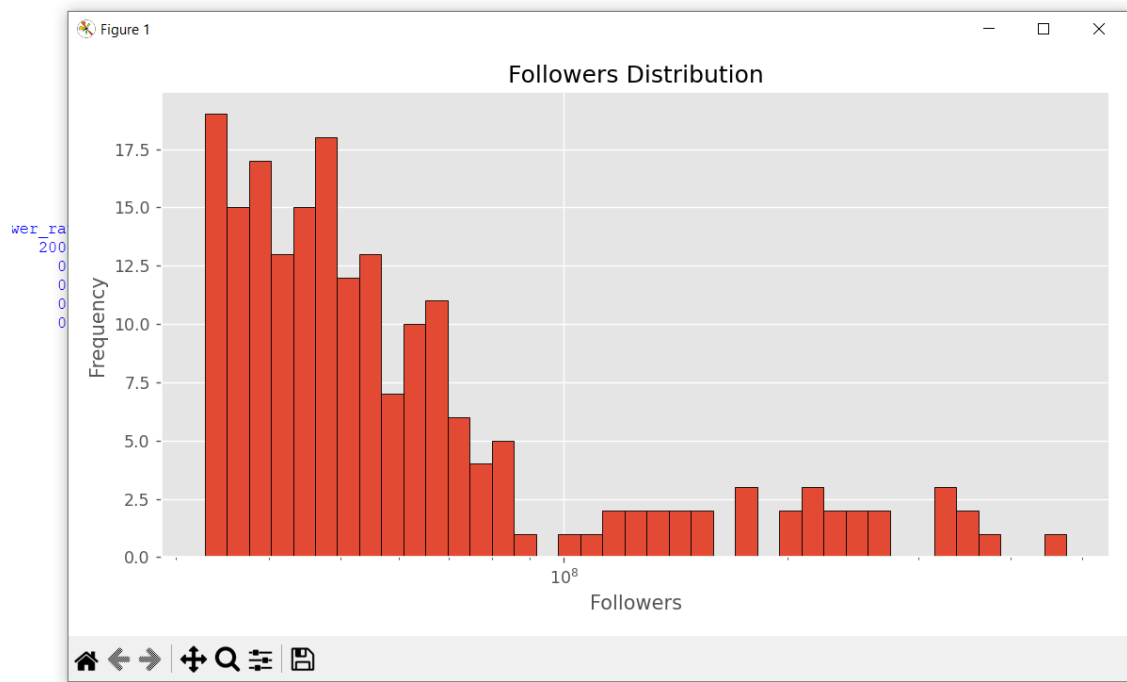This section highlights the top 15 influencers ranked by influence score.
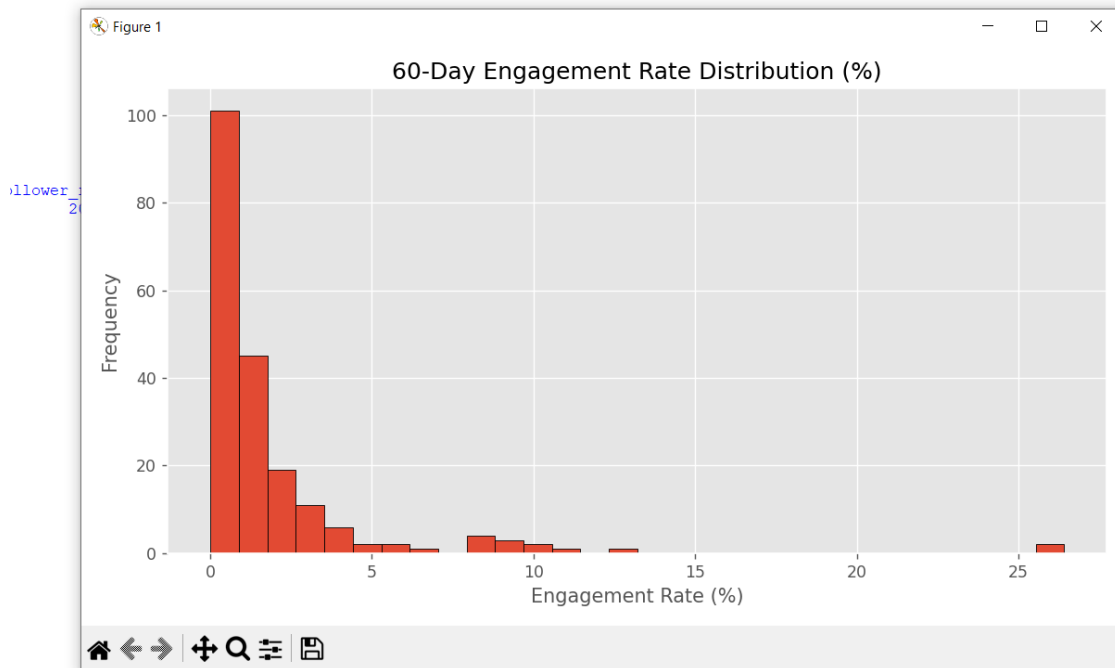
Top 15 Channels by Influence Score

