

Twitter and Facebook Scraping of Moringa School Data

BUSINESS UNDERSTANDING

Social media are interactive technologies that allow the creation or sharing/exchange of information, ideas, career interests, and other forms of expression via virtual communities and networks. While challenges to the definition of social media arise due to the broad variety of stand-alone and built-in social-media services currently available, there are some common features

Moringa School ,a higher education institution,tasked us to analyse their social media pages such as twitter and facebook for them to better understand their clients and identify patterns within data that will help them increase engagement.

MAIN OBJECTIVE

1.0 QUESTIONS MARKETING TEAM WANTS ANSWERED FROM DATA DRIVEN INSIGHTS:

1. What is the best time to post content on the Facebook platform?
2. Which posts are more receptive by the audience?
3. Which tags are used by the higher ranked bootcamps and institutions to drive engagement?
4. Who are Moringa's competitors and what are their strategies?

OTHER OBJECTIVES

1.2 DS TEAM: OBJECTIVES

1. To create a timeline graph of year and gained number of likes in Moringa to see which year had more traffic
2. Help the marketing team with insights such as best time to post and social tags that they can use to drive engagement.
3. To check based on the posts and tweets , event posts or general posts , which one gains more likes and engagement
4. To identify the polarity of the posts made by the marketing team

1.1 ANALYSIS PROCESS

1. Scrape Data from Twitter and Facebook.
2. Do sentiment analysis by categorizing the polarity in various posts,comments was a deficiency to extract.
3. Cluster the various categories that bring engagement for the different social media platforms.
4. Do data analysis on the two social media platforms.
5. Give recommendations

Twitter and Facebook Scraping of Moringa School Data

2.0 TECHNOLOGIES USED

- For loop through the facebook page and extract data from page 2014-2021
- Twitter
- Beautiful Soup
- Tweepy
- Selenium
- Authorization credentials for twitter
- Requests and tokens from Graph API
- Octoparse
- Instant data scraper

Expected outcome for data retrieval

Script to scrape Moringas Facebook Page from 2014-2021

Script to scrape Moringas Twitter Page from 2014-2021

3.0 MARKETING TEAM: MARKETING STRATEGY ADVISORY

Key Metrics Moringa Marketing deals with :

1. Ensure regular engagement on our platforms.
2. Pre-scheduled evergreen content
3. More insights on our audiences, online presence, strategy
4. Have a (long-term) structure to guide content ideas and execution

*** This excludes paid content, and one-off communication

BASED ON THE ANALYSIS

FACEBOOK CONCLUSIONS AND RECOMMENDATIONS

- Over the years Moringa has been posting on **Facebook at 7am and 1 pm** most of the time, **Studies show to get more traffic on posts post at around 9 am and 1pm**
- Best days to post on facebook is Tuesday,Wednesday and Friday
- Saturday is not the best day to post on Facebook
- 2017 was the year in which Moringa had the most engagement. The best months include **January,February,March,November and July**
- Across all years the **best months that gained engagement were March ,November,June**

Twitter and Facebook Scraping of Moringa School Data

- Across the years , 7.30 am, 9am and 2 pm had more engagement
- Something interesting: the posts that were given a polarity of positive have more likes vs the ones given a negative polarity
- The **cluster representing workshops[engagement]** has the most likes in general.
- The text with most engagement was the AIESEC Mentorship day that happened on August , it had 876 likes
Link to post: <https://facebook.com/moringaschool/posts/1945559355579577>

experience moringa school many people say learning code hard hear moringa pop head programming simplified learn code adopt ability self learn among soft skill like communicate effectively build courage teamwork among others alexona share experience read httpbitlycfdwrw moringaschoolwordpresscom experience moringa school , had 716 likes

[Link: https://facebook.com/moringaschool/posts/959265984208924](https://facebook.com/moringaschool/posts/959265984208924)

nairobi tech week march pm application day ticket sub saharan africa largest tech event open tag friend increase chance selected httpsgooglformscnerxmstzxjylzwf application free ticket close march, had 700 likes

[Link: https://facebook.com/moringaschool/posts/956867034448819](https://facebook.com/moringaschool/posts/956867034448819)

TWITTER CONCLUSIONS AND RECOMMENDATIONS

- Most tweets by the Moringa School account are at **Mid-Morning 8:01am - 10:00am and Early-Morning 6:01am - 8:00am**
- **Early-Afternoon 12:01pm - 2:00pm and Late-Morning 10:01am - 12:00am** also have a higher percentage of posts
- **Evening 6:01pm - 9:00pm, Late-Evening 9:01pm - 11:59pm and Late-Night 0:00am - 3:00am** have the **least number of posts**.
- **Tweets sent Early-Morning 6:01am - 8:00am and Mid-Morning 8:01am - 10:00am** have the **highest total number of retweets** compared to other periods of the day.
- **Tweets sent in the Evening 6:01pm - 9:00pm, Late-Evening 9:01pm - 11:59pm and Late-Night 0:00am - 3:00am** have the **least total number of retweets** compared to other periods of the day.

Twitter and Facebook Scraping of Moringa School Data

- **Tweets sent Mid-Morning 8:01am - 10:00am and Early-Morning 6:01am - 8:00am** also have the **highest total number of likes** while tweets sent in the **Evening 6:01pm - 9:00pm, Late-Evening 9:01pm - 11:59pm and Late-Night 0:00am - 3:00am** have the **least number of total likes**.
- **Tweets sent Early-Morning and Late-Morning have a higher number of retweets and likes** compared to other time periods of the day, Early-Afternoon and Late-Morning have average number of likes and retweets while tweets sent Late-Afternoon, Late-Evening, Mid-Afternoon and Towards-Morning have low number of likes and retweets. Late-Night and Evening have the least.

NB: Comparing the ratio of posts to the number of retweets, Early-Morning and Late-Morning have higher ratios. It's important to note the fact that Late-Evening and Late-Night have higher ratios due to the fact that they have a very low number of posts.

NEXT STEPS

- Deploy our findings on Heroku
- Share slide deck and documentation with Sasha and schedule a call for feedback
- Sync on items for the next sprint-if required

Attached are the resources we used for the project

Twitter Analysis Colab Notebook:

<https://drive.google.com/file/d/130DrunyYV1lqjYCzuwD3dNzJ0OCvSaNw/view?usp=sharing>

Facebook Analysis Colab Notebook:

<https://colab.research.google.com/drive/1AQno5AkMNPYbbqEX-6HWC715W-B1hJAq?usp=sharing>

Twitter Streamlit Colab Notebook:

https://colab.research.google.com/drive/1E9_mVu5J6zfcAAQrY0Vu-XF2N-jXAZQn?usp=sharing

Facebook Streamlit Colab Notebook:

<https://colab.research.google.com/drive/1CJfdRaaCfiu98lggUDddVCRJSo60-UOs?usp=sharing>

Slide Deck:

https://www.canva.com/design/DAEoGfET78E/lgsuRzMOW23LUJ4NNKcShA/view?utm_content=DAEoGfET78E&utm_campaign=designshare&utm_medium=link&utm_source=sharebutton

Twitter and Facebook Scraping of Moringa School Data

Tableau: