VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

ON

"Sentiment analysis of lockdown in USA during covid 19"

Submitted in partial fulfilment for the award of degree(18CSI85)

BACHELOR OF ENGINEERING IN INFORMATION SCIENCE

Submitted by:

GANESH SHENOY

1NT20IS055

LOGO Ex:

Conducted at
varcons technologies pvt ltd.

NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

Department of information science

Accredited by VTU
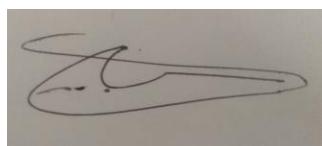
NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY
Department of information science
Accredited by VTU
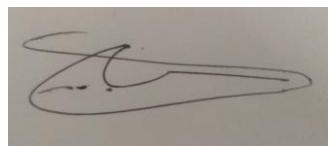
CERTIFICATE

This is to certify that the Internship titled **"Sentiment Analysis Of Lockdown In USA During Covid-19 A Case Study On Twitter using ML"** carried out by **Mr. Ganesh Shenoy,** a bonafide student of nitte meenakshi Institute of Technology, in partial fulfillment for the award of **Bachelor of Engineering**, in **information science** under VisvesvarayaTechnological University, Belagavi, during the year 2023-2024. It is certified that all corrections/suggestions indicated have been incorporated in the report.
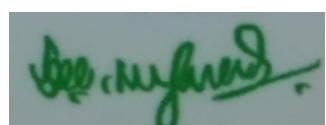
The project report has been approved as it satisfies the academic requirements in respect

of Internship prescribed for the course Internship / Professional Practice (18CSI85)

Signature of Guide　　　　　　**Signature of HOD**　　　　　　**Signature of Principal**

External Viva:

Name of the Examiner　　　　　　　　　　　　　　　　Signature with Date

1)_____

2)_____

_____

D E C L A R A T I O N

I, darshan gowda v, final year student of information science , nitte meenakshi institute of technology - 560064, declare that the Internship has been successfully completed, in **varcons technologies pvt ltd**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in information science, during the academic year 2023-2024.

Date :20/09/2023   :

Place : bengaluru

USN : 1NT20IS055

NAME : GANESH SHENOY

INTERNSHIP OFFER LETTER

SAIN INFORMATIX

Date: 11ᵗʰ August, 2023

Name: **GANESH SHENOY Y**
USN: **1NT20IS055**
Placement ID: **1308ML005**

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning with Python (Research Based)** Internship position with **Sain Informatix Pvt. Ltd.**, effective Start Date **11ᵗʰ August, 2023**, All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning with Python (Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working with you!.

Sincerely,

Nandini S
HR Manager
SAIN INFORMATIX PVT. LTD.
No. 1122, Cellar Floor, Service
Road,
Hampi Nagar
Bengaluru

Road,
Hampi Nagar
Bengaluru

A C K N O W L E D G E M E N T

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our Principal, for providing us adequate facilities to undertake this Internship.

We would like to thank our Head of Dept – branch code, for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We would like to thank our (Lab assistant name) Software Services for guiding us during the period of internship.

We express our deep and profound gratitude to our guide, Guide name, Assistant/Associate Prof, for her keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, forhelping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

ABSTRACT

The Covid-19 pandemic brought unprecedented challenges to nations worldwide, leading to the implementation of various containment measures, including lockdowns, to mitigate the spread of the virus. This study focuses on understanding the sentiment of the United States population towards lockdown measures during the Covid-19 pandemic, as expressed on the social media platform Twitter. Leveraging machine learning techniques, we analyze a vast dataset of tweets to gain insights into public sentiment.

The research employs natural language processing (NLP) and sentiment analysis to categorize tweets into positive, negative, or neutral sentiments. By examining the sentiments expressed in tweets over different phases of lockdown implementation and relaxation, we aim to provide valuable insights into the evolving public perception of lockdown measures.

Furthermore, this case study investigates the factors influencing sentiment variations, including regional disparities, political affiliations, and key events related to the pandemic. By uncovering these patterns, we can better understand how public sentiment was shaped by the dynamic circumstances of the pandemic.

The results of this study contribute to a deeper comprehension of public reactions to lockdowns during a global crisis and provide actionable insights for policymakers and public health authorities. By harnessing the power of machine learning and social media data, this research sheds light on the nuanced and evolving sentiment landscape surrounding Covid-19 lockdowns in the United States.

Table of Contents

Sl no	Description	Page no
1	Company Profile	
2	About the Company	
3	Introduction	
4	System Analysis	
5	Requirement Analysis	
6	Design Analysis	
7	Implementation	
8	Snapshots	
9	Conclusion	
10	References	

CHAPTER 1

COMPANY PROFILE

1. <u>COMPANY PROFILE</u>

A Brief History of Company

Company, was incorporated with a goal "To provide high quality and optimal Technological Solutions to business requirements of our clients". Every business is a different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions.They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Company is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements.

we strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the clients demand better. At our Company we work with them clients and help them todefiine their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstormingsession, and here they position themselves as an IT solutions consulting group comprising of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put itin one sentence " Technology helps you to Delight your Customers" and that is what we want to achieve.

CHAPTER 2

ABOUT THE COMPANY

2. <u>**ABOUT THE COMPANY**</u>

We are a Technology Organization providing solutions for all web design and development, Researching and Publishing Papers to ensure the quality of most used ML Models, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as a stakeholders to help us serve our clients with best of our capability and with at par industry standards.They have young, enthusiastic, passionate and creative Professionals to develop technological innovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solution. Motto of our organization is to "Collaborate with our clients to provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well". Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Services provided by Compsoft Technologies.

• Core Java and Advanced Java

• Research and Development/Improvise of ML Models

• Web services and development

• Dot Net Framework

• Python

• Selenium Testing

• Conference / Event Management Service

• Academic Project Guidance

• On The Job Training

• Software Training

CHAPTER 3

INTRODUCTION

3. <u>INTRODUCTION</u>

Introduction to ML

Machine Learning (ML) is a transformative field within the realm of artificial intelligence (AI) that has garnered immense attention and applications in recent years. It represents a paradigm shift in how we approach problem-solving and data analysis, enabling computers to learn from data and improve their performance over time without explicit programming. In the context of the proposed topic, "Sentiment Analysis of Lockdown in the USA During Covid-19: A Case Study on Twitter using ML," it's crucial to understand the fundamental concepts of ML as they underpin the methodology employed in this case study.

Problem Statement

During the Covid-19 pandemic, governments around the world, including the United States, implemented various measures to curb the spread of the virus. One of the most significant measures was the imposition of lockdowns, which had far-reaching social, economic, and psychological impacts on the population. This case study aims to understand the sentiment of the general public in the USA regarding these lockdown measures by analyzing Twitter data. The primary goal is to utilize machine learning techniques to extract valuable insights from a large volume of Twitter data .

CHAPTER 4

SYSTEM ANALYSIS

4. <u>SYSTEM ANALYSIS</u>

1. Existing System

In the context of system analysis, the "Existing System" refers to the current state or the way things are done before any proposed changes or improvements are made. It includes the processes, technologies, and procedures in place at the moment. Here, you would typically describe the system's limitations, issues, and drawbacks. For example, if we were analyzing an existing inventory management system, we might discuss issues such as manual data entry errors, lack of real-time tracking, or inefficient reporting processes.

2. Proposed System

The "Proposed System" represents the changes, improvements, or solutions that are being introduced to address the limitations or issues of the existing system. It outlines the new processes, technologies, or procedures that will be implemented. Continuing with the inventory management system example, the proposed system might include plans to introduce barcode scanning for accurate data entry, real-time inventory tracking using RFID technology, and automated reporting to improve efficiency.

3. Objective of the System

The "Objective of the System" refers to the goals and purposes for which the system is being analyzed and developed. It explains why the existing system needs improvement or why a new system is being introduced. The objectives can vary widely depending on the context and the specific problem or need . In the inventory management system example, the objectives might include:

a. Improve Accuracy: To reduce data entry errors and inaccuracies in inventory records.

b. Enhance Efficiency: To streamline inventory tracking and reporting processes, saving time and resources.

c. Increase Transparency: To provide real-time visibility into inventory levels and transactions.

d. Optimize Inventory Control: To enable better inventory control and reduce overstocking or understocking issues.

e. Enhance Decision-Making: To provide management with accurate and timely data for informed decision-making.

f. Reduce Costs: To lower operational costs associated with manual processes and errors.

The objectives serve as the guiding principles for designing and developing the proposed system to meet specific needs and improve upon the limitations of the existing system.

