Social Media Analytics for Business Assignment

User manual

CIS 7029-20276556

2023

Table of Contents

[Introduction 1](#_Toc148201695)

[Navigation Menu 2](#_Toc148201696)

[Dashboard 3](#_Toc148201697)

[Social Media Platforms 3](#_Toc148201698)

[**Reddit Performance** 3](#_Toc148201699)

[**YouTube Performance** 4](#_Toc148201700)

[**Reddit Like Prediction** 5](#_Toc148201701)

[**Filter of the Reddit Line Graph** 6](#_Toc148201702)

[**YouTube Views Prediction** 7](#_Toc148201703)

[**Filter of the YouTube Line Graph** 8](#_Toc148201704)

[**Polarity of the Reddit Comments** 9](#_Toc148201705)

[**Basic Network Properties** 10](#_Toc148201706)

[**Centrality Measurement:** 11](#_Toc148201707)

[Schedule 12](#_Toc148201708)

[Schedule Table 12](#_Toc148201709)

[Post Schedule Form 13](#_Toc148201710)

[Activity Diagrams 14](#_Toc148201711)

[Predictive likes 14](#_Toc148201712)

[Predictive Views 16](#_Toc148201713)

[Comment Sentiment Analysis 17](#_Toc148201714)

[Complete Network Analysis 19](#_Toc148201715)

[Schedule Posts 21](#_Toc148201716)

[Get Reddit Details (Likes, Comment) 22](#_Toc148201717)

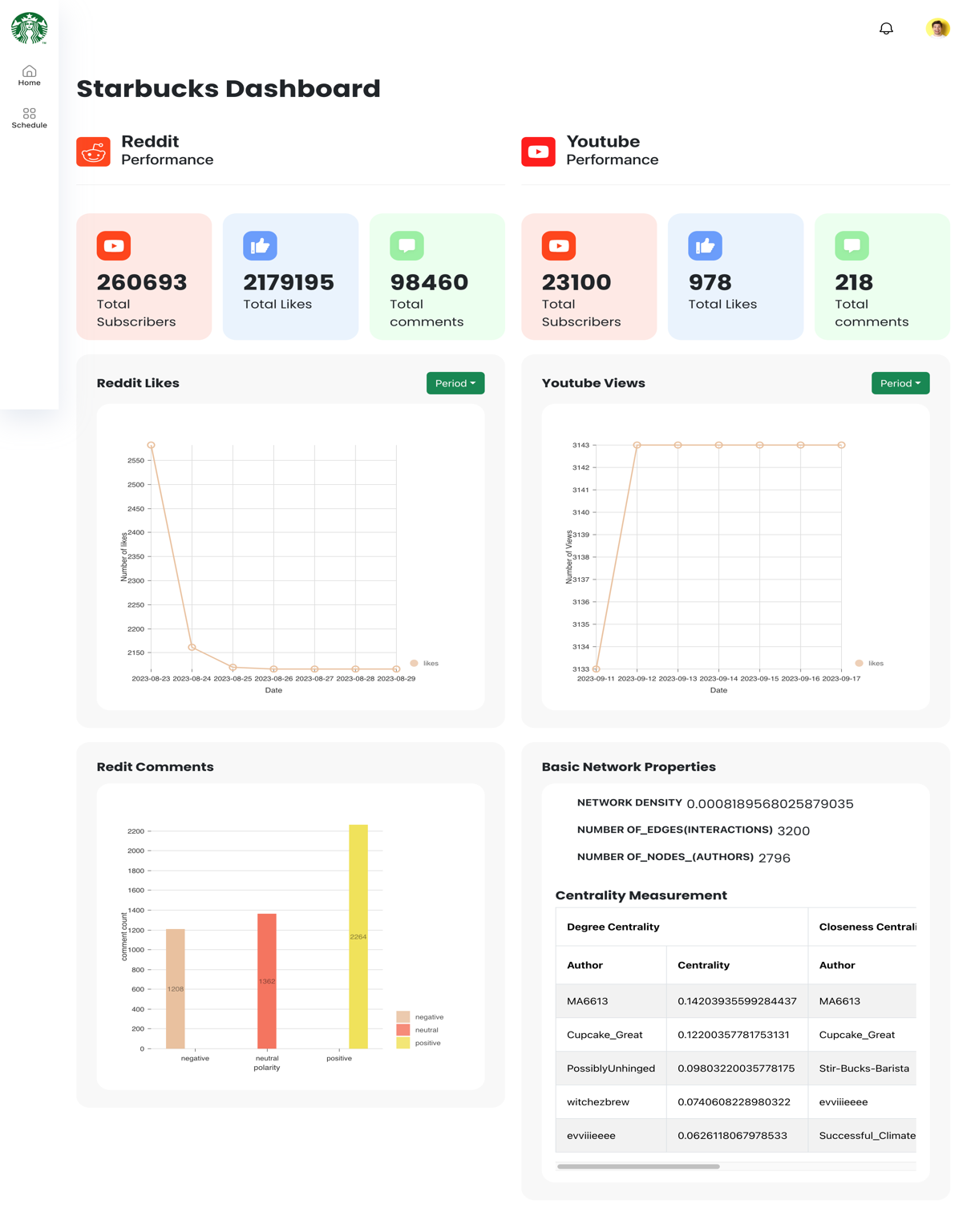
[List Schedule Posts 23](#_Toc148201718)

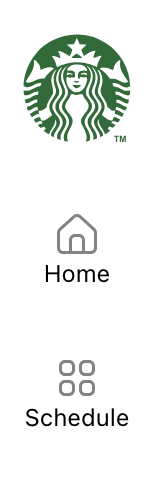
[Get YouTube Details 24](#_Toc148201719)

# **Introduction**

Welcome to the User Guide for the Starbucks Social Media Analytics Dashboard. This document describes in detail how to use this dashboard to analyse Starbucks' presence on two popular social media platforms, Reddit and YouTube, using data obtained via APIs.

From these line graphs, it is also able to predict how many likes each social media network will receive in the coming days.



