

# Twitter Data Analysis of Ministries, Department and Agencies in Nigeria

Adebola K. Ojo, Ahmed B. Olanrewaju

**Abstract--** The use of social media tools as a means of communication by various ministries, departments and agencies (MDAs) in Nigeria started very late compared with personal use or in the private or corporate sector. The social media tool of interest is Twitter which is a microblogging application and it is becoming a momentous element of the public sector social media agenda. The study revealed the interest of citizens in the activities of the various MDAs in the country, some MDAs are not known to be too active since their activities are not given the required coverage in the contemporary media or the sector seems not to be receiving adequate attention and patronage. This paper made an empirical and methodological contribution to this new body of knowledge by presenting an overview study of general Twitter accounts maintained by the various MDAs of the Nigerian government. Over 70,000 tweets were used from 64 officially available Twitter accounts. It was discovered that the gradual use of Twitter is really creating a more engaging opportunity for the citizens to have firsthand information about the activities of the government agencies. The study revealed that the trend progressed over the years, that is, there was an upward movement in the use of twitter as a means of government citizen engagement over a 10 years' period. It was discovered that the use of twitter was very high during the office hours and more activities between Mondays and Fridays. The Nigerian Football Federation led the first five MDAs. Most of the devices used by MDAs to access are laptops or desktops, followed by smart phones. Android devices are more in use than iPhone devices. The remaining devices are not popular in the public services.

## Index Terms

Twitter, Microblogging, Government Agencies, Trend, Media, Nigerian Government, Ministries

## 1. INTRODUCTION

Social media have been mainly defined to refer to “the many relatively inexpensive and widely accessible electronic tools that facilitate anyone to publish and access information, collaborate on a common effort, or build relationship” [1]. Social media are interactive computer-mediated technologies that facilitate the creation and sharing of information, ideas, career interests and other forms of expression via virtual communities and networks. [2]. In previous studies, it showed that there are about 4.4 billion Internet users as well as about 3.4 billion active social media accounts. It also stated that the average time spent on social media is about 116 minutes a day, which evidently shows that a lot of people live on social media [3] [4]. The use of social media as a means of communication by the various ministries, departments, and government agencies of Nigeria started somewhat recently. This use was mostly driven by few stakeholders. Social media classifications based on information half-life and depth, and associated marketing objectives and purposes were discussed in [4] and [5].

We are primarily interested in the use of Twitter, which has revolutionized conventional media outlets (such as radio, television, newspapers and books). [6] [7] [8]

This study reveals the activities of the various Ministries, Department and Agencies (MDAs) in the country. Some MDAs are relatively inactive. Furthermore, relationships being created with various government agencies can also be seen in terms of interactions with their Twitter accounts.

## 2. MATERIALS AND METHOD

This work studied the general use of Twitter accounts maintained by the various MDAs of the Nigerian government. Over 70,000 tweets were used from the 64 officially available Twitter accounts. The data was generated on for all the Ministries, Departments and Government agencies in Nigeria.

Figure 1 shows the proposed methodology used for the study.

- Adebola Ojo is a senior lecturer and a researcher in the Department of Computer Science of the University of Ibadan, Nigeria. She obtained her Ph.D. and MSc Degrees in Computer Science from University of Ibadan, Nigeria. Her research interests are in Digital Computer Networks, Data Mining, Text Mining and Computer Simulation. She is also into data warehouse architecture, design and data quality via data mining approach. E-mail: [adebola\\_ojo@yahoo.co.uk](mailto:adebola_ojo@yahoo.co.uk)
- Ahmed B. Olanrewaju obtained his MSc degree in Computer Science Department, University of Ibadan. His research interests are Machine Learning and Text/Data mining. E-mail: [abono2000@gmail.com](mailto:abono2000@gmail.com)

The Python libraries were used for the following pre-processing tasks:

