

## Actively Monitoring Social Media Metrics

### Summary

The intent of this exercise is to develop an approach that would allow for active monitoring of social media metrics such as volume and sentiment to identify potential anomalies. By spotting these potential shifts, brands may then investigate the drivers of these changes and assess whether a response is warranted. To illustrate this, the Excel file contains 3 years of social media metrics for Starbucks. The following metrics have been extracted for each day:

- Total number of Twitter posts
- Number of positive/neutral/negative posts
- Number of authors contributing posts
- Average number of posts per author

To illustrate how active monitoring could be implemented, our approach will be to use the previous 31 days of social media activity to establish a baseline. We will then look for activity that appears unlikely under the baseline. This will be done for each of the metrics in which we are interested, allowing us to assess combinations such as higher than usual volume and sentiment, or higher than usual volume with lower than expected sentiment.

### Identifying Deviations from Expected Volume

Beginning with the Volume tab, in cell C34, calculate the average number of Twitter posts from the previous 31 days: =average(B4:B33)

In cell D34, calculate the standard deviation of the number of posts from the previous 31 days: =stdev(B4:B33)

In cell E34, we will calculate the Z-score associated with the mean and standard deviation calculated in C34 and D34, respectively. The Z-score will allow us to determine how likely we are to observe a particular amount of volume given a mean and standard deviation:

= (B34 - C34) / D34

Assuming that our data is normally distributed, for a given mean and volume, 95% of observations are expected to fall within 2 standard deviations and 99.7% of observations are expected to fall within 3 standard deviations. These correspond to  $|Z| \leq 2$  and  $|Z| \leq 3$ . If  $|Z| > 2$  or  $|Z| > 3$ , it can be interpreted as the observation being unlikely based on the mean and standard deviation from the previous 31 days. For our purposes, we will use a threshold of  $|Z| > 3$  to be more conservative. In cell F34, use a nested IF statement to display whether the Z-score associated with the day is  $|Z| \leq 3$ ,  $Z < -3$  or  $Z > 3$ .

=IF(E34<-3,"Low",IF(E34>3,"High",""))

This provides an indication based on the previous 31 days of whether or not social media volume has exhibited a shift.

## Identifying Deviations from Expected Sentiment

The Sentiment-Data worksheet provides the number of positive, neutral and negative posts each day. Begin by constructing a composite score where each positive comment receives a value of +1, each neutral comment receives a score of 0, and each negative comment receives a score of -1. In cell E3, calculate the average score as:

= $(B3-D3)/SUM(B3:D3)$

In cell F3, calculate the ratio of positive-to-negative posts as:

= $B3/D3$

Copy the formulas in E3 and F3 down the columns. These two metrics provide alternative measures of sentiment.

Copy the values of the composite sentiment into column B of the “Sentiment composite” tab.

Copy the values of the positive-to-negative ratio into column B of the “Sentiment ratio” tab.

For each of these tabs, follow the same steps as taken on the Volume worksheet to identify those dates on which the sentiment metric appears to be inconsistent with the previous 31 days.

## Examining the Contributor Base

The tabs “Number of Authors” and “Post per Author” contain data on the number of unique posters and the average number of posts per author for each day. Changes in social media activity, including the volume, sentiment and content of posts, may arise from current contributors contributing more/less, or from new contributors joining the conversation. These two measures provide some insight into the nature of the contributor base.

Follow the instructions used for the Volume worksheet to identify aberrations in the number of authors and posts per author.

## Summarize Your Findings

On the “Summary” tab, you will find a column corresponding to each of the measures we have analyzed. From each worksheet, copy and paste the values (not the formulas) from column F corresponding to the date range (F34:F1098) onto the “Summary” tab in the appropriate column. This will allow you to more quickly assess the dates on which multiple metrics exhibit a deviation from the values we would have expected based on the previous 31 days.