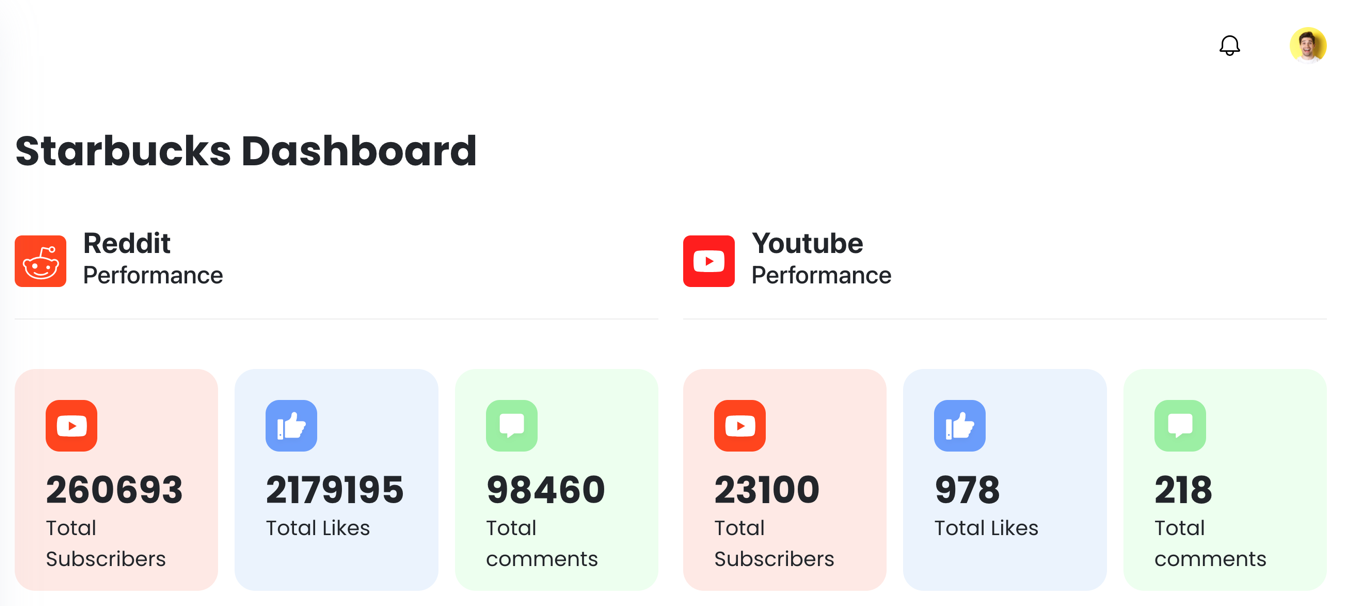


# **Navigation Menu**

1. The Starbucks logo can be shown on the top of the menu.
2. The 'Home' menu, which is the first option on the menu, allows users to access the Dashboard.
3. Regarding the last menu item on the navigation, "Schedule." The user can then proceed to the Schedule page to schedule a post for a later day on a social media platform.

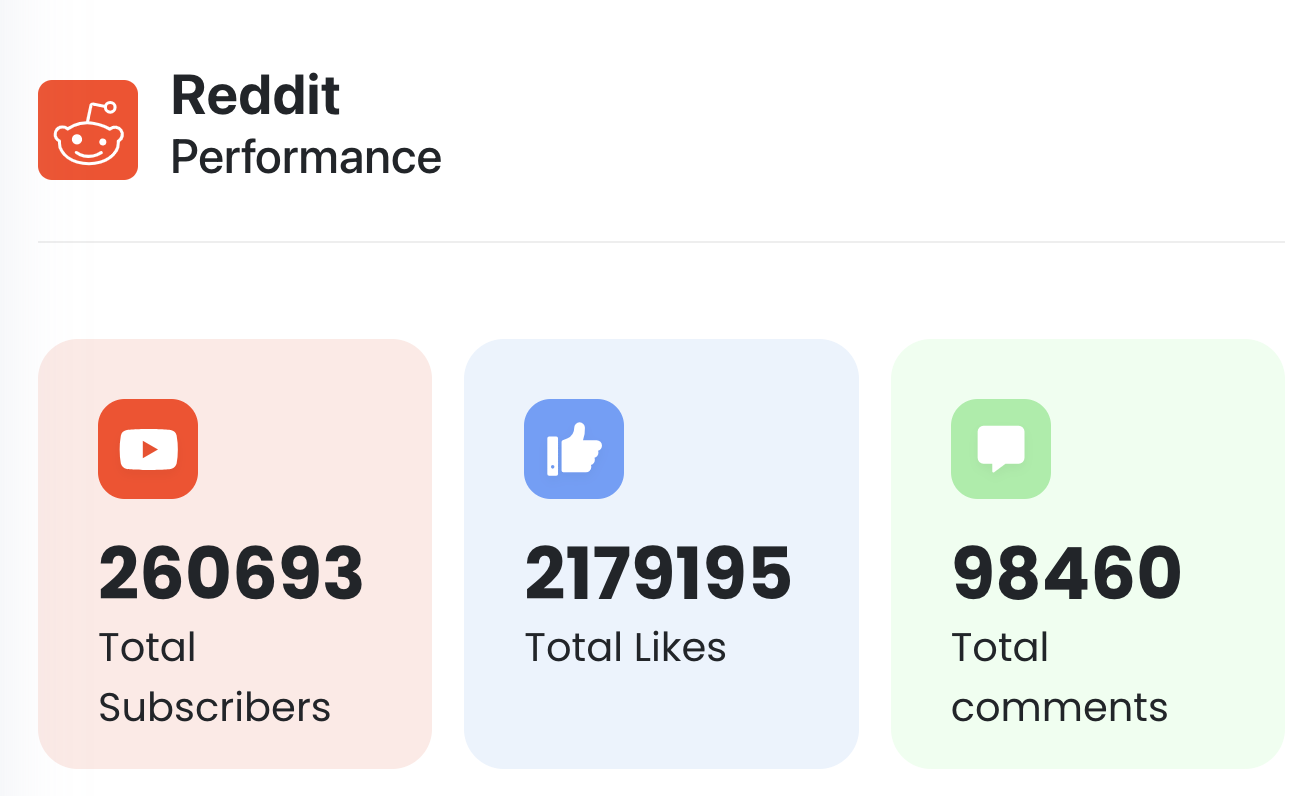
# **Dashboard**

## **Social Media Platforms**



1. The two widgets illustrate the performance of two separate social media platforms.
2. The Reddit performance display is located in the left widget.
3. The YouTube performance is visible in the right widget.

