

UNDERSTANDING CITIZEN  
OPINIONS TOWARDS ENERGY  
TRANSITION USING NATURAL  
LANGUAGE PROCESSING (NLP)  
TECHNIQUES.

VICTOR MOMODU

31<sup>st</sup> March 2022

# Table of Contents

Executive Summary .....	3
1 Introduction .....	4
1.1 Motivation .....	5
1.2 Research Question.....	5
2 Methodology .....	6
2.1 Data Collection .....	6
2.2 Data Cleansing and Pre-Processing .....	6
2.3 Method for Sentiment Analysis and Emotion Detection .....	8
3 Analysis and Results.....	9
3.1 Sentiment Analysis .....	9
3.2 Emotion Detection Analysis .....	12
3.3 Hypothesis .....	13
4 Conclusion and Limitations .....	14
5 References.....	15

# Executive Summary

The purpose of this research is to examine public perceptions on energy transition using social media data. There is a focus on understanding sentiments and emotions surrounding renewable energy in Canada and Nigeria. To mine the sentiments, data was extracted from twitter and detailed sentiment and emotion detection analysis was conducted using natural processing techniques in the RapidMiner Software. The findings are noteworthy because they show that, despite differences in economic growth, a high HDI country like Canada and a low HDI country like Nigeria have strikingly similar views on renewable energy. Both countries have very positive sentiments about energy revolution and ‘trust’ is the most elicited emotion detected. Despite disparities in the samples' socio-demographic characteristics, the analysis reflects the public's urgency and desire to integrate renewable energy into daily life. In addition to more well-known and structured polls, social media channels are effective in delivering information. Therefore, long-term monitoring of social media interactions could be valuable for policy development and tracking consensus sentiment.

# 1 Introduction

For a long time, traditional sources of energy (coal, oil, fossil fuels) powered economies of the world. However, as time passed, it had a negative impact on the environment, subsequently resulting in a decline in human wellbeing(Jain & Jain, 2019). In both industrialized and developing countries, renewable energy sources (RESs) are becoming an increasingly essential part of the energy mix(Haber et al., 2021). In May 2019, scientists, labor unions, native citizens, and youth throughout Canada convened over 150 town hall meetings across the country to establish the Pact for a Green New Deal (PGND). Delegates emphasized their wish for a world powered entirely by renewable energy, a 50% reduction in greenhouse gas emissions by 2030, and the development of one million new green jobs by 2030(MacArthur et al., 2020). Renewable energy and related sustainable technologies are predicted to reach very increased levels of adoption in energy technologies in developing nations like Nigeria, particularly in places with abundant solar and wind resources(Ram et al., 2022).

Public acceptance and awareness are required for a successful sustainable energy transition. The absence of sufficient communal knowledge regarding renewable energy resources, as well as a lack of "social acceptance" evaluation, are two major factors that can stymie widespread adoption of renewable energy systems (Abdar et al., 2020). Social perceptions of and public attitudes regarding various energy transition have been discovered to be core aspects that influence key decision-makers and end-users' actions on energy production, usage, and policy (Li et al., 2019).

The focus of this research is on people's attitudes towards renewable energy. From two different countries varying in Human Development Index (HDI), information would be gained about people's views on the environment by analyzing their tweets. Opinion Analysis or Sentiment Mining is the study of a person's qualities, feelings, reactions, and so on, in relation to entities such as products, social concerns, political processes, and their attributes supported by reaction from web sites, blogs, and social media (Sánchez-Rada & Iglesias, 2019)

## 1.1 Motivation

Public opinion and attitude on a variety of energy issues have a significant impact on our energy future. Systematic energy surveys and statistics are often used to collect this type of information, which can be costly and time-consuming. In present times, there are free communication platforms with vast volumes of public sentiments and preferences. Social media can be used as a new perspective for evaluating the dynamics of societal perception towards energy transition. Hence, critical concerns linked with social advocacy and sentiment over energy development, as well as policy implications, need to be identified. This could potentially assist the government with relevant information in proactively involving citizens in the fast-paced energy revolution.

## 1.2 Research Question

How do citizens' views on energy transition differ in nations with high and low HDI? A comparison would be made between two large oil producing countries that differ in terms of their HDI, with Canada having a high HDI (0.929) and Nigeria having a low HDI (0.539).

