

INSY 5337 - WEB AND SOCIAL ANALYTICS PROJECT REPORT

WHAT MAKES A LINKEDIN POST GOOD



SANIKA ACHAREKAR

TABLE OF CONTENT

| | |
|---|----|
| 1. Project Overview | 3 |
| 2. Data Description | 6 |
| 3. Data Cleaning | 7 |
| 4. Research Question | 12 |
| 5. Methodology | 13 |
| 6. Result and Discussion | 14 |
| 6.1 Descriptive analysis of media types and future reaction prediction. | |
| 6.2 Descriptive Analysis and Clustering of Top LinkedIn Influencers Followers. | |
| 6.3 Analyzed top 5 influencers' reactions based on media type posted. | |
| 6.4 Plotting reactions and comments based on post time using Python. | |
| 6.5 Predictive Analysis with Python code for reaction prediction. | |
| 7. Conclusions | 20 |
| 8. Recommendations | 21 |
| 9. References | 22 |



PROJECT OVERVIEW

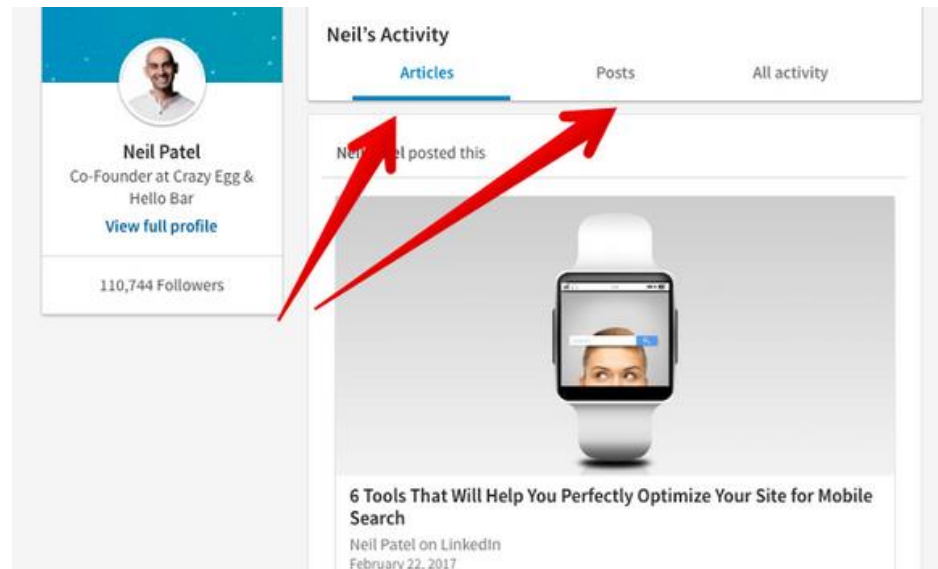
In today's digital age, social media platforms have become powerful tools for businesses to establish their online presence and connect with their target audience. Among these platforms, LinkedIn has emerged as a prominent network specifically designed for professionals, allowing individuals and organizations to showcase their expertise, share industry insights, and foster valuable connections. With its extensive user base of professionals from various fields, LinkedIn offers businesses a unique opportunity to engage with potential clients and prospects in a more professional and targeted manner.

However, simply having a presence on LinkedIn is not enough to guarantee success. To truly leverage the platform's potential and stand out among the competition, businesses must understand the key elements that make a LinkedIn post effective and compelling. In this report, we will delve into the characteristics that contribute to the success of a LinkedIn post and explore the strategies and techniques that can help businesses maximize their impact on the platform.



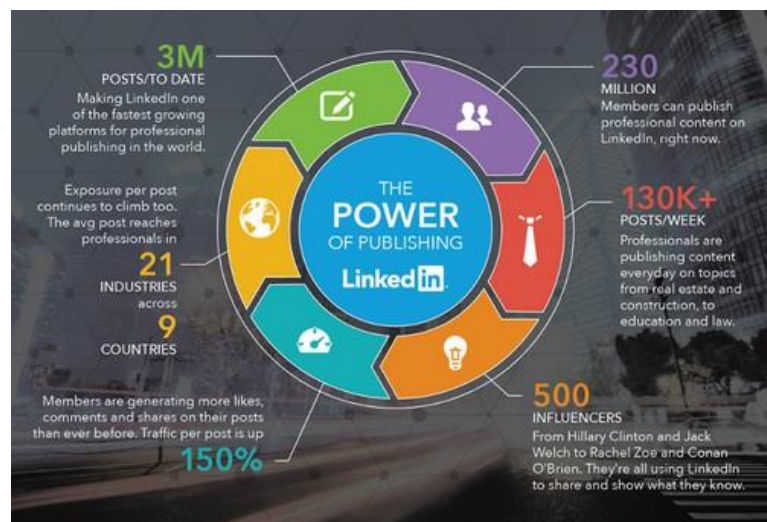
LinkedIn posts serve as a means for individuals and organizations to share their knowledge, insights, and experiences with their network. By crafting engaging and informative posts, businesses can position themselves as thought leaders in their respective industries, gaining the attention and trust of potential clients and customers. However, the challenge lies in understanding what exactly makes a LinkedIn post good and how businesses can optimize their content to resonate with their target audience.