### **Reddit Performance**

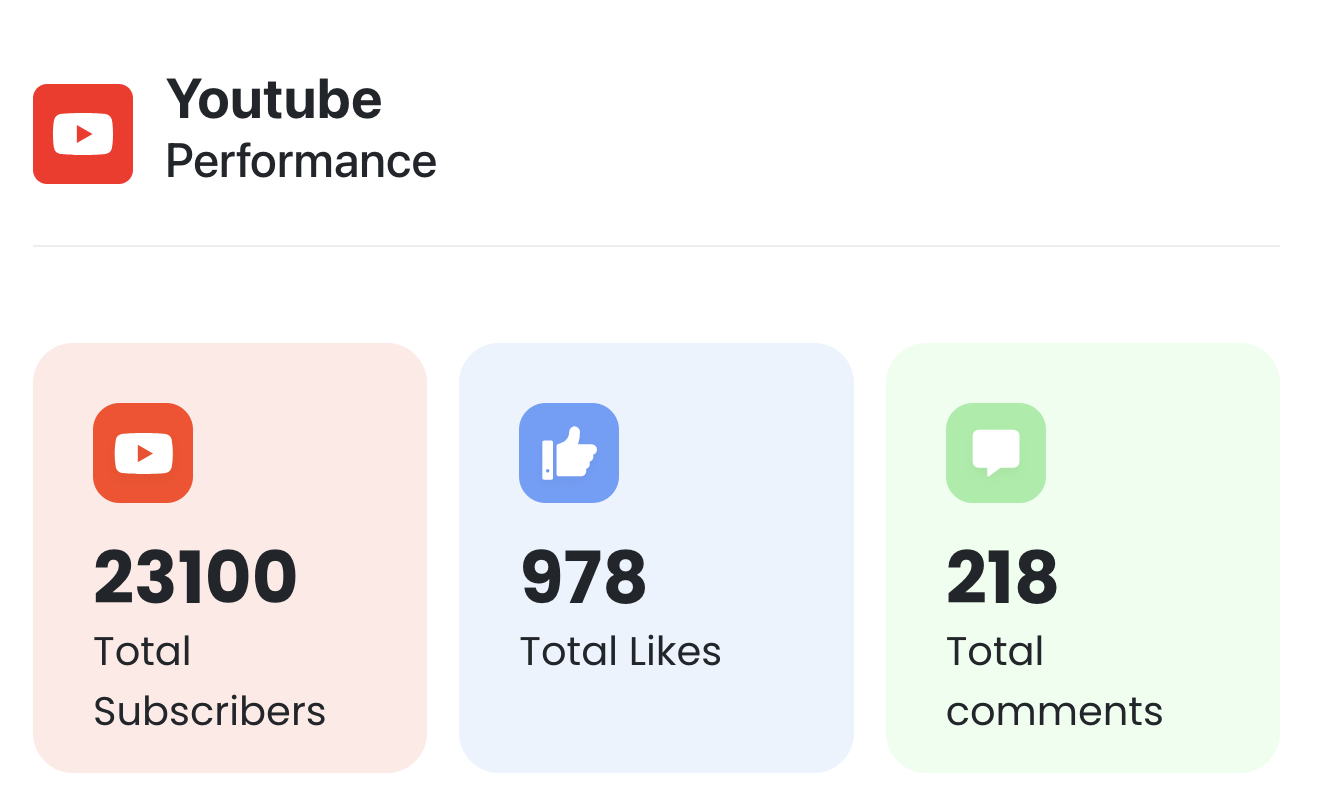


Below elements are display in the Reddit:

1. The complete count of Starbucks's subscribers.
2. Number of likes on a selection of 1000 posts.
3. Number of comments on 1000 randomly chosen posts.

The screenshot indicates a total of 2179195 Likes, 98460 Comments, and 260693 Subscribers.

### **YouTube Performance**

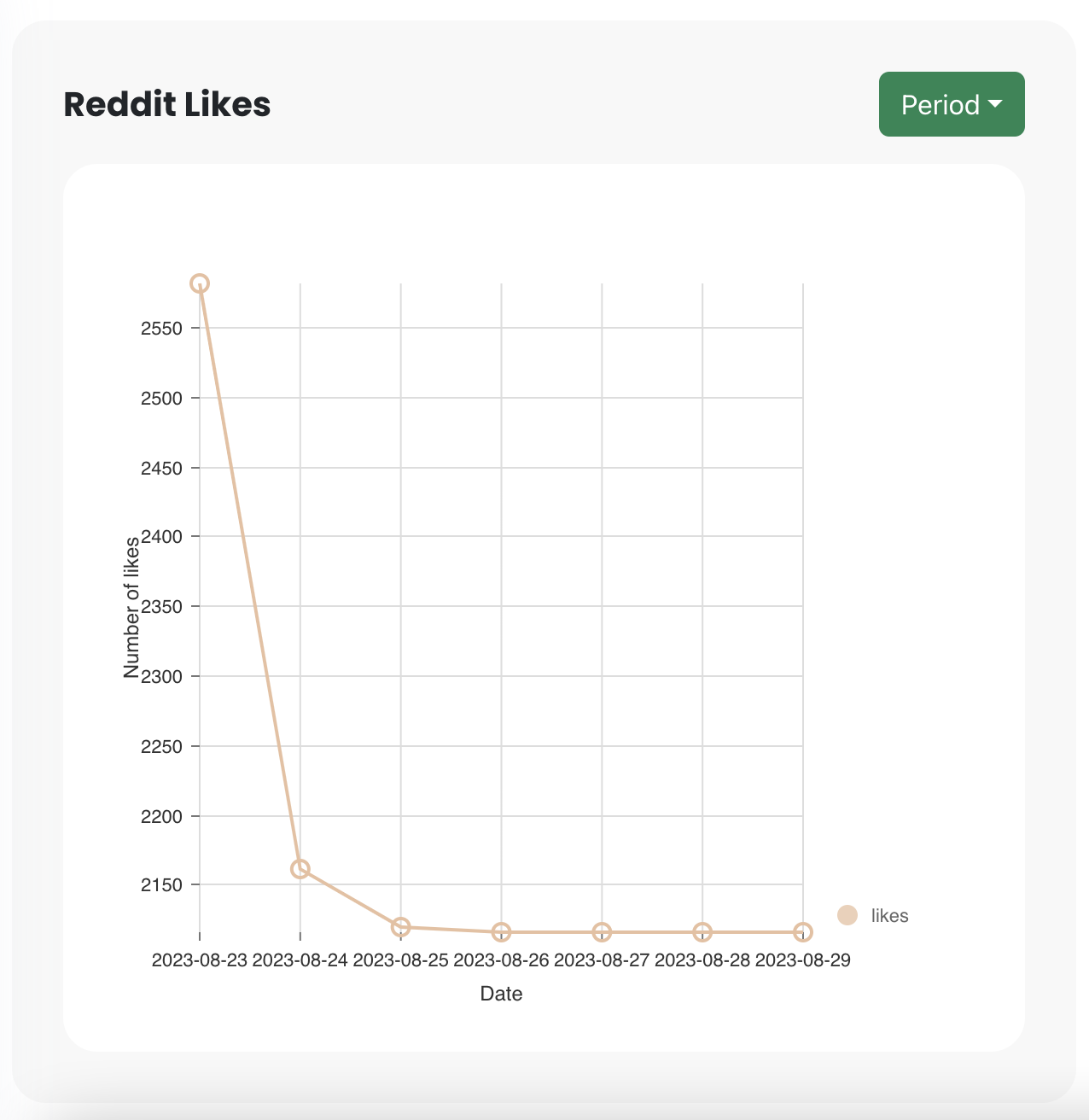


The following elements can be seen on YouTube:

1. The total number of people who are subscribed to Starbucks.
2. Likes on a random sample of 1000 posts.
3. Estimated from a sample of 1,000 posts.