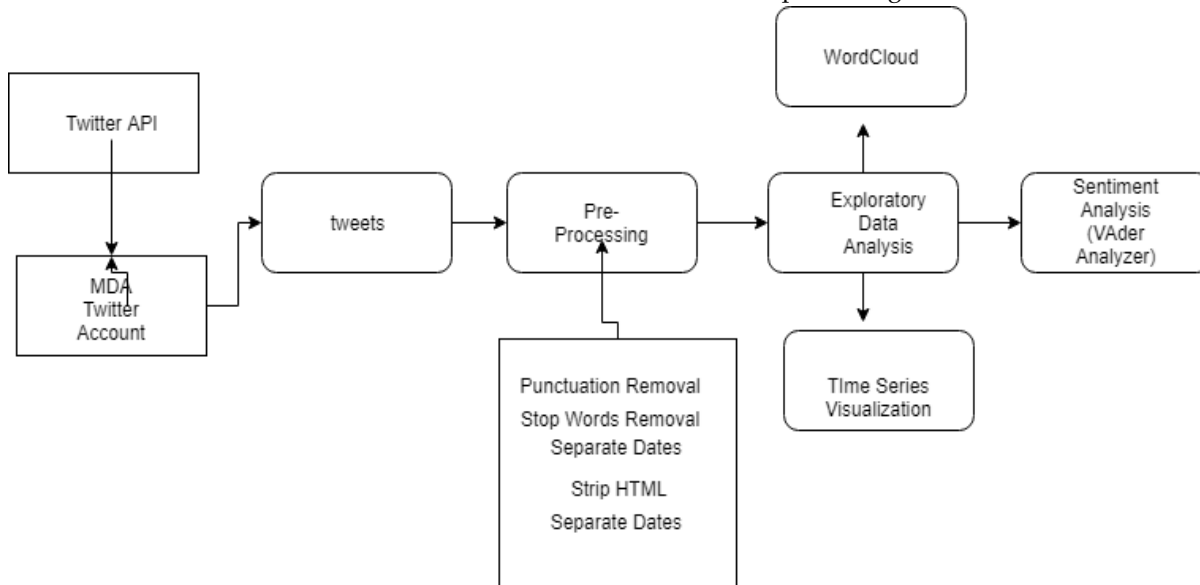


Fig 1:The Methodology

## 2.1 Document Collection

Twitter provides an application programming interface which allows for access to activities of the account of a user, the maximum number of tweets (posts) was 3200. Since we are dealing with 64 different MDAs accounts, the accounts used were Alvan Ikoku, CCB Nigeria, Central Bank of Nigeria, Copyright Commission, DMO Nigeria, EFCC Nigeria, Enactus FUDMA, Environment Ministry, Fed Min of Info & Cu, Fed Min Of Justice, Federal Character Commission, FEDERAL MEDICAL CENTRE,GUSAU, Federal Ministry of Mines and Steel Devt, Federal Ministry of PWH, Federal Ministry of Transportation, FERMA Nigeria, FIRS Nigeria, FMITI Nigeria, FUTA, Futminna Nigeria, Health Ministry NGR, IALS, U. of London, ICRC, INEC Nigeria, JAMB, Legal Aid Council NG, MDCUOIT, Ministry of Defence, Ministry of Finance, NAMA, NAPITIN, NAPITIP, National Assembly, National Centre for Women Development, Abuja, National Planning NG, Natl Pension Comm, nbcgovng, NBS Nigeria, NCAC Nigeria, NCC Consumers, ncc.gov.ng, NCPC Abuja, NDIC, NDLEA Nigeria, NEPC, NERC Nigeria, News Agency of Nigeria, Nigeria Ministry of Foreign Affairs, NIPC, NITDA Nigeria, NOA Nigeria, NPHCDA, NSCDC, NTL, NYSC HQ Nigeria, Radio Nigeria, Reforms Bureau NGR, Supreme Court of Nigeria, The Nigerian Football federation, University of Jos, Water Resources NGR.

## 2.2 Text Pre-Processing and Text Transformation

These involve tokenization (text normalization), text transformation and feature selection. Each of the accounts was saved in csv format, after which they were merged as one file for the next stage, which is the Pre-processing stage.