6. Distributions of Key Metrics

Histograms showing the distribution of followers, average likes, and engagement rates help understand the spread and skewness of these metrics.
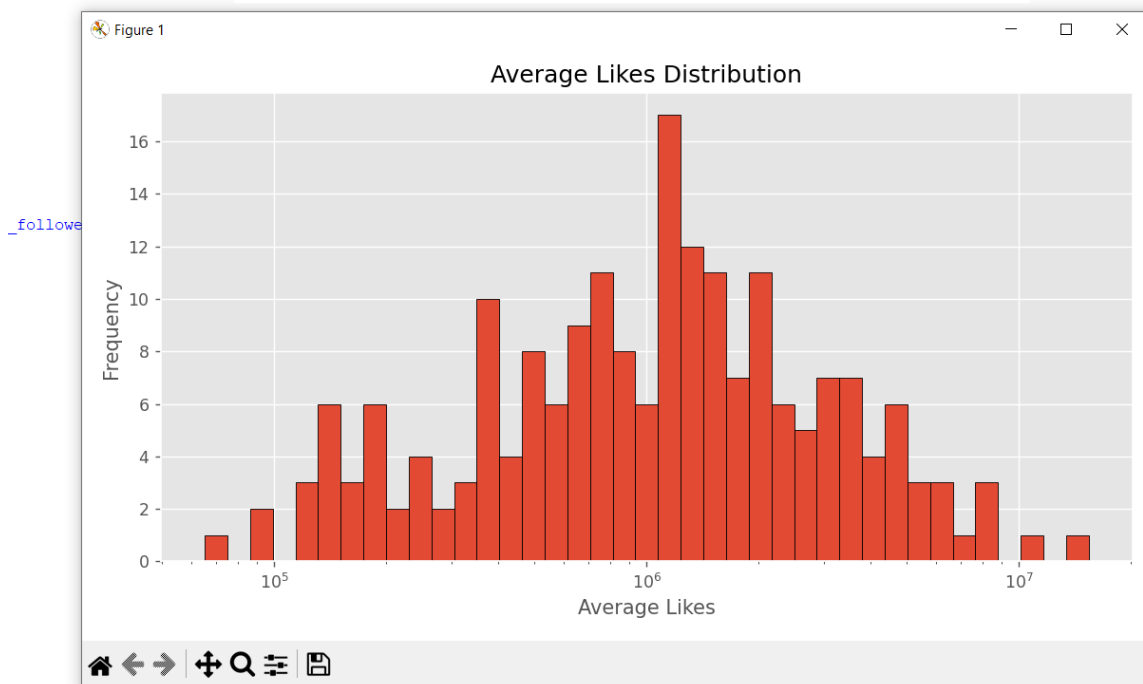
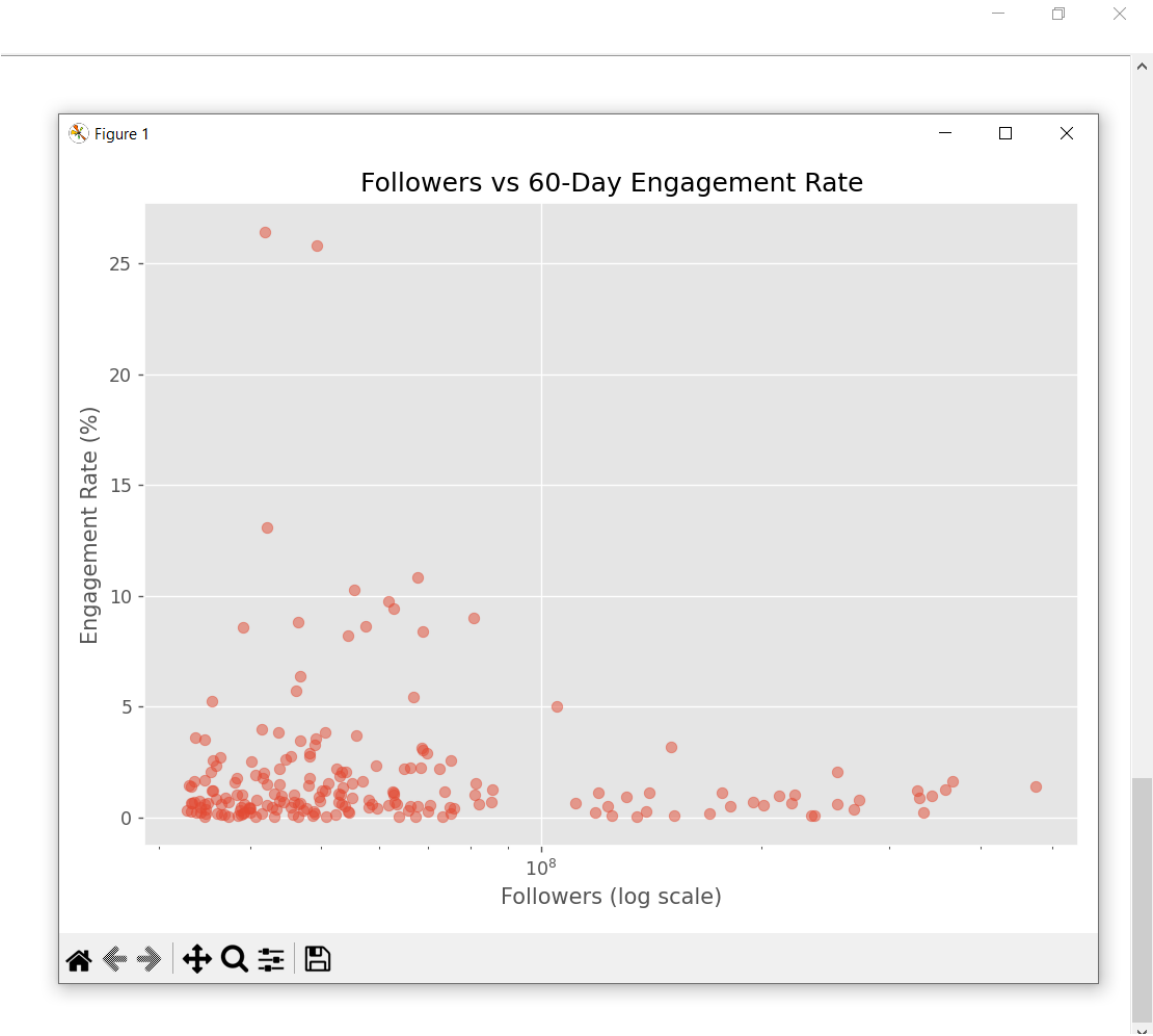# Followers Distribution