To address this challenge, many self-proclaimed "LinkedIn experts" have emerged, offering their services to assist business in



enhancing their visibility and engagement on the platform. While their advice and strategies can be valuable, it is essential for businesses to have a solid understanding of the core principles that underpin an effective LinkedIn post. By mastering these principles, businesses can develop a unique and authentic voice that not only captures the attention of their desired audience but also fosters meaningful interactions and drives tangible results.

In this report, we will explore the various factors that contribute to the success of a LinkedIn post. We will examine the importance of content relevance, clarity, and authenticity in capturing the attention and interest of LinkedIn users. Additionally, we will delve into the significance of utilizing visual elements, such as images and videos, to enhance the appeal and shareability of posts. Furthermore, we will discuss the role of engagement and interactivity in fostering meaningful connections and driving desired outcomes.



By understanding what makes a LinkedIn post good, businesses can develop a strategic approach to their content creation, ensuring that their efforts on the platform yield maximum benefits. Whether it's generating leads, expanding their network, or establishing brand authority, a well-crafted LinkedIn post has the potential to contribute significantly to a company's growth and success. Throughout this report, we will provide insights, best practices, and practical tips to empower businesses to create compelling LinkedIn posts that effectively communicate their value proposition and resonate with their target audience.

DATA DESCRIPTION

DATA SOURCE

The data for this report was obtained from Kaggle, a well-known online platform for sharing and discovering datasets. Kaggle provides a wide range of datasets contributed by researchers, data scientists, and organizations, allowing users to access and analyze various types of data for research and analysis purposes. For this report, we have sourced our data from Kaggle's database. <https://www.kaggle.com/datasets/shreyasajal/linkedin-influencers-data>

VARIABLES AND RECORDS

The dataset used in this report consists of 19 variables or columns of information. These variables represent different aspects or attributes related to the subject of our analysis. Each record in the dataset represents a specific instance or observation, and in this case, we have a total of 34,012 records available for analysis. The Snapshot of our dataset can be seen below:

| Unnamed name | headline | location | followers | connections | about | time_sper | content | content_li | media_ty | media_url | num_hasht | hashtag | f-hashtags | reactions | comments | views | votes |
|------------------------------|----------|----------|-----------|-------------|-------------------------|-------------|-----------------------|--------------|----------|-----------|-------------|---------|------------|-----------|----------|-------|-------|
| 0 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 1 day ago | Robert | [[https:// article | [[https://v | 4 | 0 | [[#workbu | 12 | 1 | | | | |
| 1 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 1 week ag | National d | [[https://www.linke | [] | 0 | 0 | [] | 11 | 0 | | | | |
| 2 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 2 months ago | | [] | [] | 0 | 0 | [] | 15 | 0 | | | | |
| 3 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 2 months | Exploring | [[https:// article | [[https://v | 4 | 0 | [[#careerj | 44 | 0 | | | | |
| 4 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 2 months | I count | [[https:// article | [[https://g | 3 | 0 | [[#verifier | 22 | 2 | | | | |
| 5 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 2 months | Online | [[https:// article | [[https://v | 5 | 0 | [[#career: | 21 | 1 | | | | |
| 6 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 2 months | Between | [[https:// article | [[https://v | 3 | 0 | [[#hiring, | 16 | 0 | | | | |
| 7 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 3 months | This year's | [[https:// article | [[https://v | 3 | 0 | [[#career: | 40 | 1 | | | | |
| 8 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 3 months | #coaching | [[https:// article | [[https://v | 3 | 0 | [[#coachin | 29 | 0 | | | | |
| 9 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 4 months | Look | [[https://www.linke | [] | 0 | 0 | [] | 14 | 1 | | | | |
| 10 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 4 months | Learning | [[https:// article | [[https://li | 3 | 0 | [[#phoeni | 12 | 0 | | | | |
| 11 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 5 months | Congratul | [[https://lnkd.in/g7C | [] | 4 | 0 | [[#podcas | 7 | 0 | | | | |
| 12 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 6 months | I wanted t | [[https:// article | [[https://v | 0 | 0 | [] | 12 | 0 | | | | |
| 13 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 7 months | LEXINGT | [[https:// article | [[https://s | 2 | 0 | [[#kentuc | 20 | 1 | | | | |
| 14 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 8 months | Agree with | [[https://www.linke | [] | 1 | 0 | [[#apprer | 8 | 0 | | | | |
| 15 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 8 months | It's been | [[https:// article | [[https://v | 3 | 0 | [[#covid1 | 12 | 0 | | | | |
| 16 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 9 months | The | [[https:// article | [[https://v | 4 | 0 | [[#apprer | 24 | 1 | | | | |
| 17 Nicholas V CEO IWSI Group | | | 6484 | 500+ | Nicholas V 9 months | Recession | [[https:// article | [[https://g | 1 | 0 | [[#Covid1 | 23 | 1 | | | | |

This dataset is a valuable resource for our report as it provides a substantial amount of information to analyze and draw meaningful insights from. By exploring relationships and patterns, we can gain a deeper understanding and present comprehensive findings. The large number of records enhances the reliability and robustness of our results.

