Analysis by :
Praneeth Jajjara

ANALYSIS ON PUMA SOCIAL MEDIA BRANDING STRATEGIES

Twitter platform

Contents

Selection of case to research

Choosing the data collection method

Data collecting

Data Analysis

Conclusion Drawing

DATA COLLECTION METHOD:

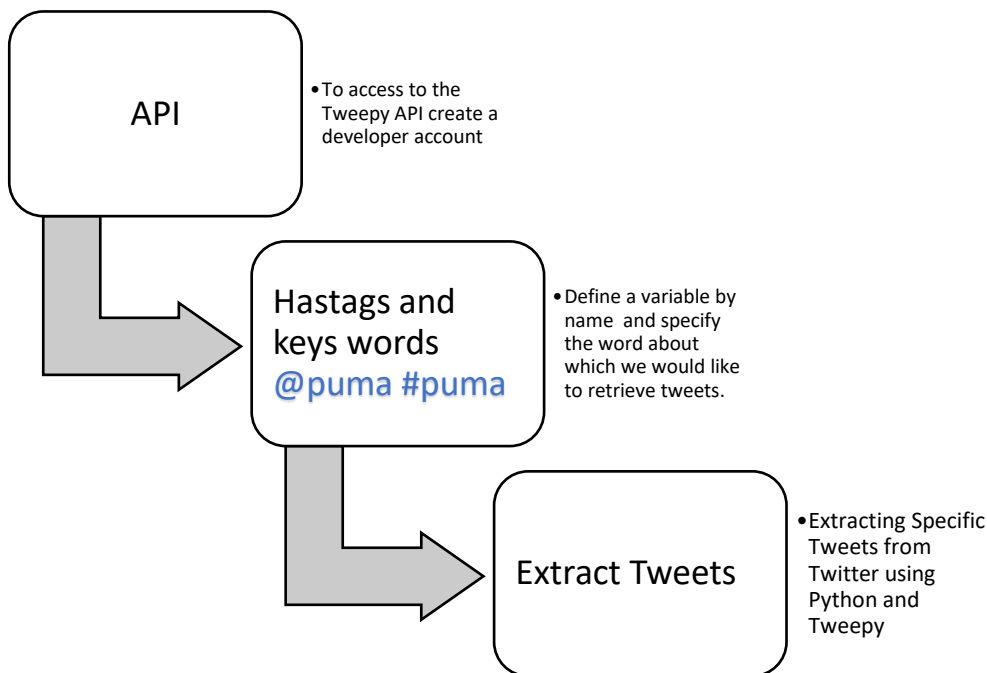Social Media Monitoring/online tracking

- Online forms are beneficial for gathering qualitative data about users, it provides insight into what content they're interested in and typically engage with.

- We have selected 'Twitter' to collect data for the brand we selected using either brand name or hashtags associated with the brand.

- One of the top benefits of Twitter for business is that it lets you share information quickly and start conversations with your target audience.

- When stored together and combined, multiple data types collected through different methods can give an informed picture of your subjects and help you make better business decisions.

- To collect the desired data, we are using web scraping.

- Few libraries like Selenium, tweepy, scrapy, requests, beautifulsoup are used to process of importing data from websites.

- The content related to the brand your desired to scrape is selected with respect to the hashtag help us in understanding market and users. Brands are usually product centric and listening to the reviews and sentiments about product, will help in creation of further products.

DATA COLLECTING:

Scraping

- Using Tweepy, we must first make sure that our Twitter credentials are ready. With that, we can connect Tweepy to our API key and begin scraping.

- Then use the four keys, here empty strings are mentioned to hide the authentication and you can use your own keys, to connect to the twitter API.

- Now, we will scrape public conversations from people on a specific trending @Puma, we have received many tweets including brand promotion taken up by Puma. Their retweets are considerable, and a comparative analysis will help to understand response whether campaign has received enough response. Which is limited in our case as we could not analyse all tweets.

- The ability to process this information has become important to deep dive into the Brand Perception. So, we need to be selective about the topic we are referring to.

- Twitter is a great source of consumer behavior data, with millions of tweets and trending topics that can be analysed for social listening and personalization engagement.

```
API
    • To access to the
      Tweepy API create a
      developer account

Hastags and
keys words
@puma #puma
    • Define a variable by
      name  and specify
      the word about
      which we would like
      to retrieve tweets.

Extract Tweets
    • Extracting Specific
      Tweets from
      Twitter using
      Python and
      Tweepy
```

Flowchart (1)

DATA CLEANING AND PREPROCESSING

In order to monitor mentions of our brand and related hashtags we need to do Data cleansing and pre-processing which allows for accurate, defensible data that generates reliable visualizations, models, and business decisions. As we identify and fixes errors, duplicates, and irrelevant data from a raw dataset

- Removing duplicate tweets.