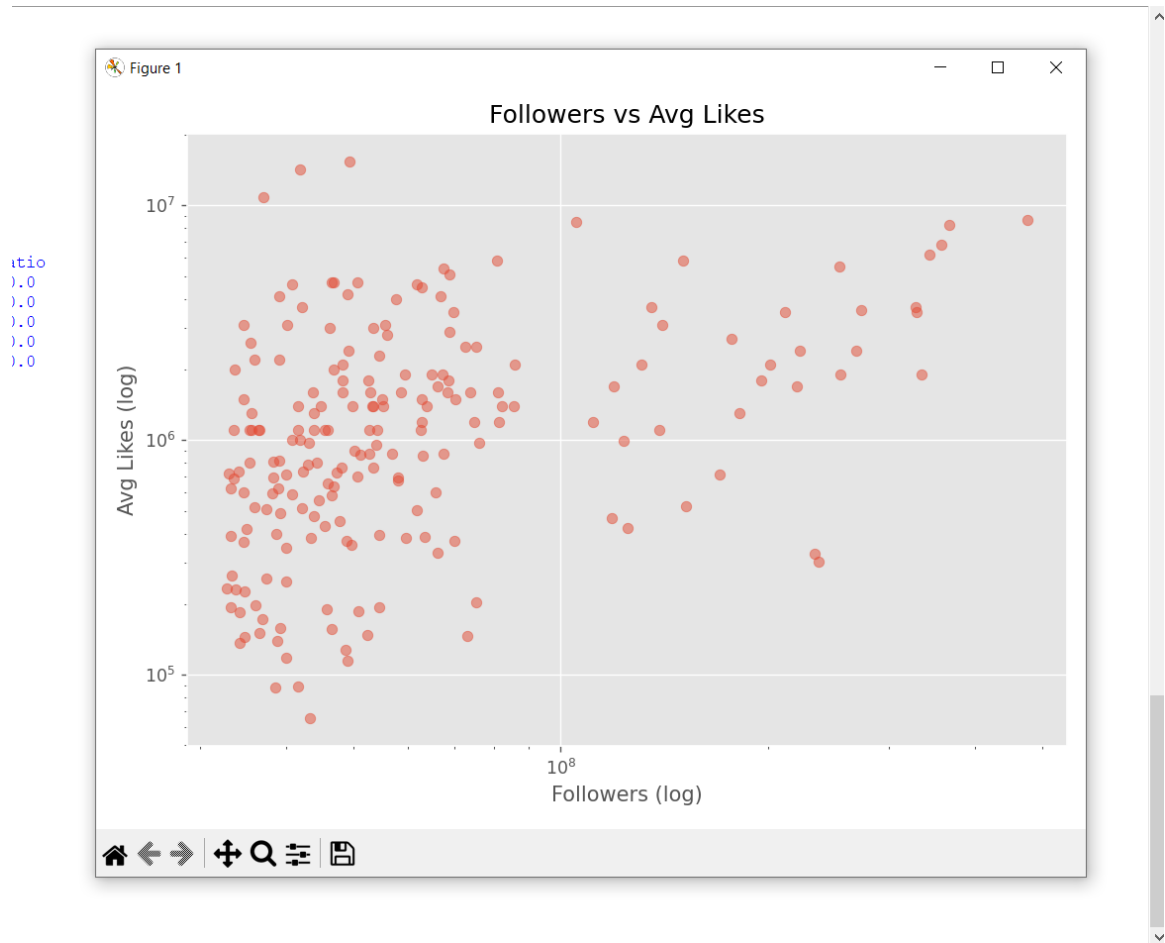# 60-Day Engagement Rate Distribution(%)

Average Likes Distribution

7. Relationships Between Metrics

Scatter plots visualize relationships such as Followers vs Engagement Rate, and Followers vs Average Likes.

Followers Vs 60-Day Engagement Rate
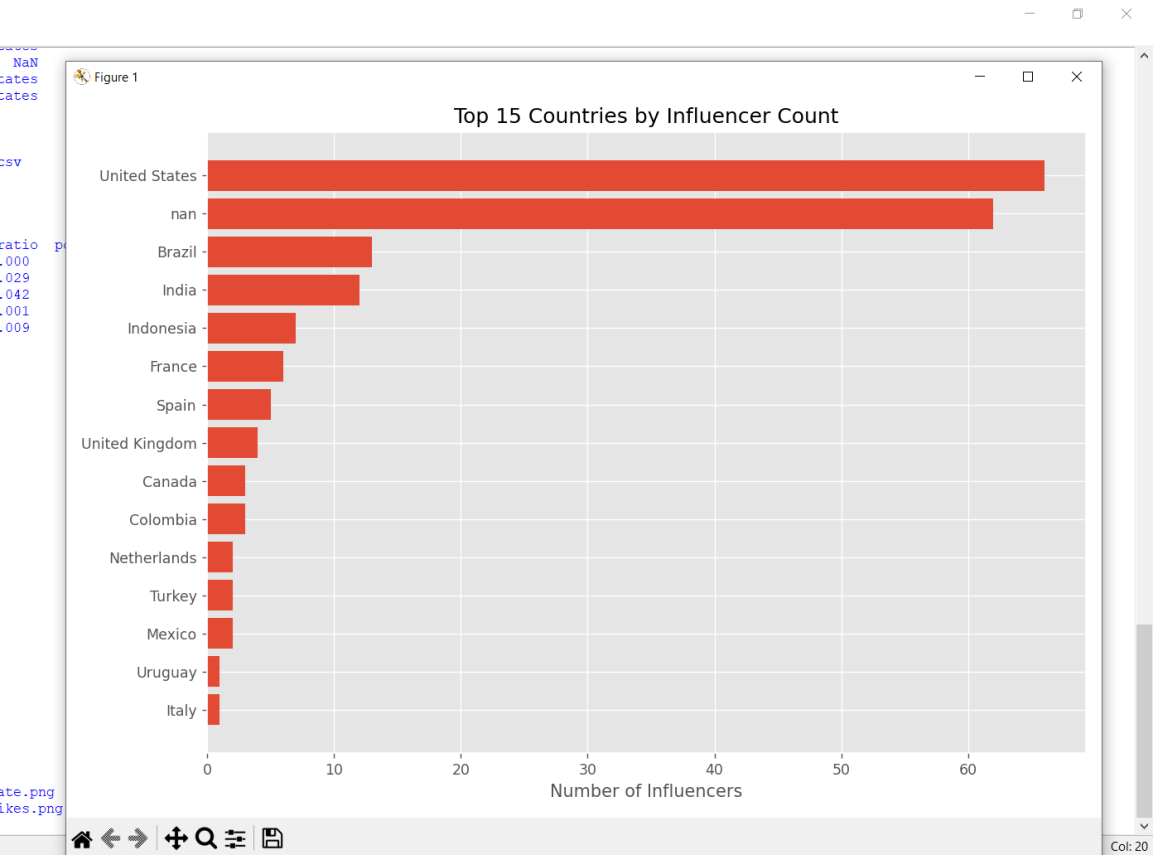
Followers Vs Avg Likes