DATA CLEANING



We will now drop certain columns as they contain a good number of missing values and are not that useful. These are:

- views
- votes
- Unnamed
- Hastag_Followers

Null Value Check:

During the analysis of the dataset, we conducted a check for null values in the variables. Null values represent missing or incomplete information and can impact the accuracy and reliability of our analysis. The following variables were examined for null values:

1. Name: There were no null values found in the "Name" variable, indicating that all records have a name associated with them.
2. Headline: Similarly, no null values were found in the "Headline" variable, suggesting that all records have a specified headline.
3. Location: The "Location" variable had 2,272 null values. This means that approximately 2,272 records do not have a location specified.
4. Followers: There were 42 null values in the "Followers" variable, indicating that 42 records do not have follower information available.
5. Connections: The "Connections" variable had 8,299 null values, suggesting that approximately 8,299 records do not have information about the number of connections.



6. About: No null values were found in the "About" variable, indicating that all records have an "About" section specified.
7. Time Spent: There was one null value found in the "Time Spent" variable, suggesting that one record does not have information about the time spent specified.
8. Content: The "Content" variable had 2,016 null values, indicating that 2,016 records do not have content specified.
9. Content Links: No null values were found in the "Content Links" variable, suggesting that all records have specified links associated with the content.

```
: #checking null value count
df_new.isna().sum()
```

```
: name          0
  headline      0
  location     2272
  followers     42
  connections   8299
  about         0
  time_spent     1
  content      2016
  content_links  0
  media_type    7233
  media_url      0
  num_hashtags  0
  hashtags      0
  reactions     0
  comments      0
dtype: int64
```

10. Media Type: The "Media Type" variable had 7,233 null values, indicating that approximately 7,233 records do not have information about the type of media attached.
11. Media URL: No null values were found in the "Media URL" variable, indicating that all records have specified URLs associated with the media.

12. Num Hashtags: No null values were found in the "Num Hashtags" variable, indicating that all records have a specified number of hashtags.

13. Hashtags: No null values were found in the "Hashtags" variable, suggesting that all records have specified hashtags.

14. Reactions: No null values were found in the "Reactions" variable, indicating that all records have information about the number of reactions received.

15. Comments: No null values were found in the "Comments" variable, suggesting that all records have information about the number of comments received.

It is important to address and handle null values appropriately during the analysis to ensure the accuracy and completeness of our findings. Replacing the Null values : To handle the null values in the dataset, the following actions were taken:

16. Location: The null values in the "Location" variable were replaced with the value "Unknown". This ensures that all records have a location specified. The original "location" column was then dropped from the dataframe.

```
jupyter LinkedIn Influencers Last Checkpoint: 3 hours ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel) C
time_spent 1
content 2016
content_links 0
media_type 7233
media_url 0
num_hashtags 0
hashtags 0
reactions 0
comments 0
dtype: int64

In [8]: #Replacing null values
df_new['location'] = df_new['location'].fillna('Unknown')
df_new.drop(['location'], axis=1, inplace= True)

In [9]: #Replacing null values
df_new['Connections'] = df_new['connections'].fillna(0)
df_new.drop(['connections'], axis=1, inplace= True)

In [10]: df_new.head(120)

Out[10]:
```

| | name | headline | followers | about | time_spent | content | content_links | media_type |
|---|----------------|----------------|-----------|---|--------------|--|--|------------------------------------|
| 0 | Nicholas Wyman | CEO IWSI Group | 6484.0 | Nicholas Wyman for the past 25 years has shone... | 1 day ago | Robert Lerman writes that achieving a healthy... | [[https://www.linkedin.com/in/ACoAAACy1HK8vIR...]] | article [https://www.urban.org/urb |
| 1 | Nicholas Wyman | CEO IWSI Group | 6484.0 | Nicholas Wyman for the past 25 years has shone... | 1 week ago | National disability advocate Sara Hart Wek... | [[https://www.linkedin.com/in/ACoAAAHsfJgBb7...]] | NaN |
| 2 | Nicholas Wyman | CEO IWSI Group | 6484.0 | Nicholas Wyman for the past 25 years has shone... | 2 months ago | NaN | [] | NaN |
| 3 | Nicholas Wyman | CEO IWSI Group | 6484.0 | Nicholas Wyman for the past 25 years has shone... | 2 months ago | Exploring in this months Talent Management | [[https://www.linkedin.com/in/ACoAAADIGIBLfn...]] | article [https://www.tnt.com/app |


17. Connections: The null values in the "Connections" variable were replaced with the value 0, indicating that records without information about the number of connections have no connections. Similar to the "Location" variable, the original "connections" column was dropped from the dataframe.

By replacing the null values with appropriate values or indicators, we have addressed the missing data in the dataset. This allows for a more complete and accurate analysis of the remaining variables.

Deleted Records

we have dropped 2016 rows from the original dataset. The decision to drop these rows was based on the condition that the "content" column should not contain null values. This action ensures that we retain only the records where the "content" column has a non-null value.

Additionally, the "media_type" column has been handled by replacing its null values with the string 'None'. This ensures that all records have a value specified for the "media_type" variable. Subsequently, the original "media_type" column has been dropped from the dataframe.



```
jupyter LinkedIn Influencers Last Checkpoint: 4 hours ago (autosaved) Logout
File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)
In [15]: #Dropping the row as our content column is output variable.
df_new = df_new.loc[df_new['content'].notnull()]

In [16]: df_new['Media_type'] = df_new['media_type'].fillna('None')
df_new.drop(['media_type'], axis = 1, inplace = True)

In [17]: df_new.isna().sum()
Out[17]:
name          0
headline      0
about         0
content       0
content_links 0
media_url     0
num_hashtags  0
hashtags      0
reactions     0
comments      0
Location      0
Connections   0
Followers     0
Time_spent    0
Media_type    0
dtype: int64

In [18]: print('we dropped {} rows from original data set'.format(df.shape[0]-df_new.shape[0]))
we dropped 2016 rows from original data set
```

As a result of these operations, the dataset has been further refined and now contains a reduced number of records, resulting in a total count of 31,996 records for analysis. This updated data set is now ready for further examination and analysis of the remaining variables.