- Detecting the language of tweets and extracting the desired language tweets (English).

- Removing repeated characters, additional accented characters from text, whitespaces, html tags, punctuations will help us to increase overall productivity and allow for the highest quality information in your decision-making

- Removing links from extracted tweets as we need to understand only the keywords, content and the topics surrounding the brand.

DATA ANALYSIS

Identify keywords, hashtags who play role in the brand campaign:

- The keyword extraction is done by finding the sub-phrases in a document that are the most like the document itself.

- KeyBERT is a easy-to-use keyword extraction technique that leverages BERT embeddings to create keywords and key phrases that are most similar to a document.

- To dive into insights of our brand related content in text we need to extract key information from a series of text data we have scraped from twitter.

- Keyword extraction is automated method for extracting relevant words from text data. This is text analysis method which involves automatically extracting the most important words.

- KeyBert extracts keyword with its respective highest cosine similarity score.

	Key_word	score
0	pumafootball	0.5438
1	pulisic	0.4713
2	pumasportswear	0.4379
3	nikefootball	0.4169
4	pumastyle	0.4090
5	pumasportstyle	0.3990
6	shoe	0.3976
7	footballer	0.3961
8	footballers	0.3929
9	puma	0.3876
10	adidas	0.3831

WORDCLOUD VISUALIZATION OF KEYWORDS

TOKENIZATION

- Tokenization helps in splitting a phrase, sentence, paragraph, or an entire text document into smaller units, such as individual words or terms. Each of these smaller units are called tokens.

- While trying to understand the topics surrounding our brand, we need understand the meaning of the text we collected, with help of tokenization we can easily be interpreted by analysing the words present in the text.

- Tokenization using NLTK, it is a library written in Python for symbolic and statistical Natural Language Processing.

- NLTK contains a module called tokenize () which further classifies into two sub-categories:

  Word tokenize:

  word tokenize () method to split a sentence into tokens or words.

  Sentences tokenize

  sent tokenize () method to split a document or paragraph into sentences.

- Tokenization is a critical step in the overall NLP pipeline. We cannot simply jump into the model building part without cleaning the text first.

- Once we can relate to the key conversations surrounding the brand then we get our content in front of a bigger audience, raise awareness about your brand, target a very specific group of people.

TOKENIZATING TEXT DATA

```python
[20]:   1  stopwords = nltk.corpus.stopwords.words('english')
        2
        3  def tokenize_only(text):
        4
        5      tokens = [word.lower()
        6               for sent in nltk.sent_tokenize(text)
        7               for word in nltk.word_tokenize(sent)
        8               if word.lower() not in stopwords]
        9
       10      filtered_tokens = []
       11
       12      for token in tokens:
       13          if re.search('[a-zA-Z]', token):
       14              filtered_tokens.append(token)
       15      return filtered_tokens
```

```python
[21]:   1  tokenize_df= []
        2  for i in range(len(df_scrape)):
        3      tokenize_df.extend(tokenize_only(df_scrape['cleaned_text'].iloc[i]))
```

Word Cloud Visualization

- To see a text-based visualization of a column we use word cloud visualization. The text height represents the scale.

- The word cloud allows you to identify the most relevant topics briefly.

- Graphical representation for series of keywords or tags associated with our brand puma organised by colour and size to form a figure as displayed below.

Puma WordCloud

Influencers most likely to play a role in the brand campaign:

Broadly, influencer marketing can be used for both awareness and persuasion campaigns. They help in growth strategies to reach new audiences. Influencers can amplify your brand if build in the right way as they can help you reach audience where your brand may not be able to, thus by giving more credibility to your brand when influencer talks about the same. Purchase decisions are driven by influencers.
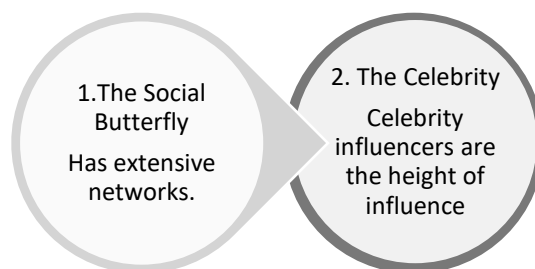
Selection of influencers:

- There are various groups, which speak about Puma products, Puma influences, competitor brands, aggressive influencers.

- Influencers can enhance audience retention. They can help us reach audience where the brand may not be able to.

<table>
<tr><td colspan="3">Characteristics of influencers</td></tr>
<tr><td>Activeness:

Number of times the influencer and their network "see" and "share"</td><td>Clout:

Number of connections and followers an influencer has</td><td>Talkativeness of receiver:

How often are messages being retweeted or shared</td></tr>
</table>

With reference to our brand PUMA the types of influencers are as follows:

1.The Social Butterfly
Has extensive networks.