- i. **Lower casing** - The first pre-processing step which was done to transform the tweets into lower case. This was to avoid having multiple copies of the same words. For example, while calculating the word count, 'Analytics' and 'analytics' would have been taken as different words.
- ii. **Stop words removal** - As discussed earlier, stop words (or commonly occurring words) were removed from the text data. For this purpose, a list of stop words was created. Predefined libraries were also used.
- iii. **Punctuation removal** - The next step was to remove punctuations; punctuations do not add any extra information while treating text data. Therefore, removing all instances of it helped to reduce the size of the training data.
- iv. **Frequent words removal** - In this stage, commonly occurring words were also removed from the text data. This is known as Pruning.
- v. **Rare words removal** - Similarly, just as the most common words were removed, this time, rarely occurring words were removed as well from the texts; because they are so rare, the association between them and other words is dominated by noise. Rare words were replaced with a more general form and then this resulted in higher counts. This is also known as Pruning.
- vi. **Spelling correction** - Tweets have a plethora of spelling mistakes. The timelines are often filled with hastily sent tweets that are barely legible at times. In that regard, spelling correction is a useful pre-processing step because this also helps in reducing multiple copies of words.

- vii. **Lemmatization** - Lemmatization is a more effective option than stemming because it converts the word into its root word, rather than just stripping the suffices (as in the case of stemming). It makes use of the vocabulary and does a morphological analysis to obtain the root word. Therefore, using lemmatization was preferred and used over stemming.
- viii. **Date Separation** - The date format used was a full date format (2018-07-18 06:55:05), in order to be able to do necessary visualization and analysis, which were later separated into year, month, day, hour.

The essence of this pre-processing was to make available data that could further be processed.

Other tasks involved after the pre-processing stage were:

- i. Comparing the source of device used for tweeting by various MDAs
- ii. Comparing tweets per hour, day, monthly and through the lifespan of the existence of the various accounts. the first account was created on September 9, 2008 which shows that at least MDAs have started using Twitter for quite some time.
- iii. Evaluation of retweets and favourite tweets in the dataset.
- iv. Exploratory data analysis to reveal insights and pattern from the dataset
- v. Word Cloud is a data visualization technique used for representing text data in which the size of each word indicates its frequency or importance.
- vi. Sentiment analysis was done using Vader Sentiment from the NLTK library.

### 3. RESULTS AND DISCUSSIONS

In this section, the results and various discussions on the work were presented. These were as follow:

#### Exploratory Data Analysis

Figure 2 shows the hourly activities of MDA confirm that they are truly government agencies and reveal office hours starting from 8am daily while Figure 3 presents the daily activities of MDA, which confirm that they are truly government agencies and have more activities between Mondays to Fridays.

#### A gradual progression in the use

Figure 4 presents an upward trend in the use of Twitter as a means of government citizen engagement over a 10 years' period.

#### Analysis of the MDA Tweets

Nigerians are football lovers, and they are the most engaging MDA followed by the Electoral body. Table 1 shows the top 5 MDAs as well as the bottom 5 MDAs.

Table 1 gives a very clear picture of the real nature of Twitter engagement by Nigerians. The first column represents the top five (5) MDAs in terms of highest number of likes i.e. interest in their tweets while the second column represents the bottom five (5) MDAs that users have fewer interest and are not interested in sharing.