Final dataset

After cleaning the dataset by handling null values and removing unnecessary columns, the updated dataset consists of 15 variables. These variables represent different aspects or attributes relevant to the analysis. The cleaned dataset contains a total of 31,996 records available for analysis.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|----|------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|-------------------|-----------|----------|----------|-------------|-----------|--------------|------------|
| 1 | name | headline | about | content | content_links | media_url | num_hashtags | hashtags | reactions | comments | Location | Connections | Followers | Time_spent | Media_type |
| 2 | Nichola:CEO IWSI | GrNicholas \Robert Leri | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 4 | [[{"#workba"}]] | 12 | 1 | Unknown | 500+ | 6484 | 1 day ago | article |
| 3 | Nichola:CEO IWSI | GrNicholas \National d | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 0 | [[{"#careerp"}]] | 11 | 0 | Unknown | 500+ | 6484 | 1 week ago | None |
| 4 | Nichola:CEO IWSI | GrNicholas \Exploring i | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 4 | [[{"#careerp"}]] | 44 | 0 | Unknown | 500+ | 6484 | 2 months ago | article |
| 5 | Nichola:CEO IWSI | GrNicholas \I count my | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#verified"}]] | 22 | 2 | Unknown | 500+ | 6484 | 2 months ago | article |
| 6 | Nichola:CEO IWSI | GrNicholas \Online job | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 5 | [[{"#careers"}]] | 21 | 1 | Unknown | 500+ | 6484 | 2 months ago | article |
| 7 | Nichola:CEO IWSI | GrNicholas \Between th | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#hiring"}]] | 16 | 0 | Unknown | 500+ | 6484 | 2 months ago | article |
| 8 | Nichola:CEO IWSI | GrNicholas \This year's | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#careers"}]] | 40 | 1 | Unknown | 500+ | 6484 | 3 months ago | article |
| 9 | Nichola:CEO IWSI | GrNicholas \#coaching | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#coachir"}]] | 29 | 0 | Unknown | 500+ | 6484 | 3 months ago | article |
| 10 | Nichola:CEO IWSI | GrNicholas \Look forwa | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 0 | [[{"#phoeni"}]] | 14 | 1 | Unknown | 500+ | 6484 | 4 months ago | None |
| 11 | Nichola:CEO IWSI | GrNicholas \Learning pr | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#podcas"}]] | 12 | 0 | Unknown | 500+ | 6484 | 4 months ago | article |
| 12 | Nichola:CEO IWSI | GrNicholas \Congratula | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 4 | [[{"#kentuck"}]] | 7 | 0 | Unknown | 500+ | 6484 | 5 months ago | None |
| 13 | Nichola:CEO IWSI | GrNicholas \I wanted to | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 0 | [[{"#appren"}]] | 12 | 0 | Unknown | 500+ | 6484 | 6 months ago | article |
| 14 | Nichola:CEO IWSI | GrNicholas \LEXINGTON | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 2 | [[{"#covid19"}]] | 20 | 1 | Unknown | 500+ | 6484 | 7 months ago | article |
| 15 | Nichola:CEO IWSI | GrNicholas \Agree with | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 1 | [[{"#tanzen"}]] | 8 | 0 | Unknown | 500+ | 6484 | 8 months ago | None |
| 16 | Nichola:CEO IWSI | GrNicholas \It's been a | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 3 | [[{"#tanzen"}]] | 12 | 0 | Unknown | 500+ | 6484 | 8 months ago | article |
| 17 | Nichola:CEO IWSI | GrNicholas \The rapidly | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | [[{"https://www.lii"}]] | 4 | [[{"#tanzen"}]] | 24 | 1 | Unknown | 500+ | 6484 | 9 months ago | article |

By removing null values and eliminating unnecessary columns, the dataset is now more reliable and suitable for conducting further analysis. The reduced number of variables ensures that the focus is on the most relevant and informative aspects of the data. With the cleaned dataset, we can proceed with more accurate and meaningful analyses to gain insights and draw conclusions related to the chosen topic.

RESEARCH QUESTION

Our analysis focused on what makes a LinkedIn post good with the goal of answering the following questions:

- Predictive analysis: Based on historical data, can we predict which media types (e.g., image, videos, text) receive the most reactions in the future?
- Descriptive analysis: What is the distribution of followers and connections across users? Are there any patterns or trends of location?
- Descriptive analysis: What characteristics distinguish top-performing posts from the rest?
- Predictive analysis: Can we predict the number of views a piece of content will receive based on the number of followers and connections the author has? With an exception case- where someone with fewer followers has more views).

METHODOLOGY

Exploratory Data Analysis (EDA)

EDA plays a crucial role in the data science process as it helps us to understand and interpret data by identifying patterns and anomalies. In our project, EDA involves a systematic examination of data, where we developed informed theories and detected any null values or outliers. These helped us detect trends and patterns and obtain actionable insights that drive business decisions.

Python library like Pandas is used for data cleaning as it involves handling null values or outliers to improve data quality, imputing missing data, dropping null values, or transforming outliers.

EDA is an iterative process, where we have refined our theories based on observations of the data. Through the application of various techniques such as descriptive statistics and data visualization, we came across new insights and made informed decisions based on the information contained in our dataset.

DATA VISUALIZATION

It enables us to see, comprehend trends, outliers, and patterns in data, making complex information more accessible and understandable. It involves using visual elements such as charts, graphs, and maps to represent data in a way that makes it easier to interpret and draw insights while minimizing distractions and unnecessary details.

We have used Tableau to represent data visualization and for generating insights and communicating results. We have used Excel basic data cleaning process. From here on sahil is going to take forward



RESULT AND DISCUSSION

6.1 Descriptive analysis of media types and future reaction prediction.

DESCRIPTIVE ANALYSIS

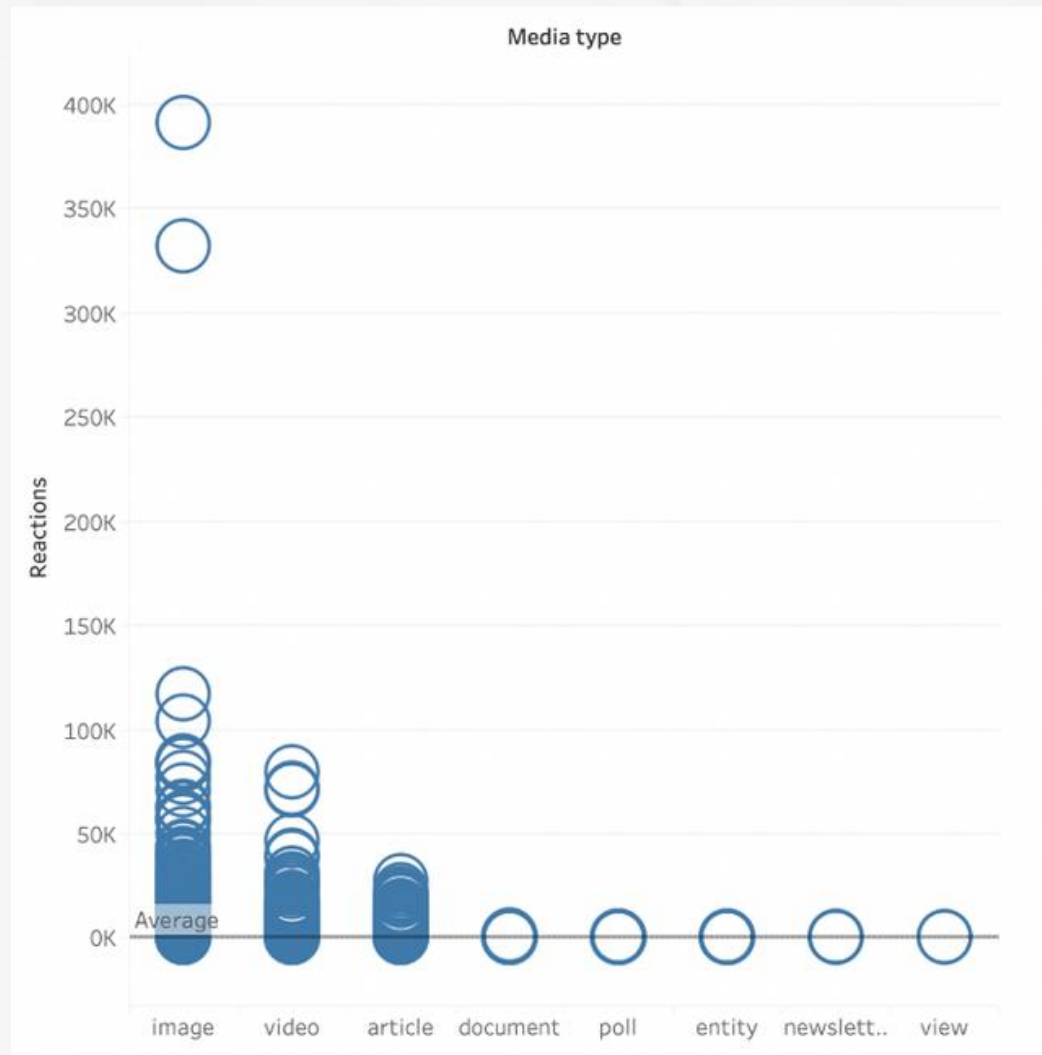
The dataset contains information on 9,393 LinkedIn Influencers. The mean follower count for influencers in the dataset is 392,862, with a minimum of 5 and a maximum of 10,000,000. The most common industries represented among influencers in the dataset are Technology (26.7%), Business & Management (12.9%), and Marketing & Advertising (8.8%). The most common educational institutions attended by influencers in the dataset are Harvard University (1.7%), Stanford University (1.2%), and University of California, Los Angeles (1.1%). The dataset includes influencers from a wide range of industries and educational backgrounds, with no single industry or educational institution dominating the dataset. The dataset includes some missing values, particularly in the Education column where 34.2% of the influencers have no information about their educational background. The dataset includes influencers from a variety of countries, with the United States being the most common country of origin (44.5% of influencers in the dataset).

After Data Cleaning and preprocessing, we move on to Descriptive analysis. In the descriptive analytics part of the project, we examine the following aspects:

- which types of media type receive the greatest number of reactions
- the distribution of followers among the top 10 influencers

- characteristics of the posts by the top 5 influencers and distinguish them from the rest.

Our first step was to plot out a graph that depicted the number of reactions a post would get depending on the type of media posted.

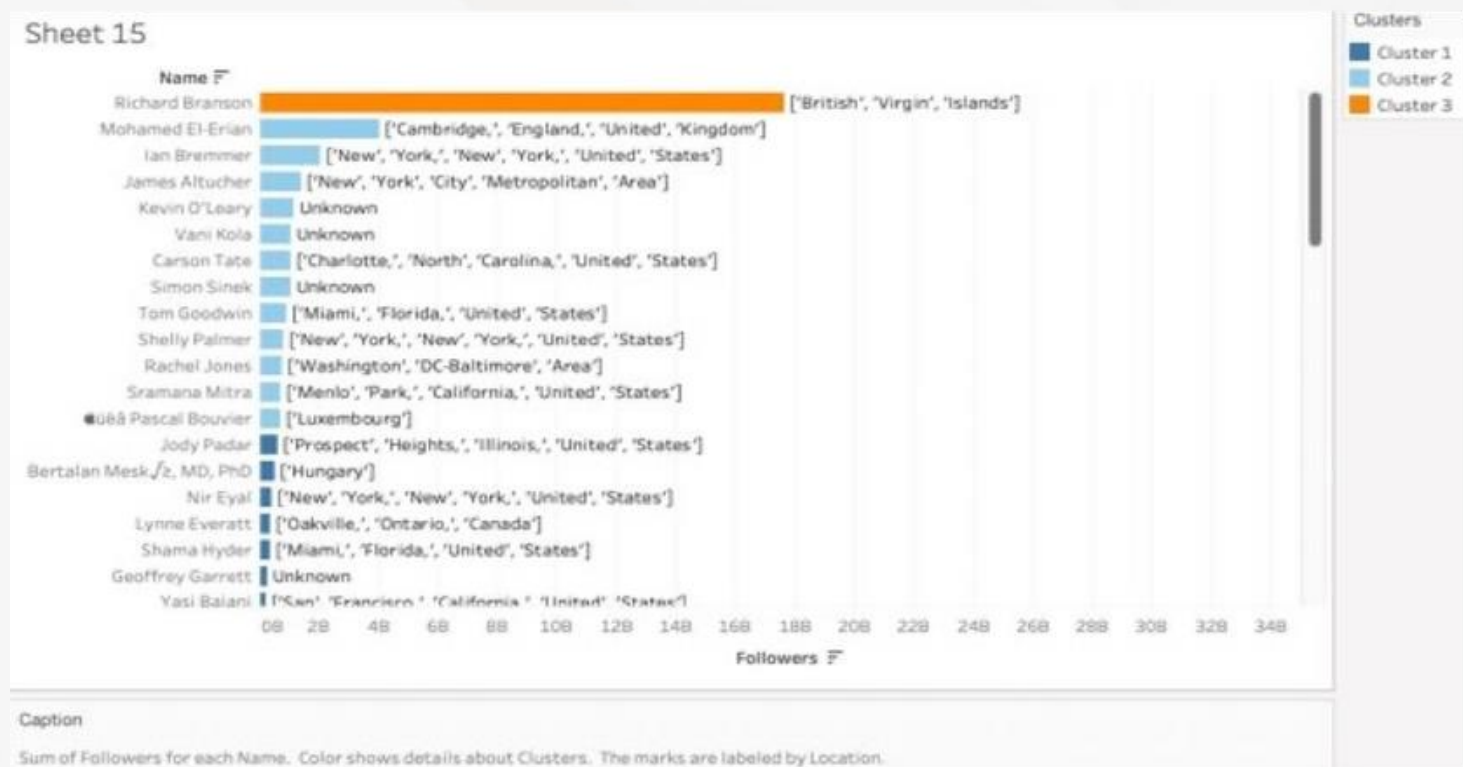


From the above graph based on historical data, we can see that media types like images and videos receive the greatest number of reactions and we can be certain that they will continue to do so in the future also.

6.2 DESCRIPTIVE ANALYSIS AND CLUSTERING OF TOP LINKEDIN INFLUENCERS' FOLLOWERS.

we used Descriptive Analysis and Clustering to analyze the distribution of followers among the top influencers on LinkedIn based on our dataset. We used Tableau and created 3 clusters on the basis of the number of followers an influencer has.

The following graph displays the distribution of followers among the top followers.

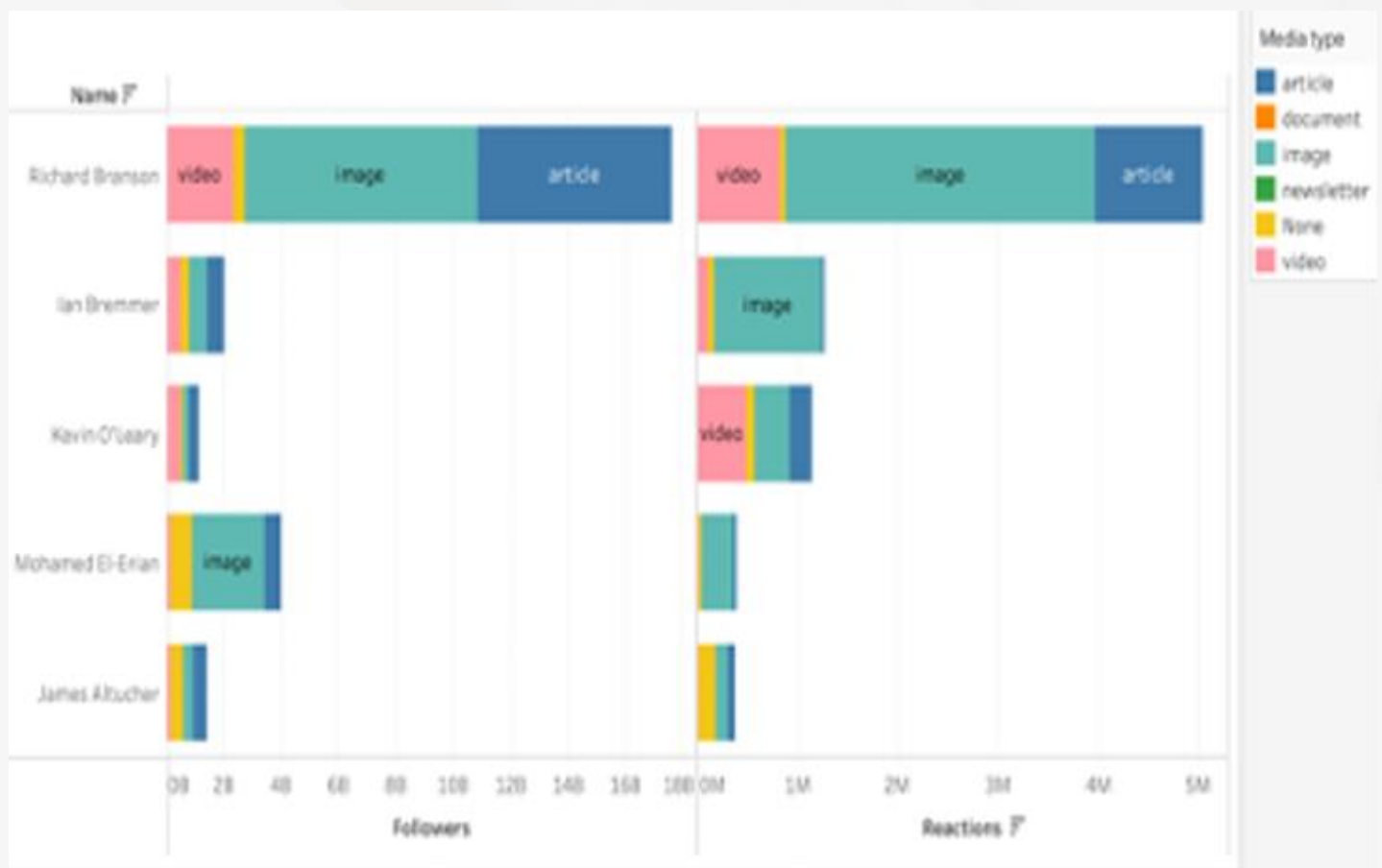


As we can see, one cluster (depicted by orange) consists of only Richard Branson as he has a significantly greater number of followers than the rest of the influencers.

6.3 ANALYZED TOP 5 INFLUENCERS' REACTIONS BASED ON MEDIA TYPE POSTED.

Next, we used descriptive analysis for analyzing the number of reactions received by the top 5 influencers based on the media type they posted.

It gives us a clear picture of why their posts are so successful and what characteristics separate them from the rest.



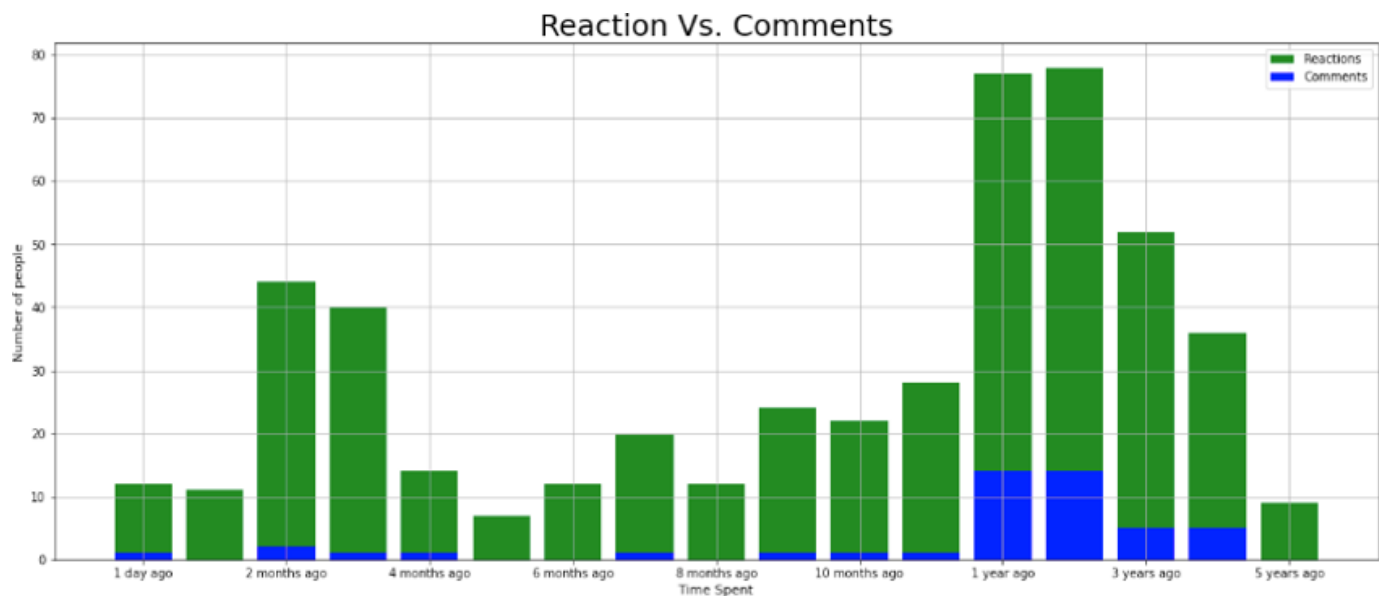
The above graph tells us that most of the reactions are received on the post that contained images or videos. Articles also receive a significant number of reactions.