## 2 Methodology

### 2.1 Data Collection

The data used for this research and analysis is extracted from a popular social media platform called Twitter. Users express their thoughts, feelings, and perceptions on twitter using “tweets”. These tweets can then be analyzed, and major findings and insights can be derived. For this research, the tweets were mined with dates ranging from 1st March 2017 to 1st of March 2022. The data was gathered by searching for tweets with the term "renewable energy". The data were acquired using the snsrape python module, which allows users to scrape data from social network services (SNS) with no restrictions by providing a "keyword" or a specific "hashtag". A total of 20,000 tweets were retrieved with 10,000 tweets from users in Canada and 10,000 tweets from users in Nigeria.

### 2.2 Data Cleansing and Pre-Processing

The major goal of the cleaning process was to remove all data from the analysis that was unrelated to our research topic. For this project, the major tool used is the RapidMiner software. Rapid miner is an open-source application that combines data analysis, data integration, and reporting into one package (Singh Hanswal et al., 2019). All pre-processing steps were implemented in rapidminer with the following rules:

- **Data Cleansing:** In this step, the data is cleaned, and unnecessary information that aren't valuable for sentiment analysis is removed. These include removal of empty tweets, URLs, HTML tags, hashtags, and RT (Retweets). Then, duplicate tweets are dealt with and removed from the data. This results in a corpus size of 9075 and 9058 for Canada and Nigeria respectively.

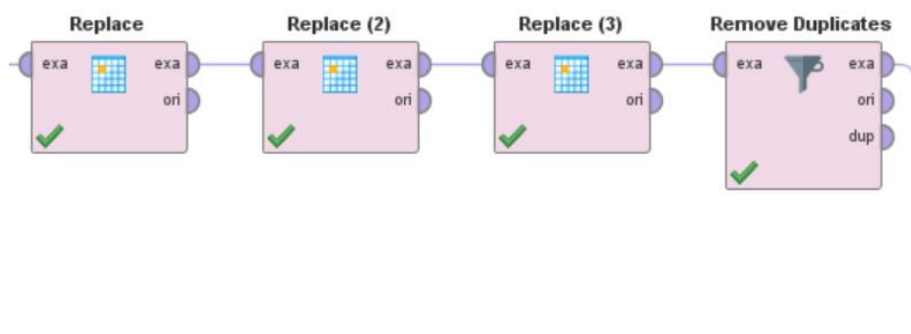


Figure 1: Operators used to perform data cleansing in Rapidminer

- The text is split into smaller chunks which are called tokens. These tokens are given as the input for the analysis.
- All words are lowercased. Punctuations and stop words are removed. Also, the tokens are filtered by length to ensure irrelevant tokens are excluded.
- Then, stemming process is implemented on the tokens which helps reduce the length of the word to its basic form. After this, these tokens would be ready for sentiment analysis.