8. Country-Level Insights

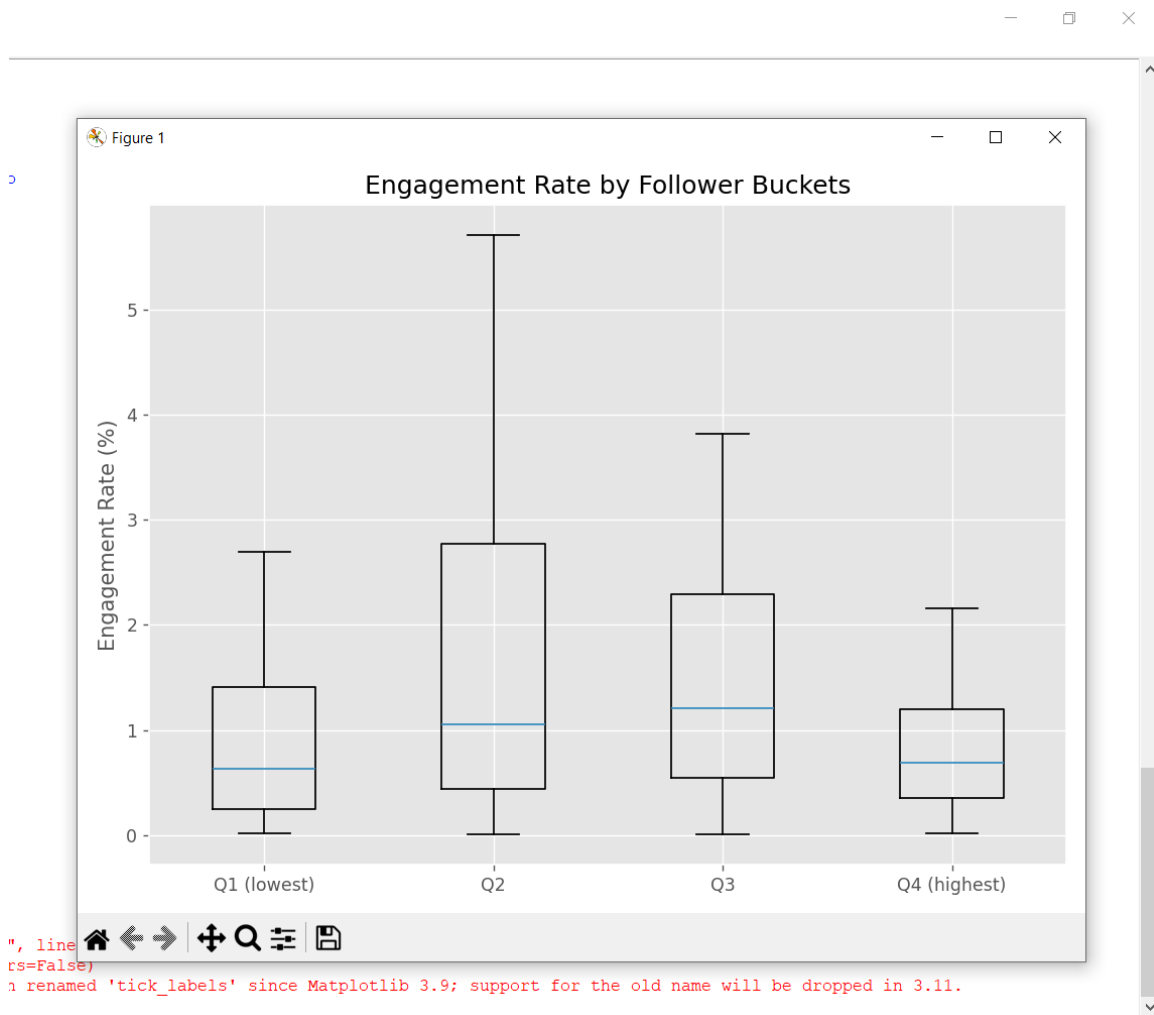A bar chart showing the top 15 countries with the highest number of influencers.