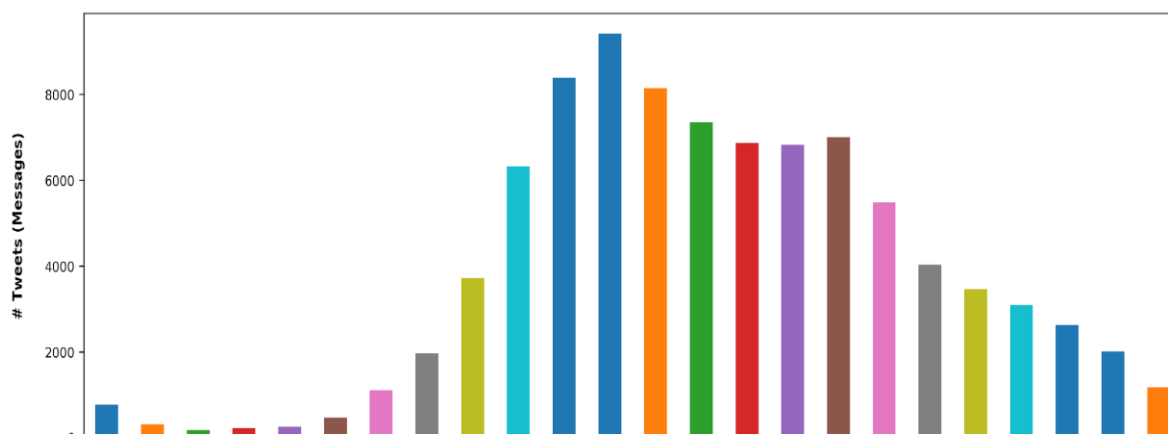


Fig. 2: Hourly activities of MDA

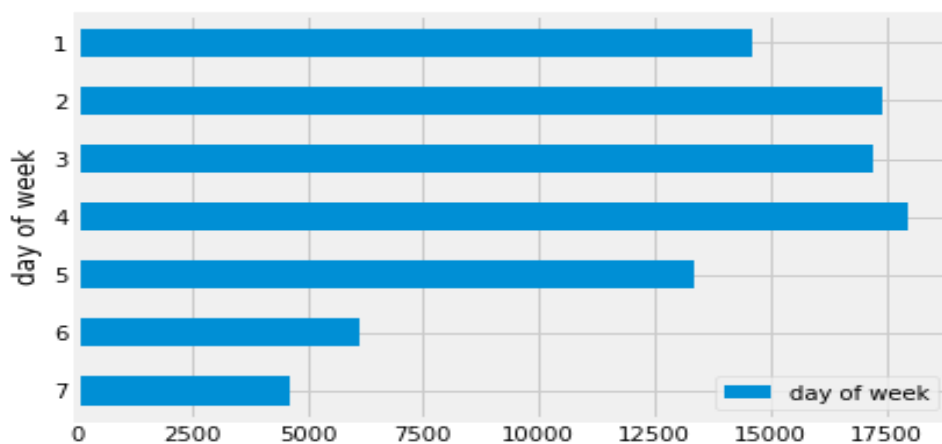


Fig. 3: Daily activities of MDA

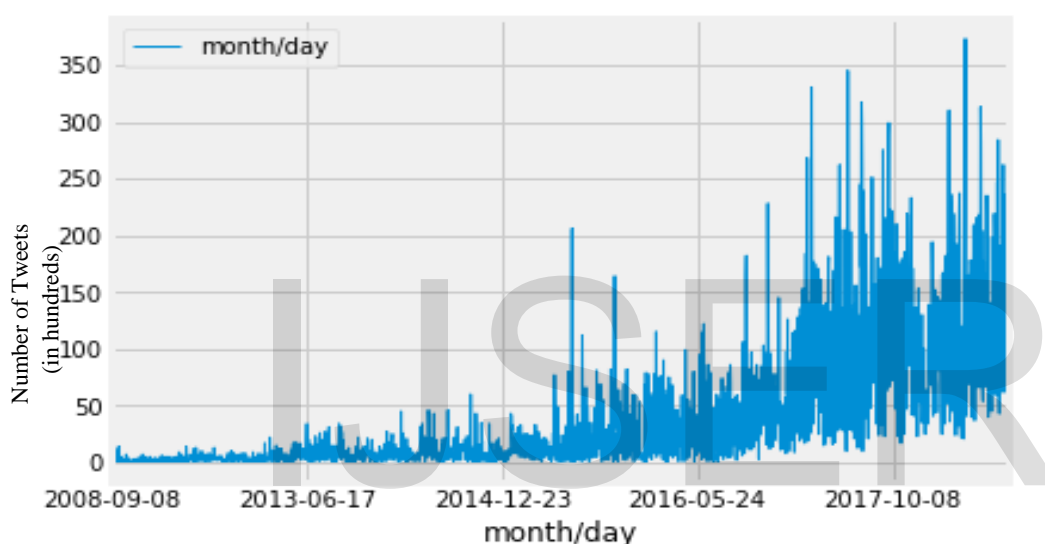


Fig.4: An upward trend in the use of Twitter as a means of government citizen engagement over a 10 years' period

TABLE 1: ANALYSIS OF THE MDA TWEETS

S/N	TOP 5 MDAs	BOTTOM 5 MDAs
1	The Nigerian Football Federation	Economic and Financial Crimes Commission
2	Independent National Electoral Commission	Legal Aid Council of Nigerian
3	Federal Ministry of Information and Culture	National Power Training Institute of Nigerian
4	News Agency of Nigerian	Supreme Court of Nigerian
5	Nigeria Bureau of Statistics	Federal Medical Centre, Gusau