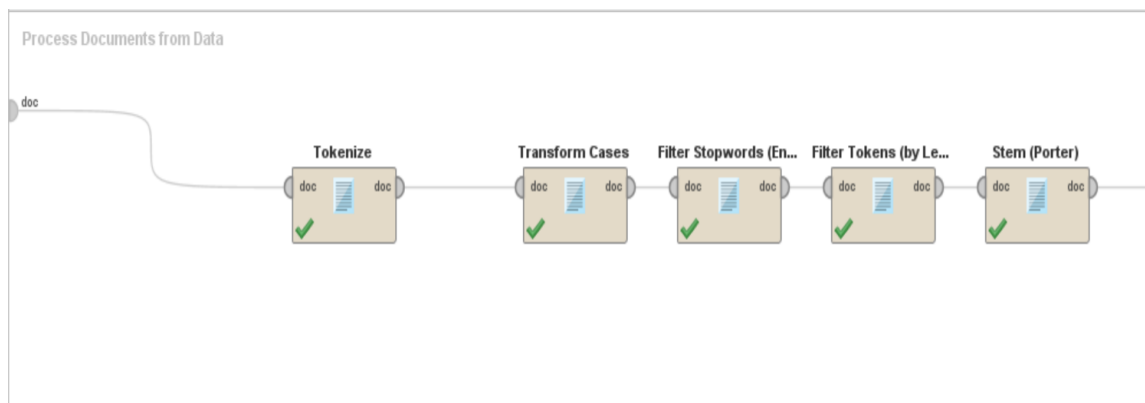


Figure 2: Operators used to perform data pre-processing

1	new vast major come canada mexico irrelev focus renew energi
2	thank marc andré pigeon martin boucher renata leonhardt present censu renewableenergi cooper canada host coopstudi jsgspp access report
3	roycesouth agre nuclear safe afford catastroph happen safe term storag radioact wast huge untap hydroelectr capac ontario canada want clean renew energi nuclear believ expert
4	curiou canada renewableenergi cooper locat check host good folk coopstudi
5	jkennei billion pipelin imagin portion funnel greenenergi canada lead exampl leav poor dinosaur grave leader renewableenergi thing learn ukrain dumb
6	commun power agenc communitypowera advocaci group social enterpris support commun own renew energi project inspir similar org scotland wale denmark germani canada austria
7	candicebergenmp stood make canada leader renew energi depend held hostag countri provinc cndpoli greenenergi
8	politicscarlo duncande jkennei enter renew energi select canada great option peopl want suffer elit crème brûlée
9	cdnenergycentr billion forget canada renew energi
10	solvest north klondik solarpow plant megawatt scale grid ti independ power product project northern canada oper detail solar solarenergi gosolar renewableenergi sustain cleanenergi
11	final comment draft green renew energi standard canada unit state propos updat elig length util tariff term contract option
12	ibensongitau keystone pipelin export sand canada fight renew energi independ secur freedom
13	garossino work canadian build compon renew energi machin take time costlii canada subsid cost foreign labour countri child labour
14	ikennei wish folk stop net fossil fuel natur start net renew energi resourc solar wind hydrr source energi abund canada

Figure 3: Sample of tokens after all data preprocessing steps

## 2.3 Method for Sentiment Analysis and Emotion Detection

In obtaining the sentiment from the tweets of users in Canada and Nigeria, a lexicon-based approach is used. A special operator called “Extract Sentiment” as contained in the RapidMiner software is deployed. This operator generates a sentiment score from an existing text attribute using open-source sentiment dictionaries. The operator contains several models and has inbuilt data pre-processing steps for accurate sentiment analysis. In the project, the text was scored using the VADER (Valence Aware Dictionary and Sentiment Reasoner) lexicon and rule-based sentiment. VADER is specifically tuned in to social media feelings and generates scores based on a lexicon of terms. The sum of all sentiment word scores in the text is calculated and then displayed by this operator (Hutto & Gilbert, 2014). This helps determine if a tweet was either a positive, negative, or neutral sentiment.



Figure 4: Extract Sentiment Operator in RapidMiner

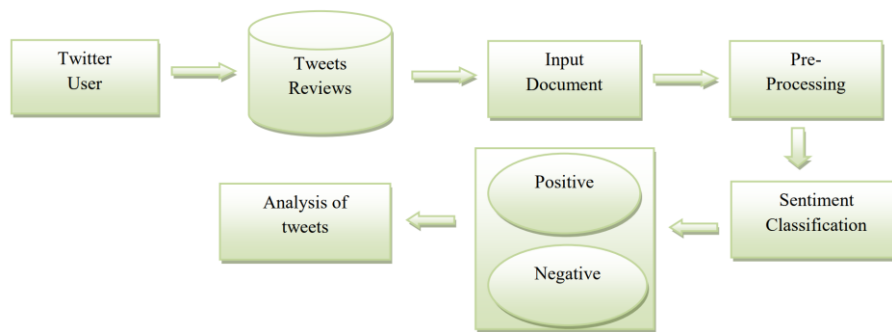


Figure 5: Process of Sentiment Analysis (Singh Hanswal et al., 2019)

Furthermore, the emotions evoked in the different tweets is detected. To analyse this, the EmoLex lexicon which was produced by the National Research Center of Canada (NRC) is employed (Mohammad & Turney, 2013). The NRC Emotion Lexicon includes a list of English terms and their relationships with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) as well as two primary sentiments (negative and positive). This system of categorizing emotions has gained a lot of traction around the world (Loureiro & Alló, 2020). The analysis is focused on the count of words categorised according to the emotions conveyed by the words used in the corpus of text.