Top 15 Countries by Influencer Count

9. Engagement Rate by Follower Buckets

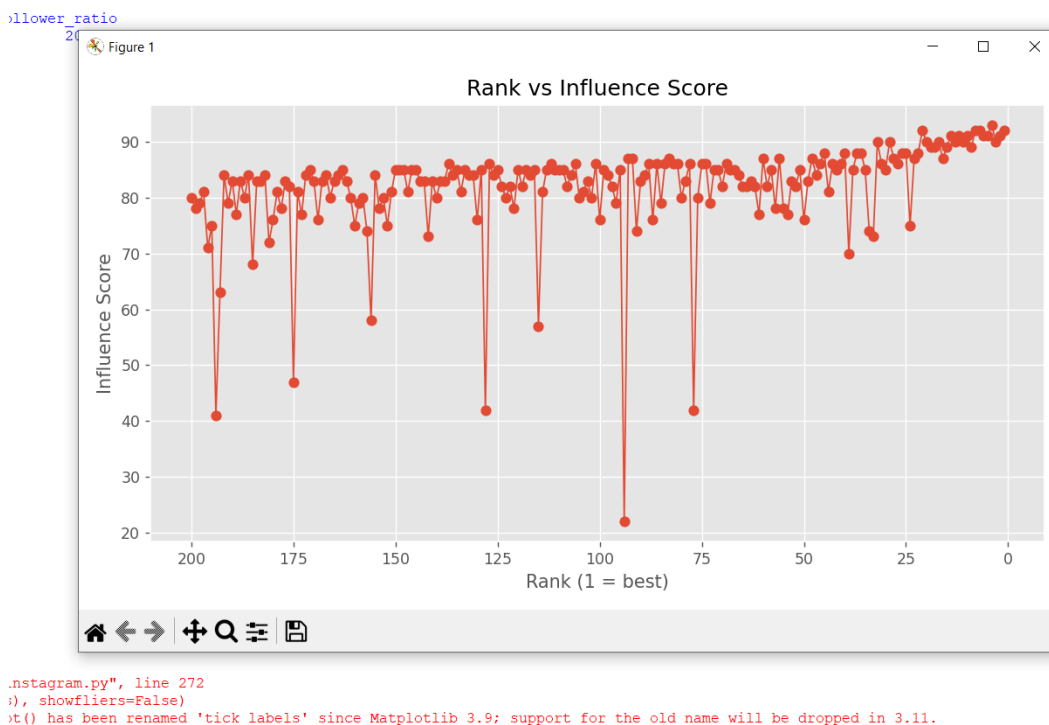A boxplot showing how engagement rate varies across different follower count quartiles.

Engagement Rate by Follower Buckets

10. Rank vs Influence Score

A line plot comparing influencer rank with their influence score.

Rank Vs Influence Score

11. Predictive Modeling

A Random Forest Regression model was built to predict influence scores based on followers, likes, and engagement metrics. Performance metrics such as RMSE, MAE, and R^2 were calculated.
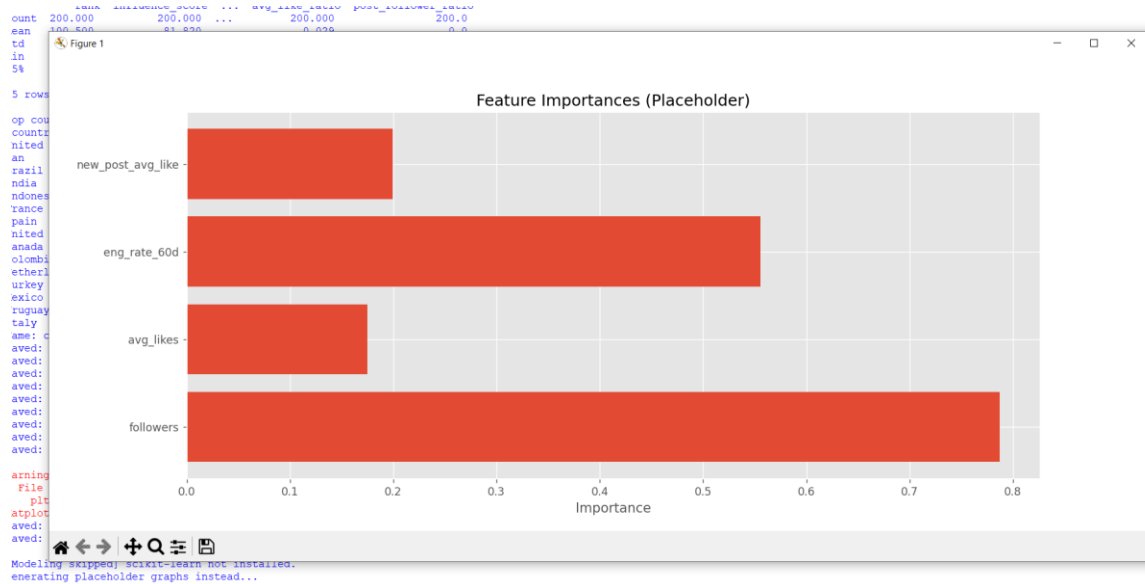
Prediction Vs Actual – Influence Score