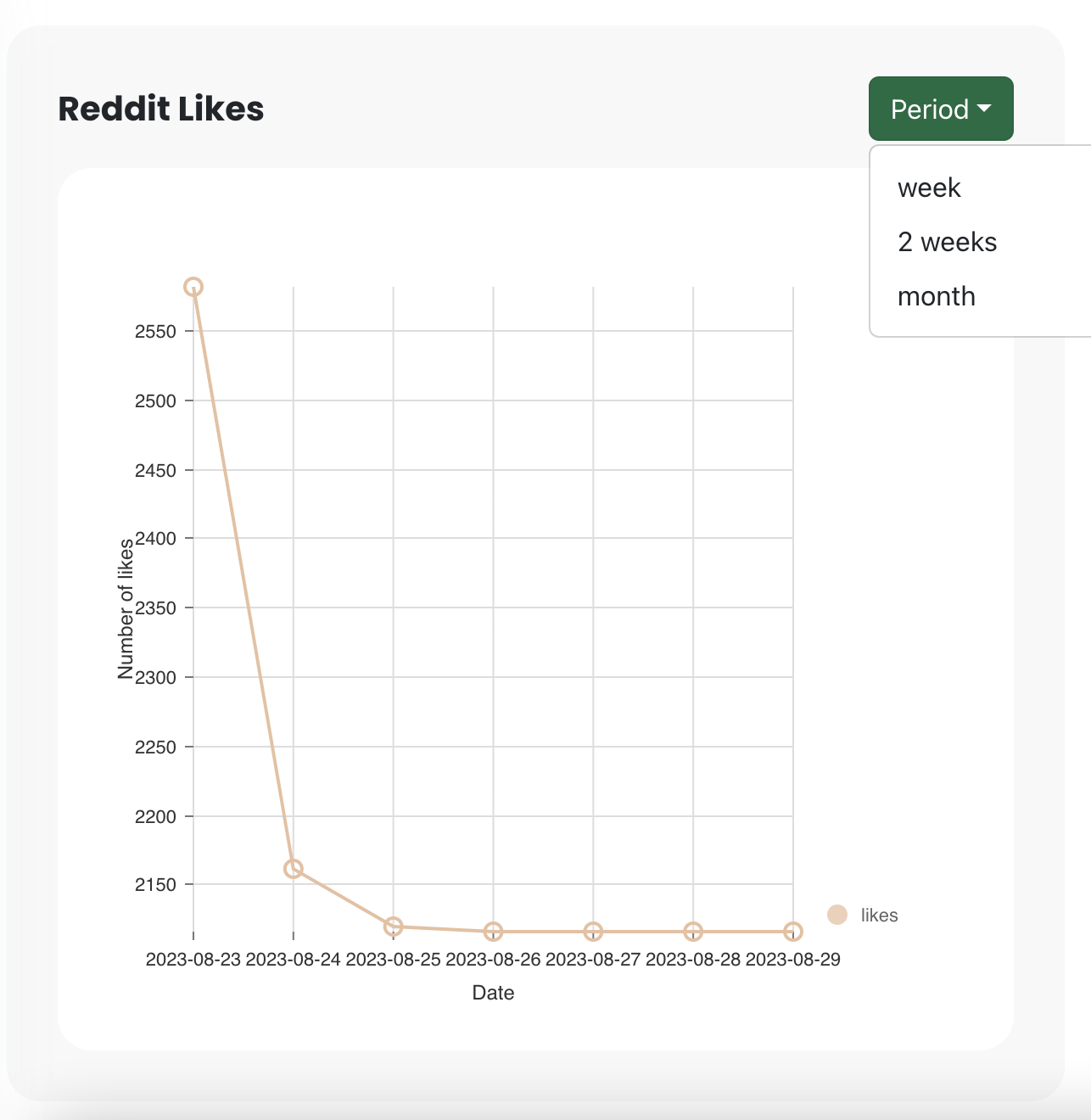
There are a total of 978 Likes, 218 Comments, and 23100 Subscribers, as seen in the screenshot.

### **Reddit Like Prediction**



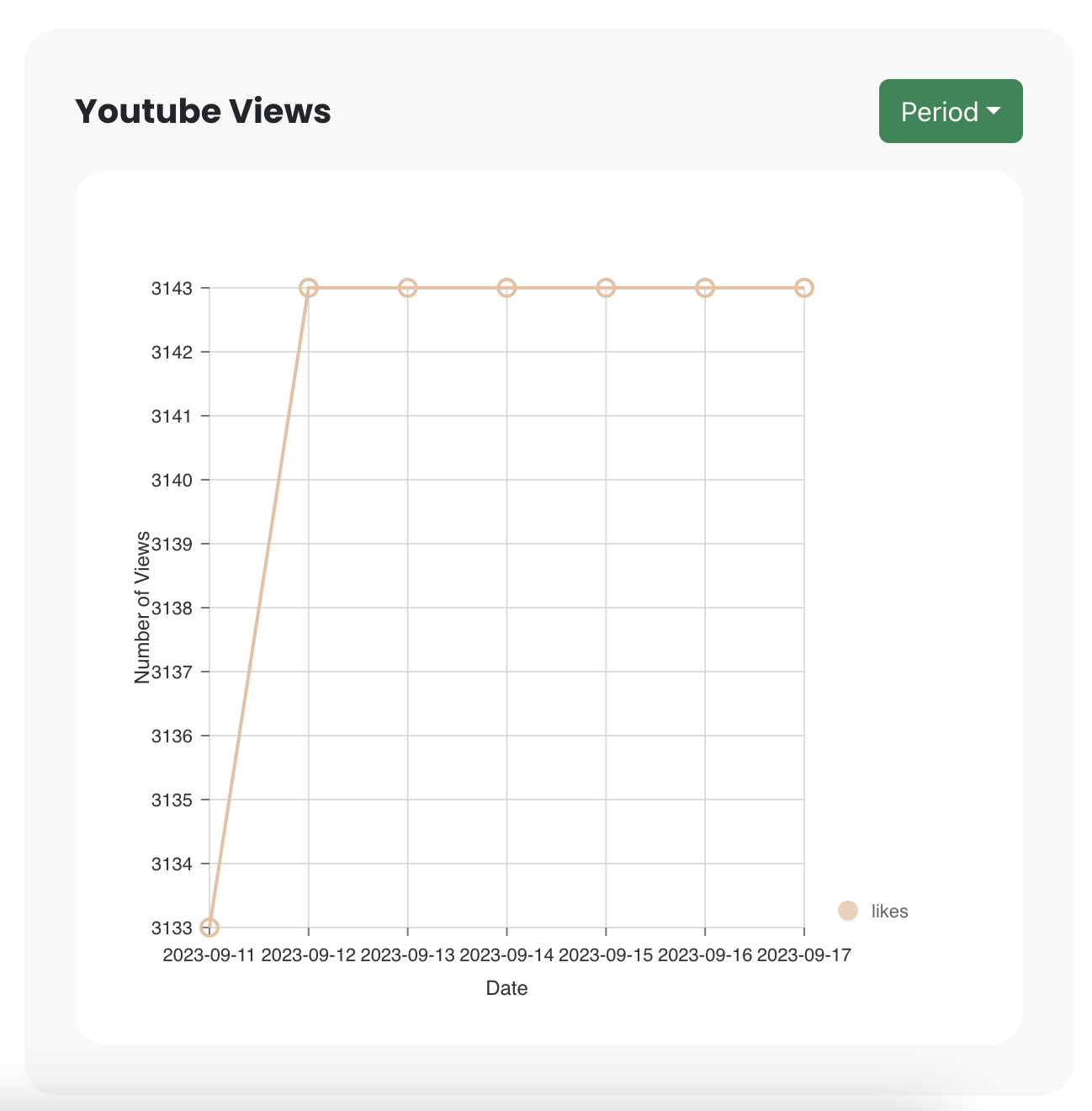
This line graph shows how many likes a post in the Starbucks subreddit on Reddit is likely to receive over the next few days. The forecast for the next week is displayed in the initially.