CHAPTER 5

REQUIREMENT ANALYSIS

5. <u>REQUIREMENT ANALYSIS</u>

Hardware Requirement Specification

Hardware requirements specify the physical components and infrastructure needed to support the proposed system. These requirements will vary depending on the nature of the system. Here is a general outline for hardware requirements:

1. Server Infrastructure:Specify the type of server(s) required (e.g., web server, database server).Mention the server specifications, including CPU, RAM, storage capacity, and network bandwidth.Discuss redundancy and failover requirements for high availability if applicable.

2. Client Devices:Describe the types of client devices that will access the system (e.g., desktops, laptops, mobile devices).Specify minimum hardware specifications for these devices (e.g., minimum RAM, processor speed, screen resolution).

3. Networking:Detail network infrastructure requirements, including switches, routers, and firewall configurations.Specify network bandwidth requirements, especially if the system relies on data-intensive operations.

4. Storage:Outline storage requirements, including the type of storage (e.g., SSD, HDD).
5. Specify the storage capacity needed for databases, file storage, and backups.
6. Discuss data retention and archival requirements.

7. Peripheral Devices:List any specialized hardware devices required, such as barcode scanners, RFID readers, or sensors.

8. Security Hardware:Mention any security-related hardware requirements, such as biometric authentication devices or encryption hardware modules.

Software Requirement Specification

Software requirements specify the software components and functionalities needed to implement the proposed system. Here is a general outline for software requirements:

1. Operating System:Specify the required operating system(s) for servers and client devices (e.g., Windows Server, Linux, iOS, Android).

2. Programming Languages and Frameworks:List the programming languages and frameworks necessary for system development (e.g., Python, Java, .NET).

3. Database Management System (DBMS):Specify the DBMS for data storage and management (e.g., MySQL, PostgreSQL, MongoDB).Define database schema requirements.

4. Web Server Software:Mention the web server software needed (e.g., Apache, Nginx, IIS).

5. Middleware and APIs:List any middleware or APIs required for integrating different components of the system or connecting to external services.

6. Development Tools:Specify development tools, including integrated development environments (IDEs) and version control systems.

7. Security Software:Detail security software requirements, such as antivirus, firewalls, and intrusion detection systems.

8. User Interface (UI) Technologies:Describe the front-end technologies needed for the user interface (e.g., HTML, CSS, JavaScript, React).

9. Third-Party Libraries and Dependencies:List any third-party libraries, modules, or dependencies that the system relies on.

10. Licensing and Compliance:Ensure that all software components are properly licensed and compliant with legal and regulatory requirements.

11. Performance and Scalability:Specify software scalability requirements to handle anticipated user loads and data volumes.

12.   Testing and Quality Assurance Tools:Identify software testing tools and frameworks required for quality assurance and testing processes.

13.   Documentation and User Manuals:Include requirements for documentation, user manuals, and help systems.

14.   Backup and Recovery Software:Specify backup and recovery software for data protection and disaster recovery.

15.   Maintenance and Support Tools:Define tools for system monitoring, logging, and remote support.The Hardware Requirement Specification and Software Requirement Specification documents serve as critical guides for system developers and stakeholders, ensuring that the system is designed, implemented, and maintained to meet specific hardware and software needs and constraints.

CHAPTER 6

DESIGN ANALYSIS

6. <u>DESIGN & ANALYSIS</u>

Design and Analysis is a critical phase in the software development lifecycle where the high-level system architecture, components, and data flow are designed, and the feasibility and efficiency of the proposed system are analyzed. This phase sets the foundation for the actual implementation of the system. Here are key aspects of this phase:

System Architecture Design:

- Define the overall system architecture, including components, modules, and their interactions.
- Choose the architectural style (e.g., client-server, microservices, monolithic) that best suits the system's requirements.
- Create architectural diagrams (e.g., UML diagrams) to visualize the system's structure.

Database Design:

- Design the database schema, including tables, relationships, and constraints.
- Determine data storage and retrieval strategies, such as normalization and indexing.
- Choose an appropriate database management system (DBMS) based on requirements.

User Interface (UI) Design:

- Create wireframes and mockups of the user interface to visualize how users will interact with the system.
- Ensure a user-friendly and intuitive design that aligns with user needs and expectations.
- Consider responsive design for various devices and screen sizes.

Algorithm and Logic Design:

- Develop algorithms and logic for core functionalities of the system.
- Specify data processing workflows, decision-making processes, and data transformation steps.
- Optimize algorithms for performance and efficiency.

Security Design:

- Identify potential security threats and vulnerabilities and design security mechanisms to mitigate them.
- Implement access control, authentication, authorization, and encryption as needed.
- Ensure compliance with security best practices and standards.

Scalability and Performance Analysis:

- Analyze the system's scalability requirements to handle increasing user loads.
- Conduct performance testing and optimization to ensure the system meets performance goals.
- Consider horizontal and vertical scaling strategies.

Data Flow and Integration:

- Define how data will flow within the system and between external systems.
- Design data integration points, APIs, and data exchange formats.
- Ensure data consistency and integrity.

Feasibility Analysis:

- Assess the technical, operational, and economic feasibility of the proposed system.
- Identify potential risks and mitigation strategies.
- Evaluate the cost-effectiveness and return on investment (ROI).

CHAPTER 7

IMPLEMENTATION

7. <u>IMPLEMENTATION</u>

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and it constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods a part from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether theobjectives have been met. Errors are noted down and corrected immediately.

2. Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is dividedinto several modules and is developed individually. So unit testing is conducted to individual modules.

3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

CHAPTER 8

SNAPSHOTS

8. SNAPSHOTS

0.0.1 Data Analysis of Covid 19 Tweets

Task

Dataset

Task

- Text
  - Text Preprocessing
  - Sentiment Analysis
  - Keyword Extraction
  - Entity Extraction

```
[6]: # Check Columns
     df.columns
```

```
[6]: Index(['user_name', 'user_location', 'user_description',
'user_created',
 'user_followers', 'user_friends', 'user_favourites', 'user_verified',
        'date', 'text', 'hashtags', 'source', 'is_retweet'],
      dtype='object')
```

```
[7]: # Datatype
     df.dtypes
```

```
[7]: user_name          object
     user_location      object
     user_description   object
     user_created       object
     user_followers      int64
     user_friends        int64
     user_favourites     int64
     user_verified        bool
     date               object
     text               object
     hashtags           object
     source             object
     is_retweet           bool
     dtype: object
```

```
[8]: # Source/ Value Count/Distribution of the Sources
     df['source'].unique()
```
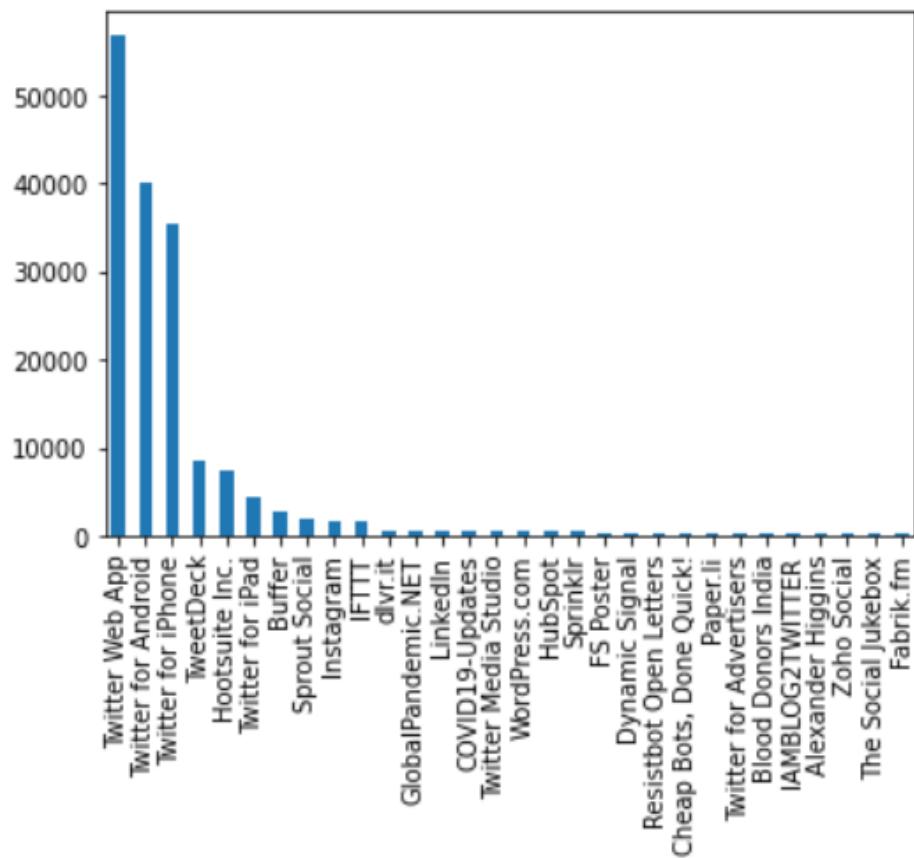
```
[8]: array(['Twitter for iPhone', 'Twitter for Android', 'Twitter Web
App',
```