Prediction vs Actual — Influence Score (Placeholder)

2

'tick_labe

Feature Importances

Feature Importances (Placeholder)
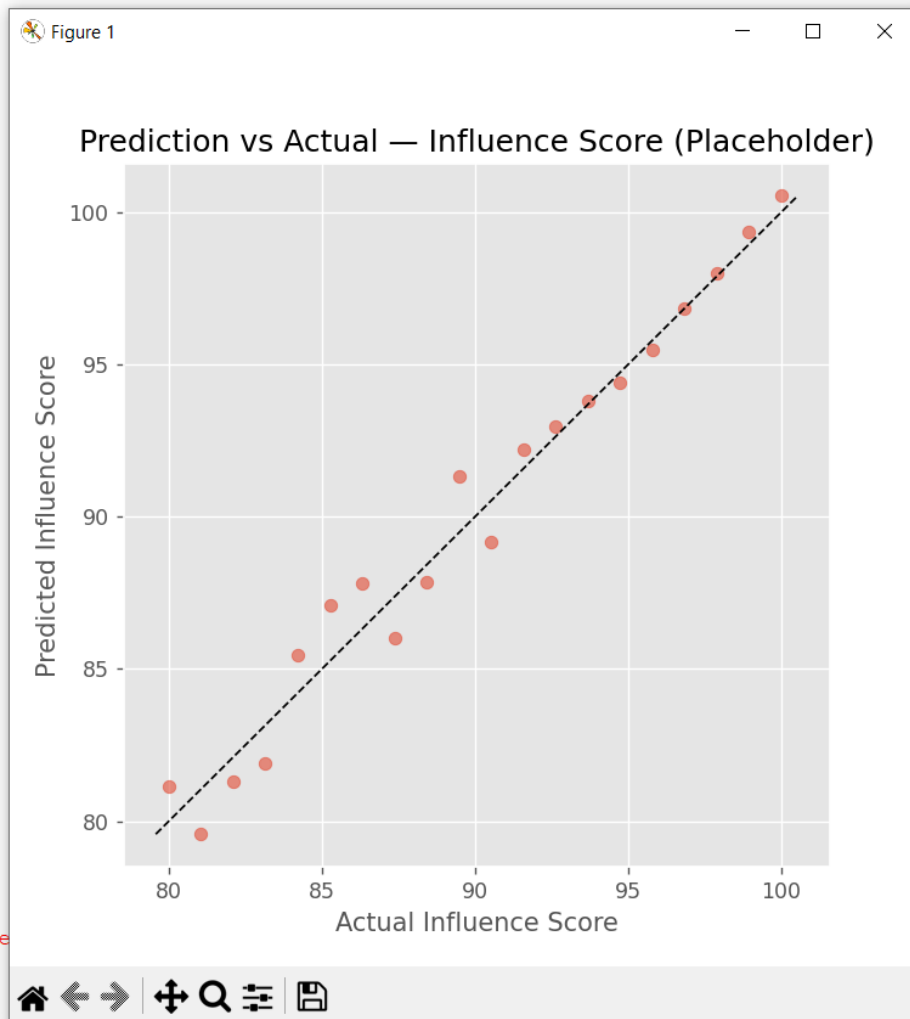
12. Conclusion

The analysis provided insights into the characteristics of top Instagram influencers, highlighted countries with the largest influencer presence, and explored relationships between engagement metrics. The predictive model demonstrated the feasibility of estimating influence scores based on engagement data.