### **Filter of the Reddit Line Graph**



In the right top of the chart there is a filter with periods. There are 3 options in the filter and they are ‘week’, ‘2 weeks’ and ‘month’ In the initial stage it is in week basis. If user select 2 weeks from the filter the prediction will appear for 2 weeks and if user select month from the filter the prediction will count the like count for the next month.

### **YouTube Views Prediction**



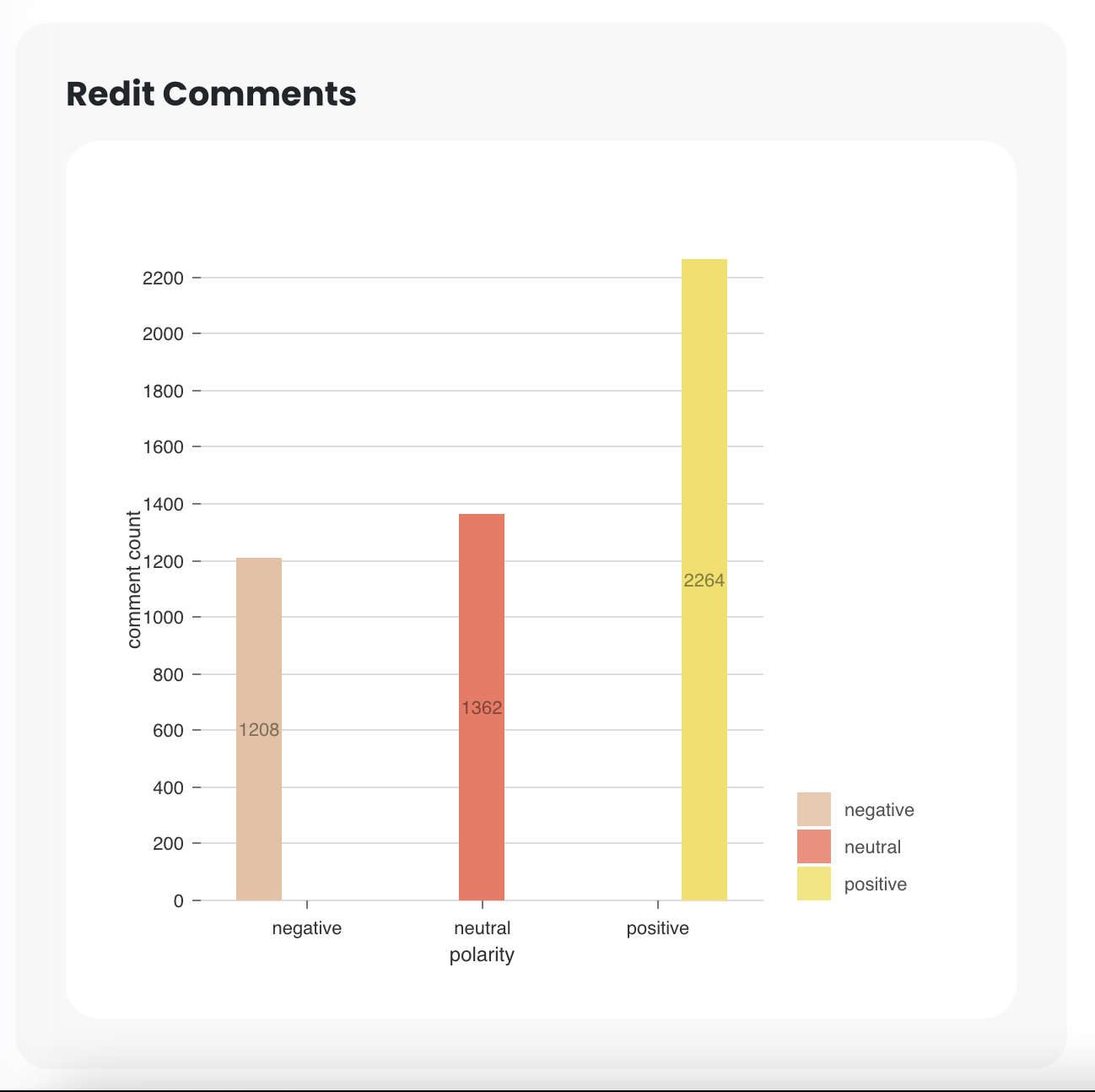
This line graph depicts, over the period of the next few days, the estimated number of views that a video posted to the Starbucks on YouTube will gain. The predictions for the upcoming week can be seen in the initialization graph.

### **Filter of the YouTube Line Graph**



A period filter that can be found in the top right corner of the chart. The filter allows you to choose between three different time intervals: a week, 2 weeks, or a month. When we launch out, we do it on a weekly basis. If the user selects month from the filter, the prediction will tally the like count for the following month. If the user selects two weeks from the filter, the prediction will appear for two weeks.