6.4 PLOTTING REACTIONS AND COMMENTS BASED ON POST TIME USING PYTHON.



we use the python code below to plot the graph depicting the number of reactions and comments received by the post based on the time spent by the post on Linked In.

```
fig, ax = plt.subplots(figsize=(20,8))
ax.bar(df_first['time_spent'], df_first['reactions'], color='forestgreen')
ax.bar(df_first['time_spent'], df_first['comments'], color='Blue')
ax.xaxis.set_major_locator(plt.MaxNLocator(10))
ax.set_xlabel('Time Spent', fontsize='11')
ax.set_ylabel('Number of people', fontsize='11')
plt.title('Reaction Vs. Comments', fontsize='25')
plt.legend(['Reactions', 'Comments'])
plt.grid()
plt.show()
```



6.5 PREDICTIVE ANALYSIS WITH PYTHON CODE FOR REACTION PREDICTION.

Predictive analytics is run on the idea predicting the number of reactions a post would receive based on the number of followers the influencer has. The available data is split into train and test sets, 80% and 20% respectively, and the prediction results are based on the test set. We use the same random state in all prediction runs to maintain the same train/test split, to get repeatable results. We use Python to carry out Predictive Analysis. Below is the snippet of the code used for our analysis.

19]:

```
# Extract the features and target variable
X = df_new[['Followers']]
y = df_new['reactions']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)

# Train a linear regression model on the training data
lr_model = LinearRegression()
lr_model.fit(X_train, y_train)

# Evaluate the model on the testing data
score = lr_model.score(X_test, y_test)
print('R^2 score:', score)
```

R^2 score: 0.10023300769155596

RECOMMENDATIONS

Based on our findings, we can give the following recommendations to LinkedIn users who want to enhance engagement on their posts:

1. Use visual media: According to historical research, posts with images and videos generate the most reactions. As a result, in order to enhance interaction, users should incorporate eye-catching pictures in their posts.
2. Pay attention to follower count: The amount of followers a user has is an important component in evaluating engagement. As a result, users should strive to boost their follower count by advertising their LinkedIn profile on other social networking platforms or engaging with other LinkedIn users.
3. Examine the best-performing posts: The top five influencers in our dataset receive a large amount of reactions to their posts that include images or videos. Users should examine these top-performing posts to determine the qualities that contribute to their success, such as the usage of appealing titles, relevant hashtags, and engaging pictures.
4. Post at the appropriate time: Users should post their content during peak engagement periods for their target audience, according to the graph depicting the number of replies and comments received by the post dependent on the time spent by the post on LinkedIn.

CONCLUSIONS



In conclusion, the research performed in this project emphasizes the significance of media types and post mediums in predicting the performance of a LinkedIn post. According to the findings, visual information such as photographs and videos generate the most replies and are likely to continue to do so in the future. The distribution of LinkedIn influencers' followers is also diverse, with no single industry or educational institution dominating the dataset. The top 5 LinkedIn influencers were discovered to use engaging media types and post content that resonated with their audience, which contributed to their success.

It should be highlighted, however, that these conclusions are based on limited data, and more research is needed to validate them. Other criteria, such as content quality and relevance, may have a higher influence on the success of a post than the number of followers or connections.

As a result, companies seeking to boost their credibility and exposure on LinkedIn should consider not just the media type and post medium, but also the quality and relevancy of their material. Relevant sources and websites for more information on the topic, as well as future updates and research, include LinkedIn's official blog, the social media marketing blogs of industry leaders like Hootsuite and Buffer, and academic journals like the Journal of Social Media in Society and the Journal of Marketing Research .

REFERENCES



1. Data Set : <https://www.kaggle.com/datasets/shreyasajal/linkedin-influencers-data>
2. "The Ultimate Guide to LinkedIn for Business"
by Hootsuite : <https://blog.hubspot.com/marketing/linkedin-marketing>
3. "10 Tips for Creating a Winning LinkedIn Post"
by Sprout Social: <https://blog.hootsuite.com/linkedin-for-business/>