'Buffer', 'TweetDeck', 'Twitter for iPad', 'Africa Newsroom',
'Blood Donors India', 'TweetCaster for Android',
'Alexander Higgins', 'IFTTT', 'Hootsuite Inc.', 'Sprout
Social',
'Sprinklr', 'assarofficial', 'IAMBLOG2TWITTER',
'CrowdControlHQ',
'COVID19-Updates', 'EveryoneSocial', 'Dynamic Signal',
'Instagram',
'TweetCaster for iOS', 'GlobalPandemic.NET', 'Venrap Radio',
'HeyOrca', 'Twitter for Advertisers', 'Paper.li',
'Twitter Media Studio', 'Twitter for Mac', 'dlvr.it',
'Cheap Bots, Done Quick!', 'Prof. Shanku', 'LaterMedia',
'SEMrush Social Media Tool', 'Twitterrific for iOS',
"Sebastian's Twitter Bot", 'Threader_client',
'COVID19FactoidBot',
'PwC UK SMART', 'tweet pro stiff', 'UK COVID-19 Alerts',
'Resistbot Open Letters', 'preprint-alert', 'ContentStudio.io',
'Peeping Moon', 'TweetAutomaticos', 'Orlo', 'AgoraPulse
Manager',
'Meltwater Social', 'Blog2Social APP',
'Social Genie by Brighter Vision', 'Social Media Publisher App
',

'Clearview Social, Inc.', 'Friends Me', 'Stuffed Productions',
'Revive Social App', 'Flamingo for Android',
'Microsoft Azure Logic Apps', 'MavSocial App', 'Smarp.',
'UberSocial for Android', 'Missinglettr', 'Canva', 'Talon
Android',
'Konnect Social', 'Siargao Guide', 'PinkVilla', 'Penname',
'Kashmir Life', "Monty's Twitter Reposter", 'SnapStream TV
Search',
'WatchDog Uganda', 'True Anthem', 'MarketingSuite', 'Corona-
Stats',
'Twitter for iPhone', 'Social Head', 'BookClubProTweet',
'covid19_counter', 'PromoRepublic', 'PolitiHUB.us', 'Raven
Tools',
'COVID19-Tweet', 'Cancer Health Auto Tweet', 'remote.io',
'Arena.im', 'Isrg', 'Constant Contact', 'Google', 'VegaLms',
'DTR Auto Post', 'CoSchedule', 'COVID19 Daily Stats',
'CovidDay',
'COVID-19MX', 'SmartNews | ', 'Spire FM', 'eClincher',
'Typepad', 'Publer ', 'Botbird tweets', 'shankar_live_bot',
'Shauntv', "Bob's Python tweetbot", 'Sprinklr Publisher',
'Commun.it Intelligence', 'Nintendo Switch Share',
"Stalin's Twitter Reposter", 'WeVideo', 'recurpost.com',
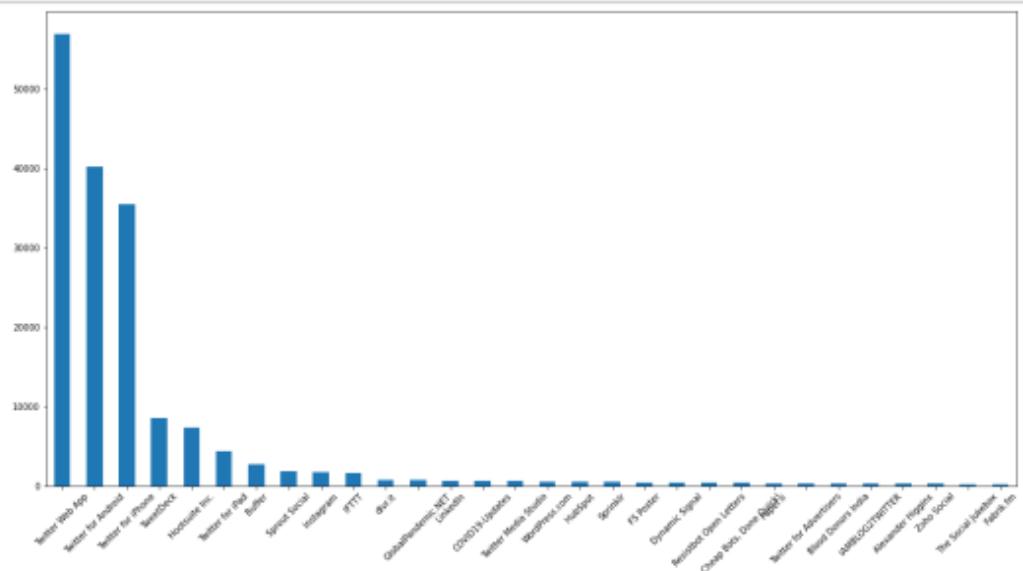
'El Cañonazo de las 9', 'Tweetings for Android',
'JoinDiaspora',
'Post Planner Inc.', 'DopeyUncle2', 'COVIDRecoveredBot',
'Qnary.io', 'Canada Covid-19 Stats', 'Twitter Test App P',
'Nelio Content', 'PlayStation®Network', 'Covid19_MV',
'Postcron App', 'TtwTimes Top News', 'Chicken Nugget',
'Plume\xa0for\xa0Android', 'Fenix 2', 'ETRetail.com',
'Campaign Share', 'Postfity.com', 'newsgovhk', 'Sociality.io',
'Echofon Android', "Rumpet's Twitter Poster", 'PostBeyond',
'dnh twitter publisher', 'news_by_gatfil', 'Pardot',
'App for IconicHipster.com', 'CovidUpdatesBot', 'KhuramKTS',
'NIAAuto', 'fovle', 'Imminent News', 'PublishBestNews',
'LFALSNAPAP', 'HW news english', 'StreamElements',
'Mediaproxy LogJam', 'dancehallaudio',
'Liberal Forum Political Chat', 'NippyTweet',
'That Best Home Automation', 'Dear_Assistant', 'Bot Libre!',
'NCoV-19 Tracker', 'tprzechlewski.app', 'DataBlogger',
'POZ Auto Tweet', 'SocialDog for Twitter', 'AtlantaTechBlogs',
'ReadyForSocial App', 'Traject Social', 'abrbuzz',
'Auto-Post (IMW)', 'ME Construction News', 'NetNaija Twiit',
'TrafficChiefNG', 'EUobserver', 'Flote.app', 'aa.com.tr',
'Twitterrific for Mac', 'Bizcommunity.com', 'Covid Robot',
'VoxPop Sync', 'stocktitan', 'StockTwits Web', 'Dr. Kill Pain',
'SociabbleApp', 'Post to Social by SHIFT1', 'PRNewswire',
'Hep Auto Tweet', 'Chorus publishing platform',
'masterdebater.net', 'A Touch of Snark', 'Login
CricketCountry',
"Let's Talk Singapore", 'RiteKit', 'Etsy', 'Zlappo.com',
'DirectorsTalk Auto Posting', 'GlobalVillageSpace',
'TeamSight Publisher', 'naalikeram', 'Amplifr', 'News Medical',
'Agenparl', 'Hacker Noon', 'Salsa Social Publishing',
'feedspora2',
'Tech_Tee_Shop', 'CensoredTodayPoster', 'Tumblr', 'Weebly App',
'science faction twitter bot', 'SocialHub by maloon',
'PNSposter',
'Social Reputation', 'RSS Masher - Dev', 'Planable', 'Trump
OwO',
"Bongani's Auto Poster", 'TTYtter', 'BAPSCharities',
'thefivedaily', 'RegulatorWatch.com', 'BrentStafford.com',
'Phil Ammann', 'flapol', 'Squarespace', 'VoiceFeed',
'Libsyn On-Publish', 'Sharpspring', 'Twitseq', 'Social Press
Kit',
'Lightful', 'Gain Platform', 'OregonCovidApp',
'Dynamics 365 for Marketing', 'yorkshire-times', 'Contento
App',

```
[9]: # Source/ Value Count/Distribution of the Sources
     df['source'].value_counts()
```

```
[9]: Twitter Web App          56891
     Twitter for Android      40179
     Twitter for iPhone       35472
     TweetDeck                 8543
     Hootsuite Inc.            7321
                              ...
     Emphatic                     1
     JustVent.LIVE                1
 Unshackled Content Poster        1
     DopeyUncle2                  1
     The Shirt List               1
     Name: source, Length: 610, dtype: int64
```

```
[14]: # Plot the top value_counts
      plt.figure(figsize=(20,10))
      df['source'].value_counts().nlargest(30).plot(kind='bar')
      plt.xticks(rotation=45)
      plt.show()
```