## 3 Analysis and Results

### 3.1 Sentiment Analysis

From the sentiment analysis, out of 9075 tweets analysed in Canada, 7441(82%) had a positive sentiment towards renewable energy. About 11.3% (1025) of the total corpus had negative sentiments towards the energy transition. However, 6.7% of the analysed text had neutral sentiments. Overall, this implies that citizens in Canada are optimistic about renewable energy.

Index	Nominal value	Absolute count	Fraction
1	positive	7441	0.820
2	negative	1025	0.113
3	neutral	609	0.067

Figure 6: Statistics of sentiment analysis result in Canada

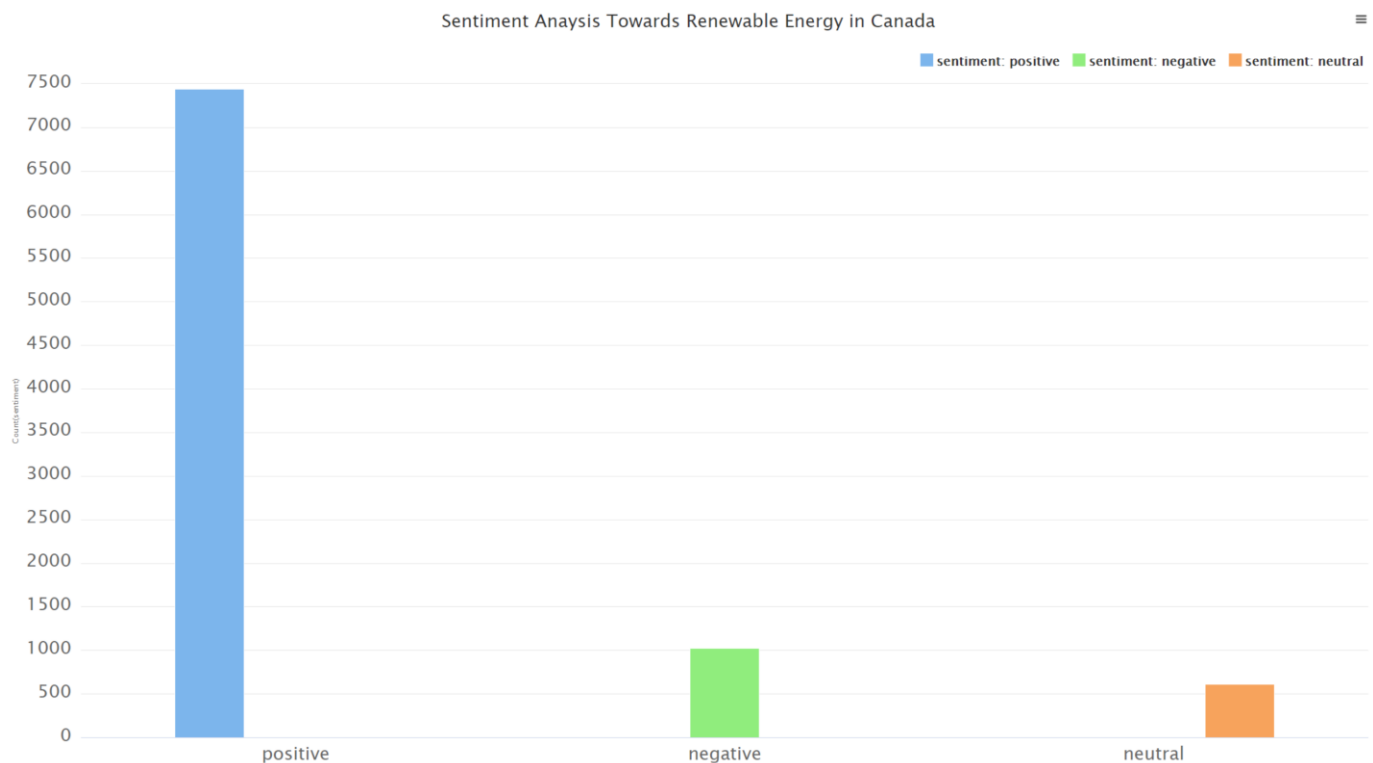


Figure 7: Plot of Sentiment Analysis in Canada

Regarding the sentiment analysis for Nigeria, 84.4% of tweets showed citizens have a positive sentiment towards renewable energy. About 8.6% of the total tweets extracted revealed negative sentiments. 630 tweets neither expressed positive or negative sentiments.