2. The Celebrity
Celebrity influencers are the height of influence

Kareena Updates

Type of influencer - The Celebrity

- Puma India has launched a campaign featuring actor Kareena Kapoor Khan alongside the Propah Lady squad to promote its new collection.
- Thus, this sort of brand campaign helps in building a brand image, introducing a new product, increasing sales of a product already on the market, or even reducing the impact of negative news.
- The twitter account is unofficial page (not verified) but significantly have a good reach as the follower count is 445801 in number.

Size official

Type of influencer – The Social butterfly

- The authentic community sneaker store since.

- Customer Influence Value: Monetary Gain (product sale)

- The twitter account is official page (verified) and have a good reach as the follower count is 201443 in number.

Mizar

Type of influencer – The Social butterfly

- Mizar is a social media influencer.

- When products are promoted through influencer It gives more credibility to your brand, when influencer talks about the same.

- They can help you reach audience where your brand may not be able to.

- Purchase decisions are driven by influencers.

- The twitter account is official page (verified) and have a good reach as the follower count is 91154 in number.

Stats of top 3 influencers and their reach in twitter are as follows:

user_screen_name	retweet_count	likes	user_follower_count
KareenaUpdates	33	135	445801
sizeofficial	0.5	8.5	201443
mizar	0	0	91154

Identify the different groups here and the conversation surrounding the brand.

Topic Modelling using LDA.

- Topic modelling is a type of statistical modelling for discovering the abstract "topics" that occur in a collection of documents. Latent Dirichlet Allocation (LDA) is an example of topic model and is used to classify text in a document to a particular topic.

- In this approach the, each word in the corpus is treated as a feature in NLP which results in feature reduction and helps us to focus on the required relevant brand related topic material rather than diving through all the data's text.

- Thus, through this we can organize the collected data from twitter get insights understand and summarize large collections of textual information.

- Output of few among the different groups surrounding the brand are follows.

LIST 1: ['rt', 'antony', 'see', 'you', 'all', 'later', 'pumafootball', 'pumaultra']

LIST 2: ['antony', 'pumafootball', 'try', 'using', 'your', 'right', 'foot', 'sometimes']

LIST 3: ['rt', 'anerobeezy', 'antony', 'pumafootball', 'repeat', 'this', 'against', 'arsenals', 'on', 'Sunday']

OUTPUT with scores:

```
1  # Print the Keyword in the 10 topics
2  pprint(lda_model.print_topics())
3  doc_lda = lda_model[corpus_lda]
```

```
[(0,
  '0.033*"rt" + 0.010*"new" + 0.010*"day" + 0.010*"us" + 0.009*"join" + '
  '0.008*"action" + 0.008*"pumafootball" + 0.007*"boycottpuma" + 0.007*"year" '
  '+ 0.007*"global"'),
 (1,
  '0.038*"good" + 0.037*"poshmarkapp" + 0.037*"loving" + 0.037*"check" + '
  '0.037*"items" + 0.037*"poshmark" + 0.036*"share" + 0.036*"fashion" + '
  '0.036*"style" + 0.034*"shopmycloset"'),
 (2,
  '0.030*"pumafootball" + 0.015*"rt" + 0.015*"loving" + 0.015*"style" + '
  '0.015*"good" + 0.014*"fashion" + 0.014*"share" + 0.014*"poshmark" + '
  '0.014*"items" + 0.014*"poshmarkapp"'),
 (3,
  '0.019*"pumafootball" + 0.009*"go" + 0.008*"rt" + 0.007*"neymarjr" + '
  '0.006*"acmilan" + 0.006*"love" + 0.006*"day" + 0.005*"today" + 0.005*"stop" '
  '+ 0.004*"well"'),
 (4,
```

CONCLUSION DRAWING:

Recommendations for PUMA Social media Strategies.

- Consumer research:

  Social media is a great communication tool used between individuals – but they also use platforms like Twitter to interact with companies. Twitter is a great source of consumer behavior data, with millions of tweets and trending topics that can be analysed for social listening and personalization engagement.

- Benefits of Influencer Marketing:

  Influencers can amplify your brand if build in the right way as they can help you reach audience where your brand may not be able to, thus by giving more credibility to your brand when influencer talks about the same. Purchase decisions are driven by influencers.

- Brand Perception Using Twitter Sentiment Analysis

  The process of determining the emotional tone behind a series of words as positive, neutral, or negative is based on sentiment algorithms. This helps in product analysis to shape the minimum viable product by following brand keywords, customer needs, and segmented audience feedback to improve how the audience perceives the brand.

  For reference our sentiment analysis results are as follows :

Number of positive tweets	121
Number of negative tweets	7
Number of neutral tweets	796

  - As the number of negative tweets are only 5% of positive tweets feedback be taken and necessary actions can be implemented.

  - More positive tweets are predominantly coming from the promotional content and less from the users   which implies that brand should focus on existing and prospective customers to interact with the brand.