### **Polarity of the Reddit Comments**

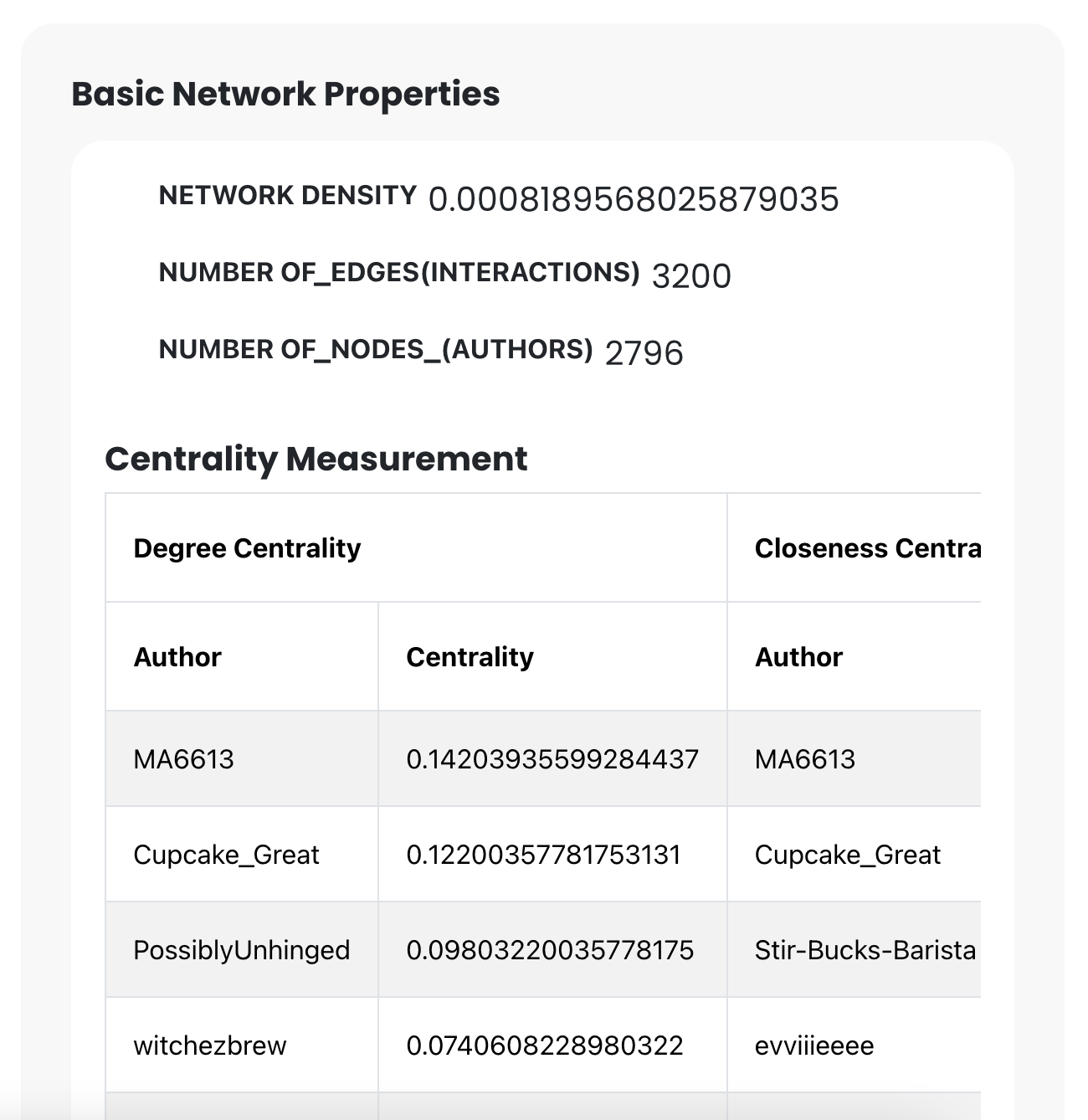


The bar chart depicts the polarity of sentiment in Reddit comments about Starbucks. The comments were divided into three sentiment categories: negative, neutral, and positive.

* Negative Comments (Beige Color): The beige bar corresponds to negative comments. There are 1208 comments classed as negative opinions about Starbucks.
* Neutral Comments (Orange Bars): The orange bars indicate the neutral comments. There are 1362 comments indicating a neutral attitude regarding Starbucks.
* Positive Comments (Yellow Bars): The yellow bars reflect comments with a favourable feeling. A total of 2264 comments have been categorized as good attitudes for Starbucks.

This visual representation provides insights into the distribution of sentiment in Reddit comments about Starbucks, giving for a rapid overview of the community's general sentiment dynamics.

### **Basic Network Properties**



1. Network Density (0.0008189568025879035): This is the fraction of network connections to total available connections. In simple terms, it displays the number of real interactions (edges) in relation to the maximum number of conceivable connections.
2. Number of Edges (Interactions) (3200): The total number of connections or interactions between network entities (nodes).
3. Number of Nodes (Authors): This is the total number of distinct entities (authors) in the network.

### **Centrality Measurement:**

Centrality is a measure of the relevance or prominence of a node (author) inside a network. To evaluate this, many centrality metrics are utilized.

1. **Centrality by Degree:**
   1. MA6613 (Author): This author has a degree centrality of 0.142, indicating a pretty high number of connections or interactions with other nodes.
   2. Cupcake\_Great (Author): This author's degree centrality is 0.122, indicating a high level of connectivity.
   3. PossiblyUnhinged (Author): This author's degree centrality is 0.098, indicating a moderate number of relationships.
   4. Stir-Bucks-Barista (Author): This author's degree centrality is 0.074, indicating a lesser level of relationships.
2. **Centrality of proximity:** 
   1. The authors' proximity centrality values are not provided in the offered information.

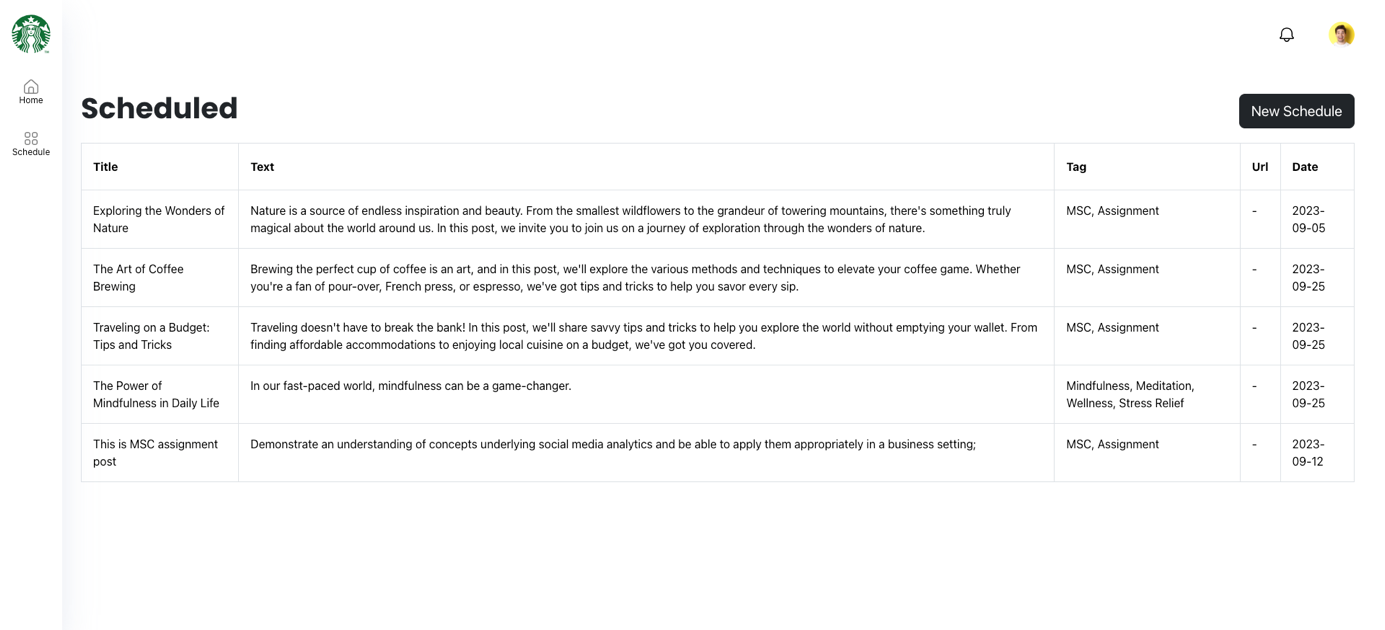
**Authors and Centrality Ratings:**

* MA6613: This author has a degree centrality of 0.142, indicating a pretty high level of network connections with other writers.
* Cupcake\_Great has a degree centrality of 0.122, indicating a high amount of connectivity.
* This author has a degree centrality of 0.098, indicating a moderate number of relationships.
* Stir-Bucks-Barista: This author has a degree centrality of 0.074, indicating that he or she has less relationships.
* evvilleeee: The author's centrality information is not supplied in the offered information.