Index	Nominal value	Absolute count	Fraction
1	positive	7647	0.844
2	negative	781	0.086
3	neutral	630	0.070

Figure 8: Statistics of sentiment analysis result in Nigeria

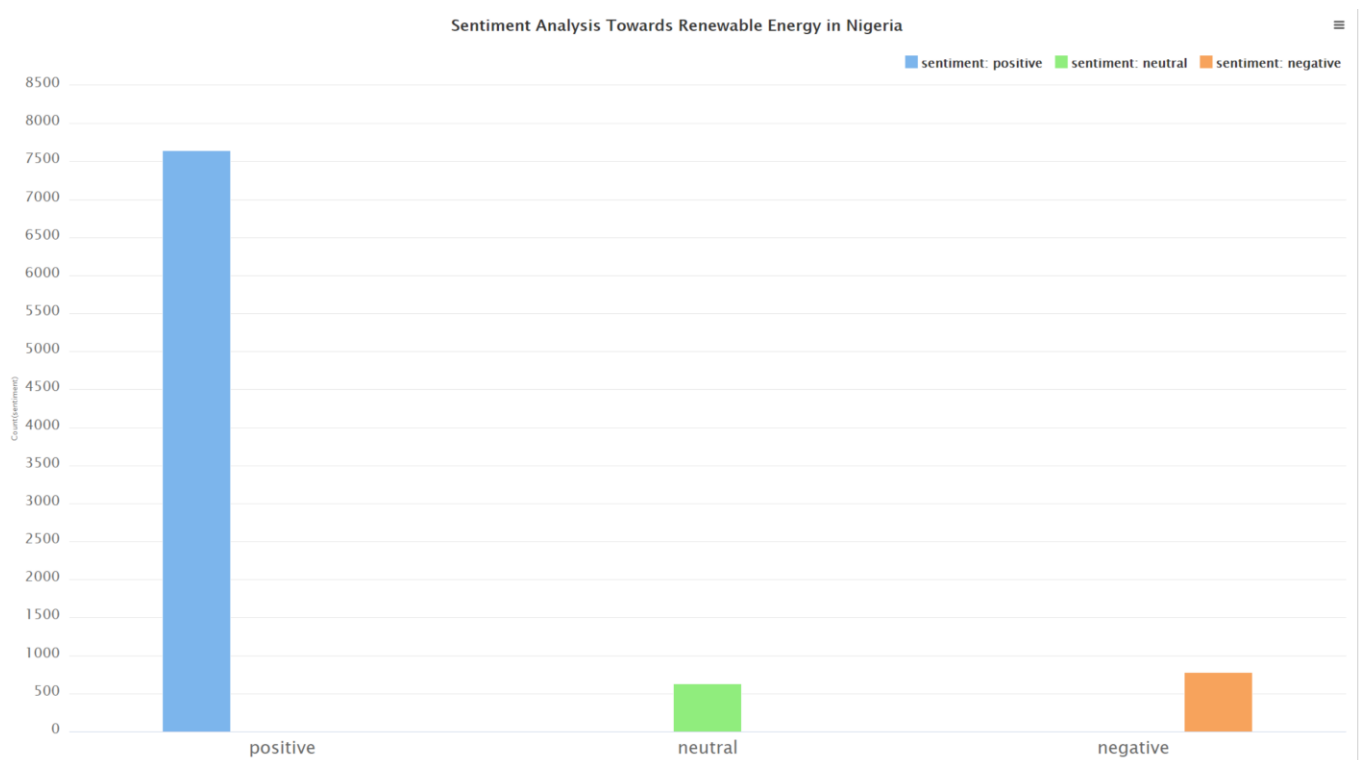


Figure 9: Plot of Sentiment Analysis in Nigeria



Figure 10: Descriptive statistics of sentiment analysis result in Nigeria.

Based on the findings above, most citizens in Canada and Nigeria have positive sentiments towards renewable energy. Nevertheless, it is important to note that the percentage of positive sentiment in Nigeria (84.4%) is slightly higher to Canada (82%). This could imply that citizens in Nigeria are more enthusiastic about renewable energy compared to citizens in Canada.

## 3.2 Emotion Detection Analysis

Detecting emotions gives a deeper understanding to the perception of the public towards renewable energy. In terms of frequency of classification of the words used in the corpus of tweets, the most popular emotion in Canada is trust accounting for 26.2% of the analysed corpus, which is followed closely by anticipation (18.54%) and joy (17.51%). The least emotion detected is disgust which constituted about 5.2% of the entire corpus. This implies that most Canadians are confident, comfortable, and very excited about the implementation of energy transition from fossil fuels to renewable energy.

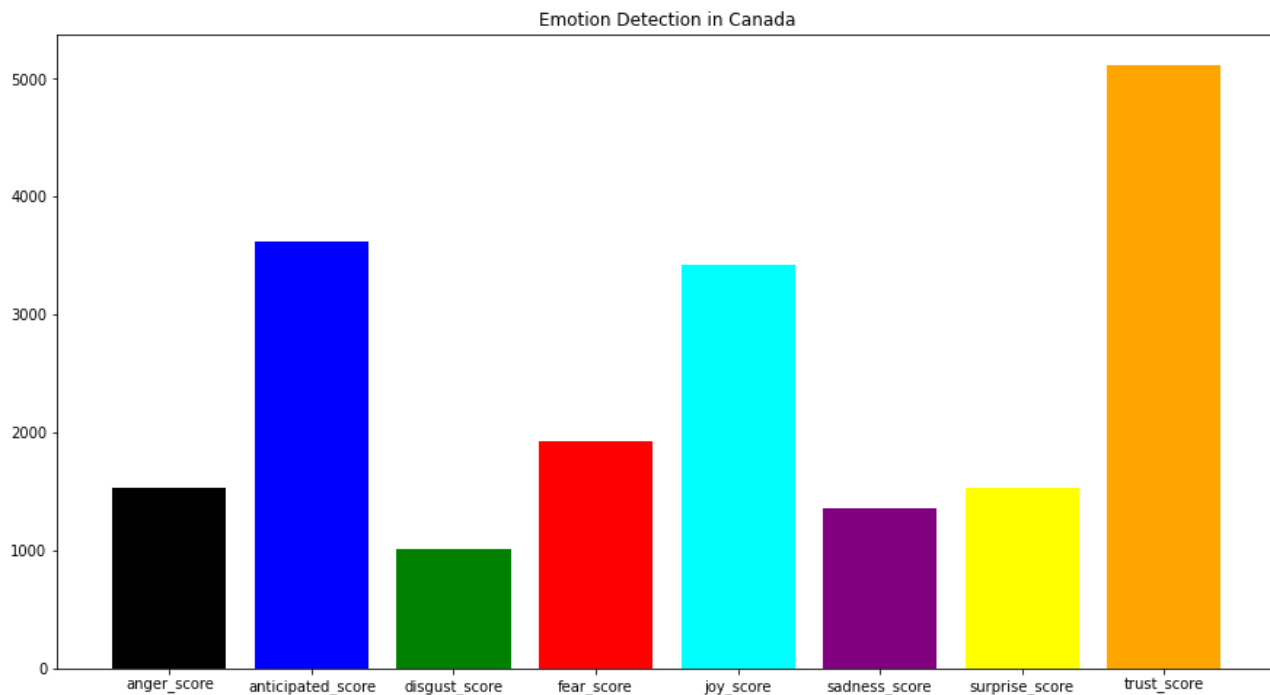


Figure 11: Emotion detection analysis in Canada.

emotion	frequency	percent
trust_score	5119	26.231104278759926
anticipated_score	3618	18.539584934665644
joy_score	3417	17.509607993850885
fear_score	1922	9.848834230079426
surprise_score	1535	7.865744299256981
anger_score	1533	7.855495772482706
sadness_score	1356	6.9485011529592615
disgust_score	1015	5.20112733794517

Figure 12: Emotion detection statistics in Canada.

Likewise, Nigeria also follows similar trend as Canada. The top three detected emotions are trust with 4503 word occurrences, which is followed by anticipation (3313) and joy (2838). The least emotions detected are sadness and disgust. This implies Nigerians are assured and happy with clean energy transition.