The dataset clearly indicates that the only thing that unites Nigerians is soccer. Being football lovers, Nigerians easily forget their differences once it comes to the game of football. Nigerians are also interested in the information being released by the electoral umpire in the country which clearly indicates main interests and ability to sustain democracy in the country. The next three most active user engagement MDAs are Federal Ministry of Information and Culture which is more like the image maker of the country, The News Agency of Nigeria which is largely responsible for major news in the public sector domain in the country and empowered by law to do so and the Nigeria Bureau of Statistics which has actively been very popular on social media, and to actually testify to that, the Statistician General Dr Yemi Kale @sgyemikale is the most active public servant on Twitter. The bottom five (5) MDAs literally have little engagement on Twitter, due to low use of Twitter and a reflection that citizens are also not likely interested in them.

## Source of Tweets

TABLE 2: SOURCE OF THE TWEETS

Source of Tweets	Percentage Distribution
'Twitter Web Client'	31%
'Twitter for Android'	27%
'Twitter for iPhone'	14%
'Facebook'	6%
'dlvr.it'	4%
'Hootsuite'	3%
'WordPress.com'	3%
'Twitter for iPad'	2%
'Twitter for BlackBerry'	2%

In Table 2, the distribution of the devices used to access twitter was being investigated here, the device with the highest use is Twitter for Web client with thirty-one percent (31%) which means that the MDAs use their laptops or desktops to access twitter and it is stationary, which truly confirms that it is being used in an office and it is used in an official capacity. The next two devices are smart phones, Android and iPhone which have twenty-seven (27%) percent and fourteen percent (14%) respectively, which is another confirmation that Android devices are more in use than iPhone devices. The remaining devices are not popular in the public services.

## Retweet Frequencies

TABLE 3: RETWEET FREQUENCIES

Retweet Count	%
0	26%
1	14%
2	9%
3	7%
4	5%

In Table 3, the Retweet frequencies showed us more needs to be done on the part of the various MDAs, when it comes to social media engagement, 26% of the tweets were not retweeted that is, shared by citizens of Nigeria on social media either through the fact that it has no meaningful effect on citizens or the various MDAs do not provide regular update that will catch the attention of twitter users. The top 5 retweets frequencies showed that half of the tweets posted by all the MDAs did not get enough sharing by Twitter users.

## Word Cloud

Software as a Service (SaaS) is cloud computing best known model and delivers applications over the Internet as subscription-based services in a pay -as- you -go model to consumer and their use by end users is accelerating because of ease of access, cost savings, operational efficiency and enhanced flexible business capabilities. Online cloud consumers use Social media platforms like Facebook, twitter, review sites, blog, discussion forum etcetera and other different web platforms to express their views and emotions on different aspects of cloud services using natural language [9]. Figure 5 presents the word clouds of tweets. Nigerian is the word with highest frequency in the tweet's dataset.



Fig. 5: Word Clouds of Tweets

## Sentiment Analysis

Vader Sentiment Analyzer was used for the Sentiment Intensity Analyzer and categorizing as Positive, Negative and Neutral. It is general believed that all government agencies will report more of positive and neutral reviews about the government. This is clear seen in Figure 7