Text Analysis of tweet

```
[15]: #!pip install neattext
```

```
[16]: # Load Text Cleaning Package
      import neattext.functions as nfx
```

```
[17]: # Methods/Attrib
      dir(nfx)
```

```
[17]: ['BTC_ADDRESS_REGEX',
       'CURRENCY_REGEX',
       'CURRENCY_SYMB_REGEX',
       'Counter',
       'DATE_REGEX',
       'EMAIL_REGEX',
       'EMOJI_REGEX',
       'HASTAG_REGEX',
       'MASTERCard_REGEX',
       'MD5_SHA_REGEX',
```

```
'MOST_COMMON_PUNCT_REGEX',
'NUMBERS_REGEX',
'PHONE_REGEX',
'PoBOX_REGEX',
'SPECIAL_CHARACTERS_REGEX',
'STOPWORDS',
'STOPWORDS_de', 'STOPWORDS_en',
'STOPWORDS_es',
'STOPWORDS_fr',
'STOPWORDS_ru',
'STOPWORDS_yo',
'STREET_ADDRESS_REGEX',
'TextFrame',
'URL_PATTERN',
'USER_HANDLES_REGEX',
'VISACard_REGEX',
'__builtins__',
'__cached__',
'__doc__',
'__file__',
'__generate_text',
'__loader__',
'__name__',
'__numbers_dict',
'__package__',
'__spec__',
'_lex_richness_herdan',
'_lex_richness_maas_ttr',
'clean_text',
'defaultdict',
'digit2words',
'extract_btc_address',
'extract_currencies',
'extract_currency_symbols',
'extract_dates',
'extract_emails',
'extract_emojis',
'extract_hashtags',
'extract_html_tags',
'extract_mastercard_addr',
'extract_md5sha',
'extract_numbers',
'extract_pattern',
'extract_phone_numbers',
'extract_postoffice_box',
```

```
'extract_shortwords',
'extract_special_characters',
'extract_stopwords',
'extract_street_address',
'extract_urls',
'extract_userhandles',
'extract_visacard_addr',
'fix_contractions', 'generate_sentence',
'hamming_distance',
'inverse_df',
'lexical_richness',
'markov_chain',
'math',
'nlargest',
'normalize',
'num2words',
'random',
're',
'read_txt',
'remove_bad_quotes',
'remove_btc_address',
'remove_currencies',
'remove_currency_symbols',
'remove_custom_pattern',
'remove_custom_words',
'remove_dates',
'remove_emails',
'remove_emojis',
'remove_hashtags',
'remove_html_tags',
'remove_mastercard_addr',
'remove_md5sha',
'remove_multiple_spaces',
'remove_non_ascii',
'remove_numbers',
'remove_phone_numbers',
'remove_postoffice_box',
'remove_puncts',
'remove_punctuations',
'remove_shortwords',
'remove_special_characters',
'remove_stopwords',
'remove_street_address',
'remove_urls',
'remove_userhandles',
```

```
[20]: '@diane3443 @wdunlap @realDonaldTrump Trump never once claimed
      #COVID19 was a hoax. We all claim that this effort to…
      https://t.co/Jkk8vHWHb3'
```

0.0.2 Noise

- remove mentions/userhandles
- remove hashtags
- urls
- emojis
- special char

```
[21]: df.head()
```

```
[21]:        user_name        user_location \
      0                        astroworld
      1    Tom Basile         New York, NY
      2 Time4fisticuffs     Pewee Valley, KY
      3       ethel mertz Stuck in the Middle
      4       DIPR-J&K      Jammu and Kashmir


                                 user_description      user_created \
      0  wednesday addams as a disney princess keepin i… 2017-05-26
         05:46:42
      1  Husband, Father, Columnist & Commentator. Auth… 2009-04-16
         20:06:23
      2  #Christian #Catholic #Conservative #Reagan #Re… 2009-02-28
         18:57:41
      3  #Browns #Indians #ClevelandProud #[]_[] #Cavs … 2019-03-07
         01:45:06 4   Official Twitter handle of Department of Inf… 2017-02-
         12 06:45:15 user_followers user_friends user_favourites
         user_verified \
      0            624          950         18775        False
      1           2253         1677           24         True
      2           9275         9525         7254        False
      3            197          987         1488        False
      4         101009          168          101        False


                date                                           text \
      0 2020-07-25 12:27:21 If I smelled the scent of hand sanitizers
      toda…
      1 2020-07-25 12:27:17 Hey @Yankees @YankeesPR and @MLB - wouldn't
         it…
      2 2020-07-25 12:27:14 @diane3443 @wdunlap @realDonaldTrump Trump
         nev…
```

```
3  2020-07-25 12:27:10 @brookbanktv The one gift #COVID19 has give
   me…
4  2020-07-25 12:27:08 25 July : Media Bulletin on Novel
   #CoronaVirus…

                        hashtags            source is_retweet
   0                        NaN Twitter for iPhone    False
   1                        NaN Twitter for Android   False
   2                ['COVID19'] Twitter for Android   False
   3                ['COVID19'] Twitter for iPhone    False
   4 ['CoronaVirusUpdates', 'COVID19'] Twitter for    False
   Android
```

[22]: `df['text'].apply(nfx.extract_hashtags)`

```
[22]: 0                              []
      1                              []
      2                     [#COVID19]
      3                     [#COVID19]
      4        [#CoronaVirusUpdates,
               #COVID19]
                        …
      179103            [#WearAMask]
      179104             [#COVID19]
      179105                     []
      179106             [#COVID19]
      179107                     []
      Name: text, Length: 179108, dtype: object
```

[23]: `df['extracted_hashtags'] = df['text'].apply(nfx.extract_hashtags)`

[24]: `df[['extracted_hashtags','hashtags']]`

```
[24]:              extracted_hashtags                        hashtags
      0                           []     NaN
      1                           []     NaN
      2                   [#COVID19]     ['COVID19']
      3                   [#COVID19]     ['COVID19']
      4        [#CoronaVirusUpdates, #COVID19]
                                        ['CoronaVirusUpdates', 'COVID19']

      …                           …                        …
      179103          [#WearAMask]               ['WearAMask']
      179104           [#COVID19]                 ['COVID19']
      179105                   []                        NaN
      179106           [#COVID19]                 ['COVID19']
      179107                   []                        NaN
      [179108 rows x 2 columns]
```

```
[25]: # Cleaning Text
      df['clean_tweet'] = df['text'].apply(nfx.remove_hashtags)

[26]: df[['text','clean_tweet']]

[26]:                                                       text \
      0          If I smelled the scent of hand sanitizers toda…
      1          Hey @Yankees @YankeesPR and @MLB - wouldn't it…
      2          @diane3443 @wdunlap @realDonaldTrump Trump nev…
      3          @brookbanktv The one gift #COVID19 has give me…
      4          25 July : Media Bulletin on Novel #CoronaVirus…
      …                                                        …
      179103     Thanks @IamOhmai for nominating me for the @WH…
      179104     2020! The year of insanity! Lol! #COVID19 http…
      179105     @CTVNews A powerful painting by Juan Lucena. I…
      179106     More than 1,200 students test positive for #CO…
      179107     I stop when I see a Stop\n\n@SABCNews\n@Izinda…

                                                   clean_tweet
      0          If I smelled the scent of hand sanitizers toda…
      1          Hey @Yankees @YankeesPR and @MLB - wouldn't it…
      2          @diane3443 @wdunlap @realDonaldTrump Trump nev…
      3          @brookbanktv The one gift  has give me is an …
      4          25 July : Media Bulletin on Novel     \n@kansa…
      …                                                        …
      179103     Thanks @IamOhmai for nominating me for the @WH…
      179104     2020! The year of insanity! Lol! https://t.c…
      179105     @CTVNews A powerful painting by Juan Lucena. I…
      179106     More than 1,200 students test positive for a…
      179107     I stop when I see a Stop\n\n@SABCNews\n@Izinda…

      [179108 rows x 2 columns]

[27]: df['clean_tweet'] = df['clean_tweet'].apply(lambda x: nfx.remove_userhandles(x))

[28]: df[['text','clean_tweet']]

[28]:                                                       text \
      0          If I smelled the scent of hand sanitizers toda…
      1          Hey @Yankees @YankeesPR and @MLB - wouldn't it…
      2          @diane3443 @wdunlap @realDonaldTrump Trump nev…
      3          @brookbanktv The one gift #COVID19 has give me…
      4          25 July : Media Bulletin on Novel #CoronaVirus…
      …                                                        …
      179103     Thanks @IamOhmai for nominating me for the @WH…
      179104     2020! The year of insanity! Lol! #COVID19 http…
```

```
179105  @CTVNews A powerful painting by Juan Lucena. I…
179106  More than 1,200 students test positive for #CO…
179107  I stop when I see a Stop\n\n@SABCNews\n@Izinda…

                                      clean_tweet
0        If I smelled the scent of hand sanitizers toda…
1        Hey and - wouldn't it have made more sen… 2
             Trump never once claimed was a hoax. W… 3    The
         one gift has give me is an appreciatio…
4     25 July : Media Bulletin on Novel  \n         …
…                                                      …
179103    Thanks for nominating me for the    challen…
179104    2020! The year of insanity! Lol! https://t.c…
179105    A powerful painting by Juan Lucena. It's a t…
179106    More than 1,200 students test positive for    a…
179107    I stop when I see a Stop\n\n \n \n \n \n \n \n…

[179108 rows x 2 columns]
```

[29]: `df['clean_tweet'].iloc[10]`

[29]: `'49K+ Covid19 cases still no response from \n  cancel the compartment exa… https://t.co/kV2ZKmumu1'`

[30]:
```
# Cleaning Text: Multiple WhiteSpaces
df['clean_tweet'] = df['clean_tweet'].apply(nfx.remove_multiple_spaces)
```

[31]: `df['clean_tweet'].iloc[10]`

[31]: `'49K+ Covid19 cases still no response from cancel the compartment exa… https://t.co/kV2ZKmumu1'`

[32]:
```
# Cleaning Text : Remove urls
df['clean_tweet'] = df['clean_tweet'].apply(nfx.remove_urls)
```

[33]:
```
# Cleaning Text: Punctuations
df['clean_tweet'] = df['clean_tweet'].apply(nfx.remove_puncts)
```

[34]: `df[['text','clean_tweet']]`

[34]:
```
                                                    text \
0        If I smelled the scent of hand sanitizers toda…
1        Hey @Yankees @YankeesPR and @MLB - wouldn't it…
2        @diane3443 @wdunlap @realDonaldTrump Trump nev…
3        @brookbanktv The one gift #COVID19 has give me…
4        25 July : Media Bulletin on Novel #CoronaVirus…
```

```
      ...                                      ...
179103 Thanks @IamOhmai for nominating me for the @WH…
179104 2020! The year of insanity! Lol! #COVID19 http…
179105 @CTVNews A powerful painting by Juan Lucena. I…
179106 More than 1,200 students test positive for #CO…
179107 I stop when I see a Stop\n\n@SABCNews\n@Izinda…

                            clean_tweet
0              If I smelled the scent of hand sanitizers toda…
1              Hey and wouldnt it have made more sense to ha…
2              Trump never once claimed was a hoax We all cl…
3              The one gift has give me is an appreciation f…
4              25 July : Media Bulletin on Novel
      ...                                      ...
179103                 Thanks for nominating me for the
                       challenge I n…
179104                 2020 The year of insanity Lol
179105                 A powerful painting by Juan Lucena Its a
                       trib…
179106                 More than 1200 students test positive for
                       at m…
179107                 I stop when I see a Stop

[179108 rows x 2 columns]
```

0.0.3 Sentiment Analysis

```
[35]: from textblob import TextBlob
```

```
[40]: def get_sentiment(text):
          blob = TextBlob(text)
          sentiment_polarity = blob.sentiment.polarity
          sentiment_subjectivity = blob.sentiment.subjectivity
          if sentiment_polarity > 0:
              sentiment_label = 'Positive'
          elif sentiment_polarity < 0:
              sentiment_label = 'Negative'
          else:
              sentiment_label = 'Neutral'
          result = {'polarity':sentiment_polarity,
                    'subjectivity':sentiment_subjectivity,
                    'sentiment':sentiment_label}
          return result
```

```
[38]: # Text
      ex1 = df['clean_tweet'].iloc[0]
```

```
[41]: get_sentiment(ex1)
```

```
[41]: {'polarity': -0.25, 'subjectivity': 0.25, 'sentiment': 'Negative'}
```

```
[42]: df['sentiment_results'] = df['clean_tweet'].apply(get_sentiment)
```

```
[43]: df['sentiment_results']
```

```
[43]: 0         {'polarity': -0.25, 'subjectivity': 0.25,
      'sen… 1       {'polarity': 0.5, 'subjectivity': 0.5,
      'sentim…
      2         {'polarity': 0.0, 'subjectivity': 0.0, 'sentim…
      3         {'polarity': 0.0, 'subjectivity': 0.3571428571…
      4         {'polarity': 0.0, 'subjectivity': 0.0, 'sentim…
                                  …
      179103    {'polarity': 0.2, 'subjectivity': 0.2, 'sentim…
      179104    {'polarity': 0.8, 'subjectivity': 0.7, 'sentim…
      179105    {'polarity': 0.3, 'subjectivity': 1.0, 'sentim…
      179106    {'polarity': 0.26325757575757575, 'subjectivit…
      179107    {'polarity': 0.0, 'subjectivity': 0.0, 'sentim… Name:
                sentiment_results, Length: 179108, dtype: object
```

```
[44]: df['sentiment_results'].iloc[0]
```

```
[44]: {'polarity': -0.25, 'subjectivity': 0.25, 'sentiment': 'Negative'}
```

```
[45]: pd.json_normalize(df['sentiment_results'].iloc[0])
```

```
[45]:  polarity subjectivity sentiment
    0    -0.25          0.25 Negative
```

```
[47]: df = df.join(pd.json_normalize(df['sentiment_results']))
```

```
[48]: df.head()
```

```
[48]:        user_name        user_location \
    0                          astroworld
    1     Tom Basile         New York, NY
    2 Time4fisticuffs      Pewee Valley, KY
    3       ethel mertz Stuck in the Middle
    4       DIPR-J&K      Jammu and Kashmir


                             user_description      user_created \
    0        wednesday addams as a disney princess keepin i... 2017-05-26
             05:46:42
    1        Husband, Father, Columnist & Commentator. Auth... 2009-04-16
             20:06:23
    2        #Christian #Catholic #Conservative #Reagan #Re... 2009-02-28
             18:57:41
    3        #Browns #Indians #ClevelandProud #[]_[] #Cavs ... 2019-03-07
             01:45:06 4  Official Twitter handle of Department of Inf...
             2017-02-12 06:45:15 user_followers user_friends
             user_favourites user_verified \
    0          624           950           18775       False
    1         2253          1677              24        True
    2         9275          9525            7254       False
    3          197           987            1488       False
    4       101009 168    101   False


                 date                                          text \
    0 2020-07-25 12:27:21 If I smelled the scent of hand sanitizers
       toda...
    1 2020-07-25 12:27:17 Hey @Yankees @YankeesPR and @MLB - wouldn't
       it...
    2 2020-07-25 12:27:14 @diane3443 @wdunlap @realDonaldTrump Trump
       nev...
    3 2020-07-25 12:27:10 @brookbanktv The one gift #COVID19 has give
       me...
    4 2020-07-25 12:27:08 25 July : Media Bulletin on Novel
       #CoronaVirus...


                        hashtags         source is_retweet \
    0                        NaN Twitter for iPhone     False
    1                        NaN Twitter for Android     False
    2               ['COVID19'] Twitter for Android     False
```

```
3                    ['COVID19'] Twitter for iPhone      False
4 ['CoronaVirusUpdates', 'COVID19'] Twitter for        False
Android
                extracted_hashtags \
0                               []
1                               []
2                          [#COVID19]
3                          [#COVID19]
4               [#CoronaVirusUpdates, #COVID19]


                                clean_tweet \
0            If I smelled the scent of hand sanitizers toda…
1            Hey and wouldnt it have made more sense to ha…
2            Trump never once claimed was a hoax We all cl…
3            The one gift has give me is an appreciation f…
4            25 July : Media Bulletin on Novel


                                sentiment_results polarity subjectivity \
0 {'polarity': -0.25, 'subjectivity': 0.25,      -0.25    0.250000
'sen…
1 {'polarity': 0.5, 'subjectivity': 0.5,          0.50    0.500000
'sentim…
2 {'polarity': 0.0, 'subjectivity': 0.0,          0.00    0.000000
'sentim…
3 {'polarity': 0.0, 'subjectivity':               0.00    0.357143
0.3571428571…
4 {'polarity': 0.0, 'subjectivity': 0.0,          0.00    0.000000
'sentim…
  sentimen
t 0
Negative
1   Positive
2   Neutral
3   Neutral
4   Neutral
```

```
[49]: df['sentiment'].value_counts()
```

```
[49]: Neutral 75490
      Positive 74154
      Negative   29464
      Name: sentiment, dtype: int64
```
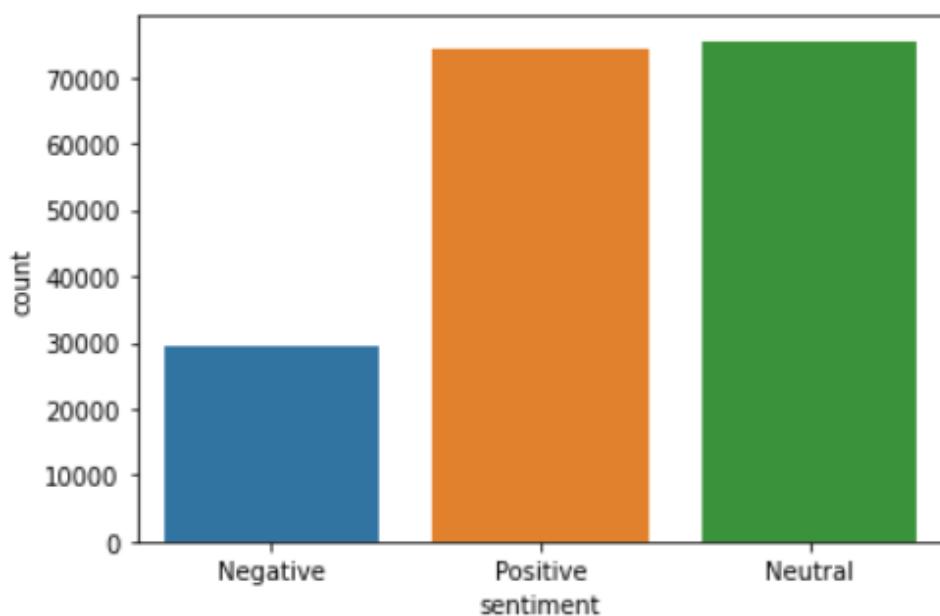
```
[50]: df['sentiment'].value_counts().plot(kind='bar')
```

```
[50]: <AxesSubplot:>
```

```
[51]:  # Plot with seaborn
       sns.countplot(df['sentiment'])
```

```
[51]: <AxesSubplot:xlabel='sentiment', ylabel='count'>
```

```
[ ]:  ### Keyword Extraction
      + For Positive and Negative Sentiment
      + General
```

```
[56]:  positive_tweet = df[df['sentiment'] == 'Positive']['clean_tweet']
```

```
[53]:  neutral_tweet = df[df['sentiment'] == 'Neutral']['clean_tweet']
       negative_tweet = df[df['sentiment'] == 'Negative']['clean_tweet']
```

```
[54]:  positive_tweet
```

```
[54]: 1        Hey and wouldnt it have made more sense to
      ha… 6 How Will Change Work in General (and recruitin…
      8             Praying for good health and recovery of
      12          — no one will be safe from until everyone i…
      13          Lets all protect ourselves from Its real and t…
                              …
      179096 We were really bummed we couldn't cop one of t…
      179103              Thanks for nominating me for the challenge I
                          n…
      179104              2020 The year of insanity Lol
      179105              A powerful painting by Juan Lucena Its a
                          trib…
```

```
    179106                          More than 1200 students test positive for at
                                    m…
    Name: clean_tweet, Length: 74154, dtype: object
```

```
[60]:  # Remove Stopwords and Convert to Tokens positive_tweet_list =
       positive_tweet.apply(nfx.remove_stopwords).tolist()
```

```
[57]:  negative_tweet_list =
       negative_tweet.apply(nfx.remove_stopwords).tolist()
       neutral_tweet_list =
       neutral_tweet.apply(nfx.remove_stopwords).tolist()
```

```
[61]:  positive_tweet_list
```

```
[61]:  ['Hey wouldnt sense players pay respects A…',
        'Change Work General (and recruiting specifically) via/',
        'Praying good health
        recovery', ' — safe safe
        commit ensure…',
        'Lets protect real numbers climbing fast Continent Lets n…',
        'Second wave Flandersback',
        'COVID Update: infection rate Florida following natural curve
       experts predicted initial cu…',
        'Good Patriots Volunteer Election Judge Polls open without…',
        'comprehensive review amp Analysis: key ways WASH help reduce
        transmission…',
        'crosses 2 lakh mark 150055 ppp recovered far positive today
        6988/total…',
        'Actor father MrGKReddy tested positive 15/20 days s…',
        'safe place visit guests said hotel meticulous applying hand
        sanitation als…',
        'CEO live WSU talking "The effects Covid19 students"…',
        'Actionables healthy recovery',
        'understood hadnt missed meeting In…',
        'released new podcast episodes week: Technology Platforms Conduct
       Telehealth Visits C…',
        'reports 1142 positive cases 29 deaths Saturday Death toll rises
       3806 reports ANI quot…',
        'Venezuelas Economy/Budget tied Oil like Louisianas Funny rightwing
       example of…',
        'reporting yesterday 1145 people died USA 73K new cases reported
        Serio…',
        'People relatives died file class action lawsuit Sinclair
        Broadcasting',
        '4249584 cases 148521 deaths reported United States date 1257 new
        cases 31…',
        ':  agree mask powerful accessories fight t…',
```

'miss isopropyl alcohol Ethanol hand sanitizer smells like poured
tequila hands
',
'GVMC sanitation staff carrying regular sanitation activities city
clean prevent spread…',
'Local government sales tax revenue dropped 27 percent yeartoyear
second quarter thanks shutdo…',
'Look Lil' Donny Micro Dick astounding achievement brilliance +
cognitive capabilities Somebody…',
'Check Angels beautiful mask hair bow designs',
'1 municipality reports new fatality compared yesterday 1 new
fatality located 0 fa…',
'current situation calls awareness stress facilitate families our…',
'economic measures governments taking respond Key actions taken 193
 countries
listed…',
'" everyone's fight covered major disasters like vi…',
'"I think government hugely overestimated danger disease" wonder
he'd brave en…',
'Ireland built COVID19 contract tracing app successful U states want
 use it…',
'Rakesh Sharma best business Coach India Website:',
'(year safe distancing requirements result o…',
'Learn new things period 2d 3d animations training ongoing Click
link join What…',
'literally head decisions safety government creating conf…',
'PR : father MrGKReddy tested positive 15/20 days s…',
'debates ensue reopening schools amid Young Leader says…',
'Companies protect workforce's physical mental health crisis use
smart technology…',
'Highest number new cases yesterday',
'CM Madhya Pradesh tested positive COVID Scindia tested',
'Bihar witnesses biggest singleday spike 2803 new cases',
'Lunch away mass tourists invaded quickly pandemic',
'Weve captured snapshot amazing VCSE response pandemic Salford want
 thank…',
'22000 nonprofits eventually close doors good result current
economic crisis acco…',
'Wells Fargo Tightens Purse Strings Ride Coronavirus Pandemic READ
 MORE:',
'informative interesting Lots think especially school age kids…',
'Veteran Affairs reports 252 new cases 13 new deaths bringing total
confirmed cases 35563 1980 to…',
'Irish home / commercial building model perfect breathing groung
virus trade treat…',

'residents - develop symptoms mild ones book test immediately at…',
'Actor father Mr GK Reddy tested positive couple weeks ago f…',
'Sounds like Plandemic second life Here's thread it's run debunks major cl…', 'know better Trump…',
'affected social anxiety wearing face mask amazing feel slightly hidden fro…',
'Vietnam's high level preparedness prevent Covid19 spread: Deputy PM',
'Republican Group Uses Trumps Words Searing SwingState Ad…',
'course absolutely diabolical government triple…',
'Learning fun interesting surprisingly easy Let's join WLF Nigeria Ambassadors promote learning a…',
'*Excellent data charts map CV progressing societies also…',
'coincidence influenza cases unusually down…',
'7813 new positive cases reported 24hours',
'(1) quick estimates Arizona post apex recovery rate: 7/03: 3849 7d avg Peak Confirmed Cases…',
'Trump hates looking like loser — right looks like biggest losers history: oped…',
'Private hospitals treatment rates General Ward 2300 HDU 3300 ICU 6500 ICU ventilator 11500 In…',
'epidemic Im giving $80 $1000 person looking able round i…',
'395 new cases 3 new deaths Uzbekistan [13:22 GMT]',
'"Until day medical vaccine food best vaccine chaos" David Beasley Am…',
'Yes "we're learning day" amp "we're listening science"…',
'Puerto Rico reports 573 new cases 10 new deaths bringing total confirmed cases 14540 201 total de…',
'1142 positive cases 2137 patients recovered/discharged/migrated 29 deaths Delhi today are…',
'Dying Dying huge difference usi…',
'design recovery plans humancentered mindset likely best position th…',
'unlike federal government largely depends funding fees collects citizenship gree…',
'Rutlands supporting testing',
'time beloved survivors a…',
'100% feeling good 0% symptoms Long ',
'New York retail',
'Delhi reported 1142 new cases 29 deaths today taking total number cases 129531 death tally 3806',
'Successful mattolson21 thinking: FEET ce…',