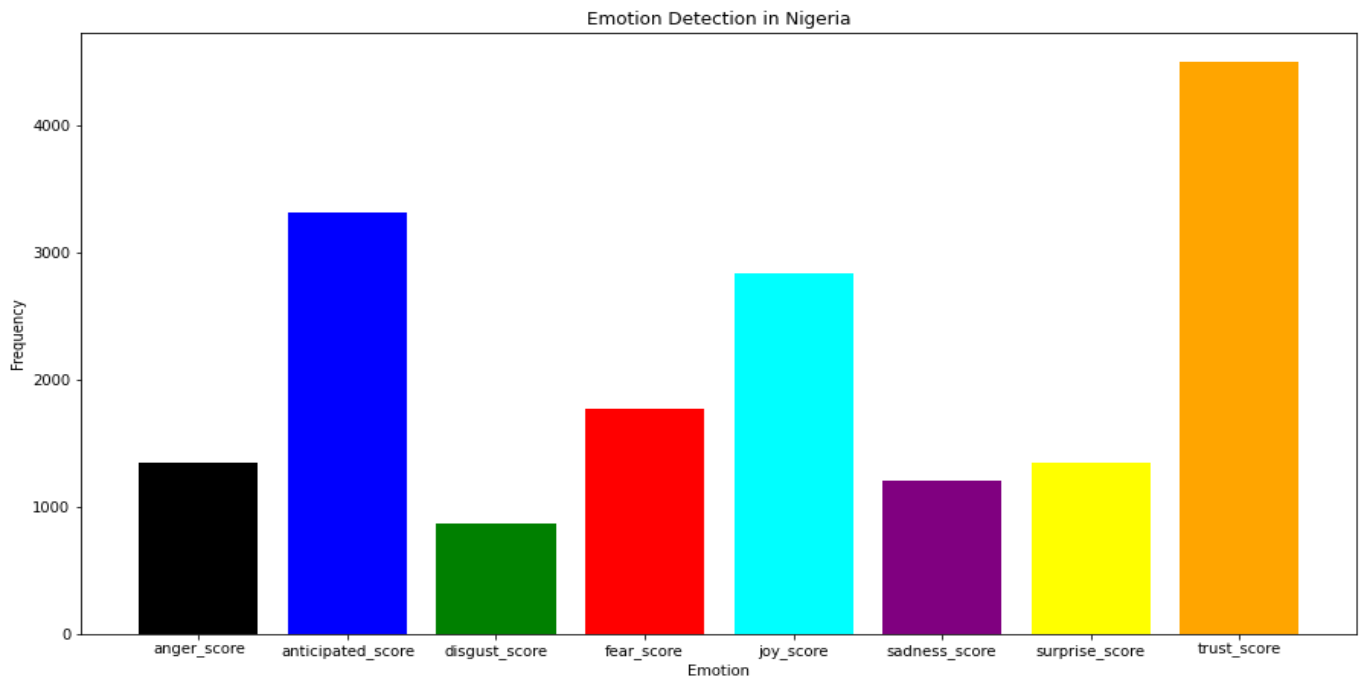


Figure 13: Emotion detection analysis in Nigeria.

emotion	frequency	percent
trust_score	4503	26.204608938547487
anticipated_score	3313	19.279562383612664
joy_score	2838	16.515363128491618
fear_score	1773	10.317737430167599
surprise_score	1347	7.838687150837989
anger_score	1342	7.809590316573558
sadness_score	1204	7.006517690875233
disgust score	864	5.027932960893855

Figure 12: Emotion detection statistics in Nigeria.

### 3.3 Hypothesis

Prior to the analysis, the following hypothesis was postulated:

- H0 - Nigerians are not enthusiastic about renewable energy
- **H1 - Nigerians and Canadians are enthusiastic about renewable energy**

From the sentiment and emotion detection analysis, Canadians and Nigerians are super excited about renewable energy. Hence, H0 will be rejected and H1 is accepted.

## 4 Conclusion and Limitations

The main purpose of this research is to analyse the public perception towards energy transition from twitter data extracted from Canada and Nigeria. There is a focus on understanding sentiments and emotions related to renewable energy in these countries. These findings are intriguing because they reveal that, although there are disparities in economic development, the view of renewable energy in a high HDI country like Canada and a low HDI country like Nigeria is remarkably similar. Both countries have trust as the most invoked emotion and have very positive sentiments towards energy transition. Despite the differences in the socio-demographic characteristics of the samples, the steady messages on energy transition concerns show the urgency of the public to integrate renewable energy to daily living. This study adds to the current body of knowledge in comparing the social perceptions of renewable energy in industrialized and developing countries. This research reveals that government and decision makers can harness public perceptions on societal topics to facilitate policy discussions and enhance decision making.

There are several limitations in this research. Although the methodology approach used in the current analysis is relatively robust in terms of assessing social media content, the irony or sarcasm in expressions and dialogues could not be identified. Additionally, the focus was exclusively on text during the data cleaning process, avoiding symbols and emoticons, which may create additional valuable information, especially in terms of emotional expression. Future improvements will incorporate the usage of deep learning approaches to overcome these flaws. Another major limitation to this research study is the accessibility of internet and social media in some regions. In a developing country like Nigeria, older people and the less educated have limited access to social media platforms as a result, lead to a selection bias.

In conclusion, social media platforms are useful in providing information in combination to other more well-known and structured surveys. As a result, long-term monitoring of social media interactions could be useful for policy formulation and societal opinion monitoring.

## 5 References

- Abdar, M., Basiri, M. E., Yin, J., Habibnezhad, M., Chi, G., Nemati, S., & Asadi, S. (2020). Energy choices in Alaska: Mining people's perception and attitudes from geotagged tweets. *Renewable and Sustainable Energy Reviews*, 124. <https://doi.org/10.1016/j.rser.2020.109781>
- Haber, I. E., Toth, M., Hajdu, R., Haber, K., & Pinter, G. (2021). Exploring public opinions on renewable energy by using conventional methods and social media analysis. *Energies*, 14(11). <https://doi.org/10.3390/en14113089>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International AAAI Conference on Weblogs and Social Media* 216. <http://sentic.net/>
- Jain, A., & Jain, V. (2019). Sentiment classification of twitter data belonging to renewable energy using machine learning. *Journal of Information and Optimization Sciences*, 40(2), 521–533. <https://doi.org/10.1080/02522667.2019.1582873>
- Li, R., Crowe, J., Leifer, D., Zou, L., & Schoof, J. (2019). Beyond big data: Social media challenges and opportunities for understanding social perception of energy. In *Energy Research and Social Science* (Vol. 56). Elsevier Ltd. <https://doi.org/10.1016/j.erss.2019.101217>
- Loureiro, M. L., & Alló, M. (2020). Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the U.K. and Spain. *Energy Policy*, 143. <https://doi.org/10.1016/j.enpol.2020.111490>
- MacArthur, J. L., Hoicka, C. E., Castleden, H., Das, R., & Lieu, J. (2020). Canada's Green New Deal: Forging the socio-political foundations of climate resilient infrastructure? In *Energy Research and Social Science* (Vol. 65). Elsevier Ltd. <https://doi.org/10.1016/j.erss.2020.101442>
- Mohammad, S. M., & Turney, P. D. (2013). CROWDSOURCING A WORD-EMOTION ASSOCIATION LEXICON. In *Computational Intelligence* (Vol. 29, Issue 3). <http://crowdsourcing.typepad.com/cs/2006/06>
- Ram, M., Bogdanov, D., Aghahosseini, A., Gulagi, A., Oyewo, A. S., Odai Mensah, T. N., Child, M., Caldera, U., Sadovskaia, K., Barbosa, L. D. S. N. S., Fasihi, M., Khalili, S., Traber, T., & Breyer, C. (2022). Global energy transition to 100% renewables by 2050: Not fiction, but much needed impetus for developing economies to leapfrog into a sustainable future. In *Energy* (Vol. 246). Elsevier Ltd. <https://doi.org/10.1016/j.energy.2022.123419>
- Sánchez-Rada, J. F., & Iglesias, C. A. (2019). Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. *Information Fusion*, 52, 344–356. <https://doi.org/10.1016/j.inffus.2019.05.003>
- Singh Hanswal, S., Pareek Sr Asst Professor, A., & Sharma Sr Asst Professor, A. (2019). Twitter Sentiment Analysis using Rapid Miner Tool. In *International Journal of Computer Applications* (Vol. 177, Issue 16).