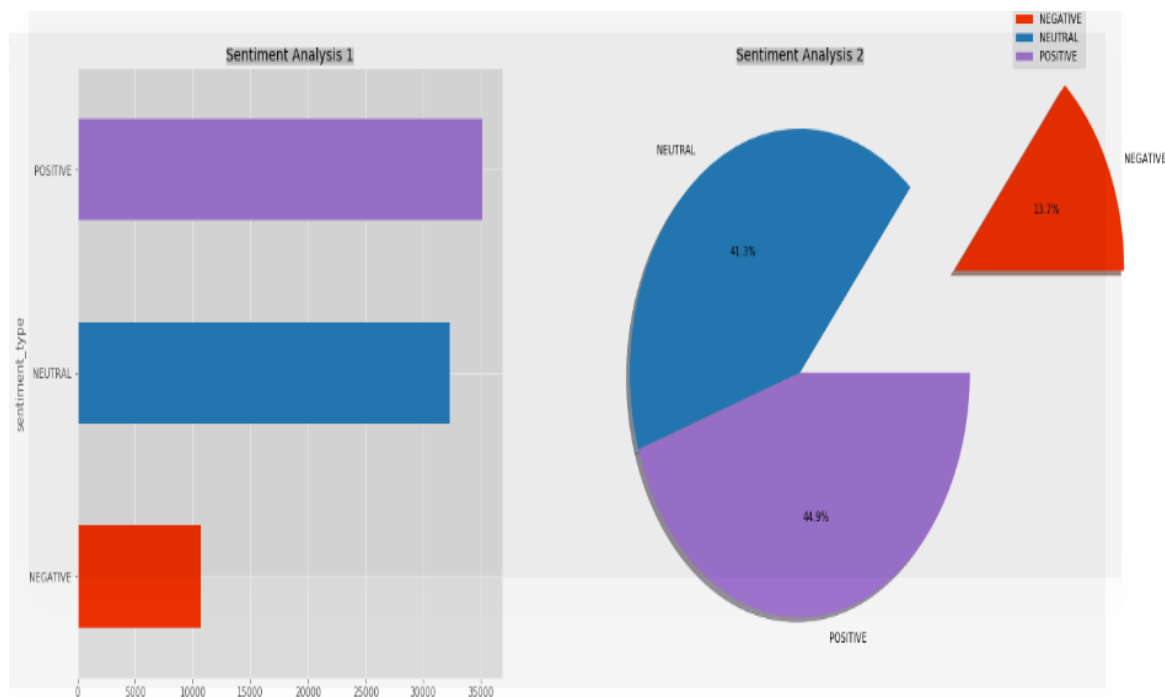


Fig.6: Sentiment Analysis

#### 4. CONCLUSION

It can be concluded that the gradual use of Twitter is really creating a more engaging opportunity for the citizens to have firsthand information about the activities of the government agencies. It was indicated that one of the major things that unite Nigerians is soccer. Most of the devices used by MDAs to access are laptops or desktops, followed by smart phones. Android devices are more in use than iPhone devices. The remaining devices are not popular in the public services.

#### REFERENCES

- [1] Shabnoor Siddiqui and Tajinder Singh, "Social Media its Impact with Positive and Negative Aspects," *International Journal of Computer Applications Technology and Research*, vol. 5, no. 2, pp. 71-75, 2016.
- [2] "Wikipedia," [Online]. Available: [https://en.wikipedia.org/wiki/Social\\_media](https://en.wikipedia.org/wiki/Social_media).
- [3] K. Smith, "126 Amazing Social Media Statistics and Facts," 13 June 2019. [Online]. Available: <https://www.brandwatch.com/blog/amazing-social-media-statistics-and-facts/>.
- [4] A. K. Ojo, "Improved Model for Detecting Fake Profiles in Online Social Network: A Case Study of Twitter," *Journal of Advances in Mathematics and Computer Science*, vol. 33, no. 4, pp. 1-17, 2019.
- [5] Tamás CSORDÁS, Éva MARKOS-KUJBUS, Mirkó GÁTI, "The Attributes of Social Media as a Strategic Marketing Communication Tool," *Journalism and Mass Communication*, vol. 4, no. 1, pp. 48-71, 2014.
- [6] Lovejoy, K., Waters, R. D., and Saxton, G. D., "Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. *Public Relations Review*," Elsevier, vol. 38, no. 2, pp. 313-318, 2012.
- [7] Lovejoy, K., and Saxton, G. D., "Information, community, and action: How nonprofit organizations use social media.," *Journal of Computer-Mediated Communication*, vol. 17, no. 3, pp. 337-353, 2012.
- [8] Kavanaugh, A. L., Fox, E. A., Sheetz, S. D., Yang, S., Li, L. T., Shoemaker, D. J., Natsev, A. and Xie, L., "Social media use by government: From the routine to the critical," Elsevier: *Government Information Quarterly*, vol. 29, no. 4, p. 480-491, 2012.
- [9] A. K. Ojo and Ako Ayito, "Detection Of Fake Product Reviews Using Bi-Directional Long Short-Term Memory And One-Dimensional Convolution Neural Network," *Egyptian Computer Science Journal*, vol. 43, no. 3, 2019.