'tested positive says latest test result negative',
'1000 deaths fourth day row questions global supply chains ability
to…',
'1142 positive cases 2137 patients recovered/discharged/migrated 29
deaths today a…',
'Check latest article: Supply Chain Risks Thwart Successful COVID19
Vaccine…',
'PR : father MrGKReddy tested positive 15/20 days s…',
'7813 new cases 52 deaths 53681 samples tested 3208 discharged E…',
'uncle tested positive lost 3 family members wear masksocial
distance',
'Venrap Radio Playing Venrap Radio Stay safe love Download app
listen on…',
'Absolutely wouldnt standing 3rd position worldwide 12 lakh cas…',
'update 7/25/2020: *Nearly 40 countries report record singleday
increases infections *US records o…',
'insurance face possible economic burdens pandemic Tens millions
America…',
'Undercover inspectors deployed sure anti measures foll…',
'Hospital cleaners porters amp catering staff unsung heroes Tell pay
fair wage…',
'AP records 7813 new cases 24 hours taking tally 88671 52 deaths
wer…',
'Economies participating AsiaPacific Economic Cooperation (
reaffirmed prioritization economic…',
'WORLD SHORTENED BUSY PEOPLE',
'reaffirms economic recovery priorities movement essential goods',
'Cautious optimism Arizona plateauing Potential good news',
'Want Test It's Easier Wealthier Whiter Neighborhoods',
'realized trump amp win (God Forbid) finally stripped ACA…',
'Pleased share interesting results worldwide survey impact bariatric
surgery
Need rob…',
'Warning death trap: second spike has…',
'Cs term hadn't heard Good advice',
'| 1142 positive cases 2137 patients recovered/discharged/migrated
29 deaths today are…',
'Great yesterday',
'Yes real',
'pleased announce inclusion Virology Laboratory University Nigeria
Teaching
Hospital Enugu…',
'Great tips friends',
'football jersey count MD NFL opting care',
'Pandemic push N Carolina families home school | Miami Herald',

'Government (Tabla player) countries',
'everything's changing start playing version Sweet Caroline games
Th…',
'use sources guidance use CDC things fact check sour…',
'brilliant people awake accepting mainstream narrative',
'100% feeling good Sounds like soul song',
'Thanks \u2066 stressing personal responsibility limiting spread',
'development worrying continue monitored closely RKI exacerbati…',
'new beginning 6yrs ago wonderful new addition family Marley brought
muc…',
'vaccine face coverings stay feels like massive shift a…',
'Dr Ebrahim Janahi CEO Tamkeen confirmed supporting sectors affected
coronavirus pandemic reac…',
'183 coronavirus cases linked AlTaqwa College Melbourne commentary
nearly non existent Imagine th…',
'Yes hoody mask far followers dont real friends l…',
'Absolutely Exceptional situations need extraordinary solutions
fight battle usual measu…',
'Oh yes sure perfectly happy selfcertification',
'Hello Twitter Delighted share great innovators creators thought
leaders Visit website an…',
'Amazing took regret choice hey better late never…',
'took weeks limited population members test positive weve
followed…',
'Wow surprised impressed Mia understands rule (for waiting brother
Fraser p…',
'new way lifeuntil Nov 4 day election lockdown charade meani…',
'help fight important threats addressing issues sanitization
Pandemic A…',
'can't coffee shop (re)imagine outdoor space early morning
reading…',
'today 77 new cases 5 new death active cases 881',
'textile reusable mask proven ability inactivate new coronavirus
responsible fo…',
'mishandling pandemic response lead excess deaths premature death
losing game…',
'reports 7813 new cases Saturday tally goes 88671 death toll
increase…',
'Great reminder important nature – physically e…',
'1005 new cases 18 new deaths United States [13:16 GMT]',
'best teachers dont ignore chaos world outside classroom writes
Instead us…',
'Latest updates covid19 Qatar',
'enjoy',
'new cases Sorsogon include teenagers',

```
'intersect ways Right expected bring severe flood…',
'recovery home requires space prevent infection living roof 1 i…',
'Fiji active cases novel coronavirus disease/COVID19',
'Barcelona legend Xavi Hernandez tested positive head coach Qatars
Al Sadd said soc…',
'government Rajasthan poor pro Look kind slum sitting what…',
'University Kentucky working Somerset Community College create new
advancements mask filtrati…',
'safely restart nonemergent pandemic Join 910:30am EDT o…',
'love wicked \u200d \u200d \u200d \u200d ',
'Wearing Sunscreen year kills',
'Trump cared protecting adults school children pandemic protec…',
'thing distinct',
'60 countries reported disruptions childrens immunizations Decades
progress …',
'Running Balance SMART number recharge online care',
'online retail sector busier outbreak 10 easy steps improving y…',
'dont wanna miss',
'retail stores limiting hours deep cleaning limit wages',
'Andhra Pradesh 25/07/2020 New positive cases 7813 New discharges
3320 Ne…',
'started inviting worlds Phase III clinical trials Heres how…',
'Surely carried correct risk assessments allow children play
football are…',
'Amazing partnership period provide knowledge an…',
'Revovered corona B+ve plasma donor interested donate plasma DM
urgently required',
'institutions forging ahead oncampus inperson classes + activities
fall terms making…',
'know Dutch government legally mandate parents send children school
c…',
'7813 new cases 52 deaths 53681 samples tested 3208 discharged E…',
'Compared rest California labor force Asian Americans high school…',
'strength idolo',
'okay can't visit granny nursing home doors visit',
'Inline UKs goals promoting quality mentorship opportunities young
persons  w…',
'Wear Mask Stay safe safe',
'resuming movie shootings latest interview',
'know foreseen far better position lo…',
'Delhi Update Jul 25 New Cases 1142 Total 129531 July 43452 (3354%
total)
Deaths 29 Total 3806 Disch…',
'Catalonia Spain reports 1493 new cases ',
```

```
# Tokenization
for line in positive_tweet_list:
#     print(line)
    for token in line.split():
        print(token)
```

```
[64]: pos_tokens = [token for line in positive_tweet_list  for token in line.split()]
```

```
[66]: neg_tokens = [token for line in negative_tweet_list  for token in line.split()]
      neut_tokens = [token for line in neutral_tweet_list  for token in line.split()]
```

```
[67]: pos_tokens
```

```
[67]: ['Hey',
 'wouldnt',
 'sense',
 'players',
 'pay',
 'respects',
 'A…',
 'Change',
 'Work',
 'General',
 '(and',
 'recruiting',
 'specifically)',
 'via/',
 'Praying',
 'good',
 'health',
 'recovery',
 '', '_
 ',
 'safe',
 'safe',
 'commit',
 'ensure…',
 'Lets',
 'protect',
 'real',
 'numbers',
 'climbing',
 'fast',
 'Continent',
```

'Lets',
'n…',
'Second',
'wave',
'Flandersback',
'COVID',
'Update:',
'infection',
'rate',
'Florida',
'following',
'natural',
'curve',
'experts',
'predicted',
'initial',
'cu…',
'Good',
'Patriots',
'Volunteer',
'Election', 'Judge',
'Polls',
'open',
'without…',
'comprehensive',
'review',
'amp',
'Analysis:',
'key',
'ways',
'WASH',
'help',
'reduce',
'transmission…',
'crosses',
'2',
'lakh',
'mark',
'150055',
'ppp',
'recovered',
'far',
'positive',
'today',
'6988/total…',

```
'Actor',
'father',
'MrGKReddy',
'tested',
'positive',
'15/20',
'days',
's…',
'safe',
'place',
'visit',
'guests',
'said',
'hotel',
'meticulous',
'applying',
'hand',
'sanitation',
'als…',
'CEO',
'live',
'WSU', 'talking',
'"The',
'effects',
'Covid19',
'students"…',
'Actionables',
'healthy',
'recovery',
'understood',
'hadnt',
'missed',
'meeting',
'In…',
'released',
'new',
'podcast',
'episodes',
'week:',
'Technology',
'Platforms',
'Conduct',
'Telehealth',
'Visits',
'C…',
```

```
'revenue',
'dropped',
'27',
'percent',
'yeartoyear',
'second',
'quarter',
'thanks',
'shutdo…',
'Look',
'Lil'',
'Donny',
'Micro',
'Dick',
'astounding',
'achievement',
'brilliance',
'+',
'cognitive',
'capabilities',
'Somebody…',
'Check',
'Angels',
'beautiful',
'mask',
'hair',
'bow',
'designs',
'1',
'municipality',
'reports',
'new',
'fatality',
'compared',
'yesterday',
'1',
'new',
'fatality',
'located',
'0',
'fa…',
'current',
'situation',
'calls',
'awareness',
```

```
            …]

[68]:   # Get Most Commonest Keywords
        from collections import Counter

[73]:   def get_tokens(docx,num=30):
            word_tokens = Counter(docx)
            most_common = word_tokens.most_common(num)
            result = dict(most_common)
            return result

[74]:   get_tokens(pos_tokens)

[74]:   {'cases': 10310,
         'new': 9947,
         'amp': 5298,
         'positive': 3888,
         'people': 3655,
         'New': 3325,
         'deaths': 3292,
         'COVID19': 2745,
         'pandemic': 1929,
         'total': 1783,
         'reported': 1773,
         'safe': 1677,
         'today': 1661,
         'confirmed': 1522,
         'latest': 1468,
         'like': 1455,
         'time': 1455,
         '|': 1454,
         'good': 1449,
         'tested': 1398,
         'need': 1393,
         '24': 1380,
         'coronavirus': 1362,
         'day': 1331,
         'number': 1270,
         'right': 1266, 'reports': 1264,
         'India': 1230,
         'Total': 1227,
```

```
        '2020': 1222}
```
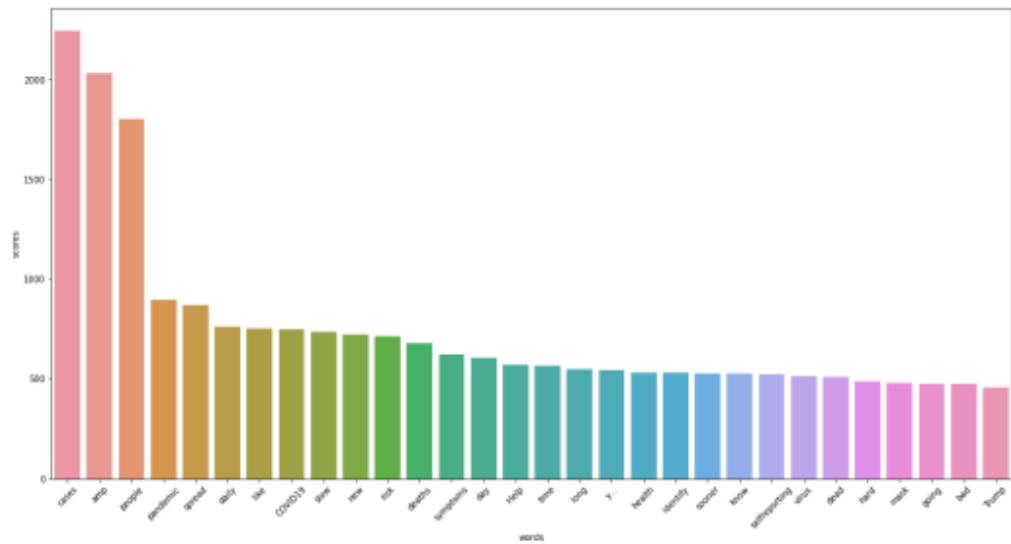
```
[75]: most_common_pos_words = get_tokens(pos_tokens)
      most_common_neg_words = get_tokens(neg_tokens)
      most_common_neut_words = get_tokens(neut_tokens)
```

```
[80]: # Plot with seaborn
      neg_df = pd.DataFrame(most_common_neg_words.items(),columns=['words','scores'])
```

```
[81]: neg_df
```
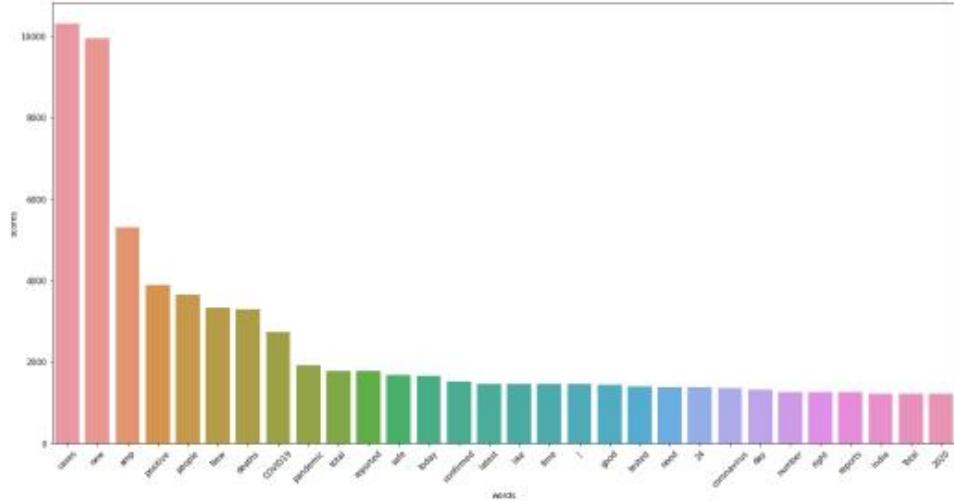
[81]:
|	words	scores
0	cases	2245
1	amp	2033
2	people	1801
3	pandemic	895
4	spread	870
5	daily	760
6	like	754
7	COVID19	748
8	slow	734
9	new	723
10	risk	713
11	deaths	679
12	symptoms	621
13	day	604
14	Help	573
15	time	568
16	long	550
17	y…	547
18	health	534
19	identify	533
20	sooner	529
21	know	527
22	selfreporting	522
23	virus	516
24	dead	510
25	hard	489
26	mask	478
27	going	475
28	bad	474
29	Trump	459
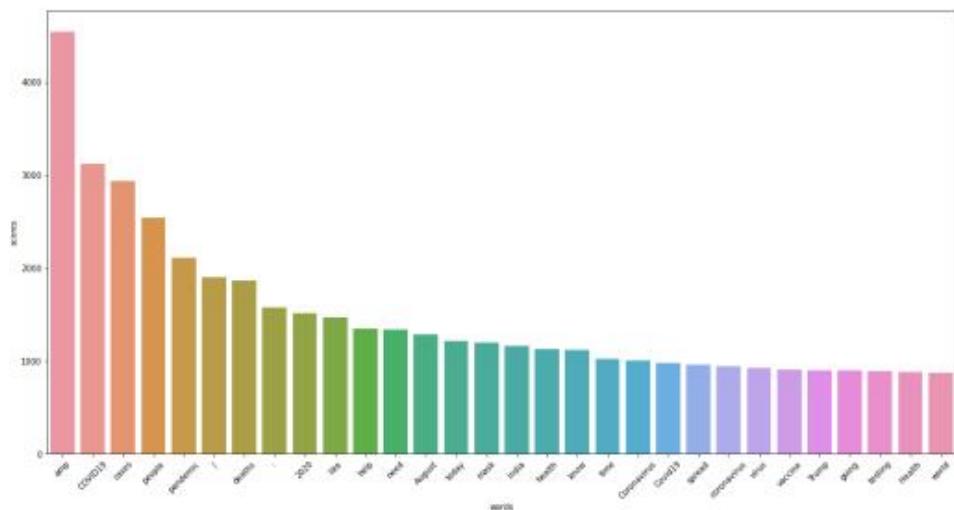
```
[83]: plt.figure(figsize=(20,10))
      sns.barplot(x='words',y='scores',data=neg_df)
      plt.xticks(rotation=45)
      plt.show()
```



```
[84]: # Plot with seaborn pos_df =
      pd.DataFrame(most_common_pos_words.items(),columns=['words','scores'
      ]) plt.figure(figsize=(20,10))
      sns.barplot(x='words',y='scores',data=p
      os_df) plt.xticks(rotation=45)
      plt.show()
```

```
[86]: # Plot with seaborn
      neut_df = pd.DataFrame(most_common_neut_words.
       → items(),columns=['words','scores'])
      plt.figure(figsize=(20,10))
      sns.barplot(x='words',y='scores',data=neut_df)
      plt.xticks(rotation=45)
      plt.show()
```

CHAPTER 9

CONCLUTION

9. <u>CONCLUTION</u>

The package was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

❖ Automation of the entire system improves the efficiency

❖ It provides a friendly graphical user interface which proves to be better when compared to the existing system.

❖ It gives appropriate access to the authorized users depending on their permissions.

❖ It effectively overcomes the delay in communications.

❖ Updating of information becomes so easier

❖ System security, data security and reliability are the striking features.

❖ The System has adequate scope for modification in future if it is necessary.

10. <u>**REFERENCE**</u>

[1] S. Raheja and A. Asthana, "Sentimental Analysis of Twitter Comments on Covid-19," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 704-708, doi: 10.1109/Confluence51648.2021.9377048.

[2] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund and J. Kim, "COVIDSenti: A Large-Scale Benchmark Twitter Data Set for COVID-19 Sentiment Analysis," in IEEE Transactions on Computational Social Systems, vol. 8, no. 4, pp. 1003-1015, Aug. 2021, doi: 10.1109/TCSS.2021.3051189.

[3] COVID-19 related sentiment analysis using state-of-the-art machine learning and deep learning techniques.

[4] Public Sentiment Analysis and Topic Modeling Regarding COVID-19's Three Waves of Total Lockdown: A Case Study on Movement Control Order in Malaysia.