From this user can found who is the best influencer that they can use for their promotion.

# **Schedule**

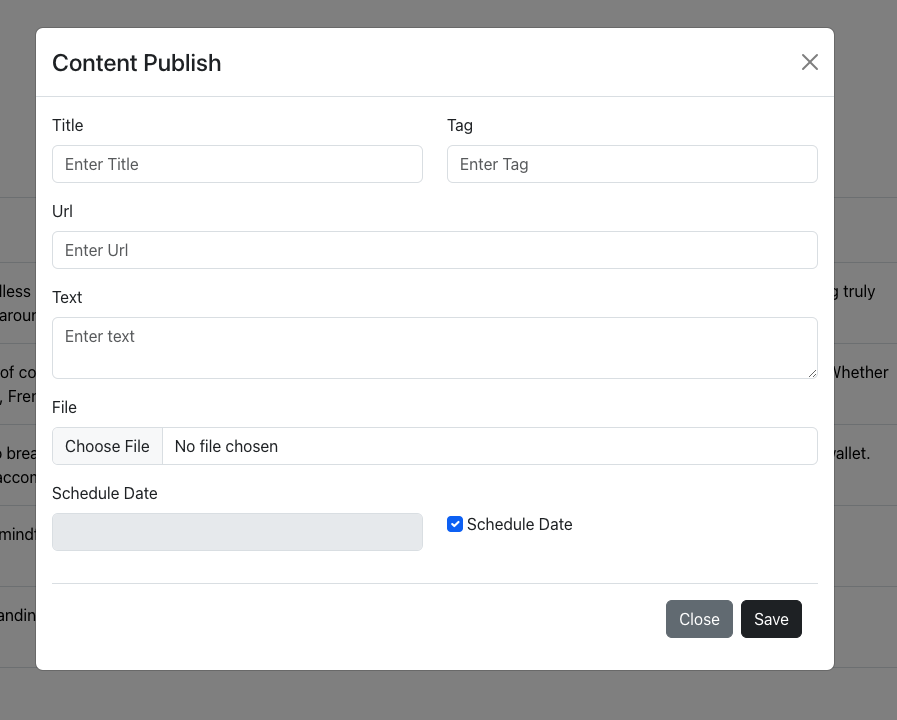
## **Schedule Table**



A "Schedule Table" is a well-structured table scheduling content for social media platforms such as Facebook and Reddit. The table includes important elements such as the post title, text, tags, url and the scheduled date for posting the post on the various social networking networks. Here's a more detailed explanation of each element as well as the overall functioning of the schedule table:

* Title: The post title is a brief and appealing headline for the content that will be posted on Facebook and Reddit. It acts as a hook to draw the audience in and urge them to read the piece.
* Text: The major body of the post that contains the intended message, information.
* Tags: Tags are keywords or phrases related to the post that help with categorization and discoverability. They can include themes, subjects, or pertinent keywords that describe the post's content. To increase visibility and reach, many tags can be added to a single post.
* URL: Can add multiple urls to the post and they are appear of the body of the post.
* Date: The scheduled date is the time and day that the post is set to be shared on Facebook and Reddit. This feature enables strategic planning and optimisation of post visibility while taking into account peak interaction hours on each platform.

## **Post Schedule Form**

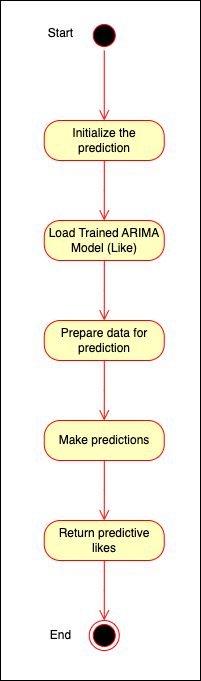


The schedule form is a user-friendly interface designed to facilitate the addition of new scheduled posts for sharing on Facebook and Reddit. It encompasses essential fields to collect relevant information for each post. Here's an expanded description of the fields and functionalities of the schedule form:

* Post Title: Users can input a captivating and descriptive title for the post. The title should succinctly represent the essence of the content being shared.
* URL: This field allows users to enter a web link that is associated with the post. The link can direct readers to additional information, a webpage, or any relevant online resource.
* Post Text: Users can input the main content or text of the post. This is where the primary message or information to be shared with the audience is composed.
* Tags: Users can input keywords or phrases relevant to the post's content, which act as tags. Tags aid in categorization and ease of search for the post. Multiple tags can be added to categorize the post effectively.
* File Attachment: This field allows users to upload and attach files such as documents, images, or any other relevant media associated with the post. These attachments can enhance the post's engagement and provide additional context.
* Scheduled Date: Users can select a specific date from the calendar for when they want the post to be shared on Facebook and Reddit. Additionally, there is an option to select a checkbox to add the post on the current day, providing flexibility and convenience.

# **Activity Diagrams**

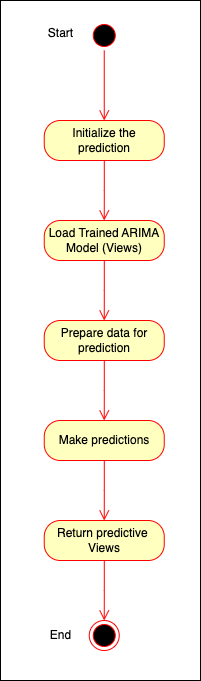
## **Predictive likes**

****

1. The process begins with the "Start" activity, initializing the predictive likes calculation.
2. The system loads the pre-trained ARIMA model in the "Load Trained ARIMA Model" activity.
3. Data preparation for prediction is performed in the "Prepare Data for Prediction" activity, which involves selecting the relevant data and features.
4. The "Make Predictions" activity uses the ARIMA model to make predictions based on the prepared data.
5. The predictions may require post-processing, so the "Post-process Predictions" activity handles and refines the results.
6. The calculated predictive likes are returned in the "Return Predictive Likes" activity.
7. Finally, the process is completed at the "End."

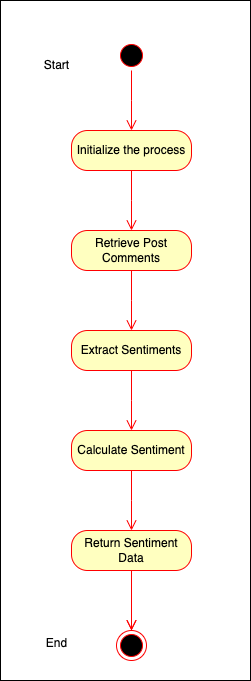
This activity diagram represents the steps involved in using a trained ARIMA model to generate predictive likes.

## **Predictive Views**



1. The process begins with the "Start" activity, initializing the predictive views calculation.
2. The system loads the pre-trained ARIMA model in the "Load Trained ARIMA Model" activity.
3. Data preparation for prediction is performed in the "Prepare Data for Prediction" activity, which involves selecting the relevant data and features.
4. The "Make Predictions" activity uses the ARIMA model to make predictions based on the prepared data.
5. The predictions may require post-processing, so the "Post-process Predictions" activity handles and refines the results.
6. The calculated predictive views are returned in the "Return Predictive Views" activity.
7. Finally, the process is completed at the "End."

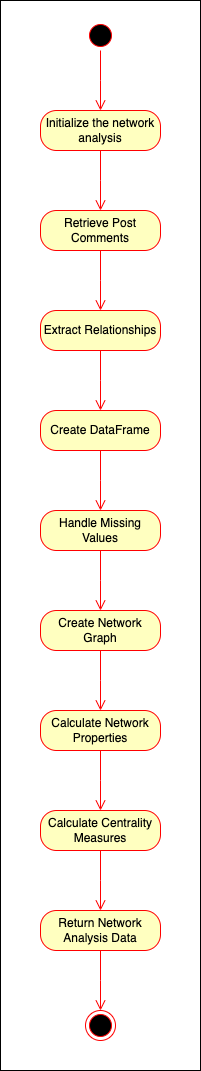
## **Comment Sentiment Analysis**

****

1. The process still starts with the "Start" activity, initializing the sentiment analysis.
2. The system retrieves Reddit post comments from the database in the "Retrieve Post Comments" activity.
3. Sentiment analysis is performed in the "Sentiment Analysis" activity, which includes both the analysis of sentiment and storing the sentiment value in the database.
4. Sentiment counts are calculated in the "Calculate Sentiment" activity.
5. The final sentiment data is formatted and used to create a response in the "Return Sentiment Data" activity.
6. The process is completed at the "End."

This activity diagram represents the major steps involved in the get\_sentiment\_data function, from data retrieval to sentiment analysis and response generation.

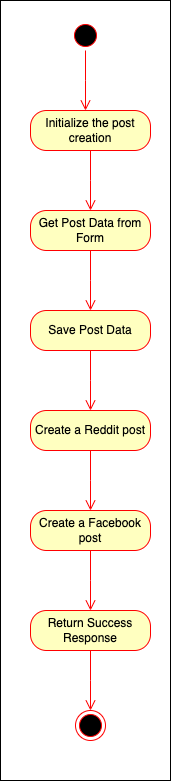
## **Complete Network Analysis**

****

1. The process starts with the "Start" activity, initializing the network analysis.
2. The system retrieves Reddit post comments from the database in the "Retrieve Post Comments" activity.
3. Relationships between authors are extracted and structured in the "Extract Relationships" activity.
4. The data structure is converted into a pandas Data Frame in the "Create DataFrame" activity.
5. Rows with missing data in the Data Frame are removed in the "Handle Missing Values" activity.
6. An empty network graph is created in the "Create Network Graph" activity, and nodes and edges are added based on the Data Frame.
7. Various network properties (e.g., number of nodes, edges, and density) are calculated in the "Calculate Network Properties" activity.
8. Centrality measures (degree centrality, betweenness centrality, and closeness centrality) are calculated in the "Calculate Centrality Measures" activity.
9. The results and network analysis data are packaged and returned in the "Return Network Analysis Data" activity.
10. The process is completed at the "End."

This activity diagram illustrates the sequence of actions and activities involved in the get\_network\_analysis\_details function, from data retrieval to network analysis and return results.

## **Schedule Posts**



1. The process starts with the "Start" activity, initializing the post creation.
2. Post data is retrieved from the incoming request in the "Get Post Data from Form" activity.
3. The post data is saved to the database in the "Save Post Data" activity.
4. A decision point checks whether the post is a one-time post.
5. If it is a one-time post, both Reddit and Facebook posts are created in the "Create One-Time Reddit Post" activity.
6. A success response is packaged and returned in the "Return Success Response" activity.
7. The process is completed at the "End."

This activity diagram represents the sequence of actions and activities involved in the create\_schedule\_post function, from retrieving data to saving it to the database and creating social media posts.

## **Get Reddit Details (Likes, Comment)**

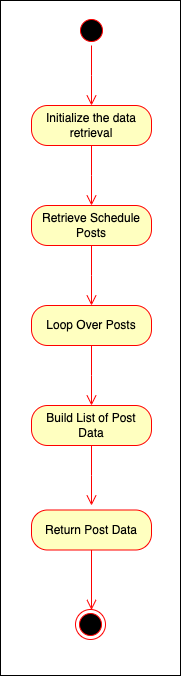
**A diagram of a diagram

Description automatically generated**

1. The process starts with the "Start" activity, initializing the data retrieval.
2. The system retrieves Reddit posts from the database in the "Retrieve Reddit Posts" activity.
3. The number of subscribers is obtained from the database in the "Retrieve Subscribers Count" activity.
4. The total like count and comment count are calculated in the "Calculate Like and Comment" activity.
5. A data structure is created with like count, comment count, and subscriber count in the "Create Data Structure" activity.
6. The data structure is packaged and returned in the "Return Reddit Data" activity.
7. The process is completed at the "End."

This activity diagram represents the sequence of actions and activities involved in the get\_reddit\_data function, from data retrieval to data calculation and result preparing.

## **List Schedule Posts**

****

1. The process starts with the "Start" activity, initializing the data retrieval.
2. Scheduled posts are retrieved from the database in the "Retrieve Schedule Posts" activity.
3. A data structure is created to store post information in the "Create Data Structure" activity.
4. The system iterates over the retrieved scheduled posts in the "Loop Over Posts" activity, extracting relevant information.
5. A list of post data is built by appending extracted data for each scheduled post in the "Build List of Post Data" activity.
6. The list of post data is returned in the "Return Post Data" activity.
7. The process is completed at the "End."

This activity diagram represents the sequence of actions and activities involved in the get\_all\_schedule\_posts function, from retrieving scheduled posts to building and returning a list of post data.

## **Get YouTube Details**

**A diagram of a diagram

Description automatically generated**

1. The process starts with the "Start" activity, initializing the data retrieval.
2. YouTube posts are retrieved from the database in the "Retrieve YouTube Posts" activity.
3. The number of subscribers is obtained from the database in the "Retrieve Subscribers Count" activity.
4. The total like count, comment count, and view count are calculated in the "Calculate Like, Comment, and View Counts" activity.
5. A data structure is created with like count, comment count, view count, and subscriber count in the "Create Data Structure" activity.
6. The data structure is packaged and returned in the "Return YouTube Data" activity.
7. The process is completed at the "End."

This activity diagram represents the sequence of actions and activities involved in the get\_youtube\_data function, from retrieving YouTube posts to calculating statistics and returning the data structure.