# Brief work Report on sentimental Analysis

Problem statement: To know what the people think about the Kashmir, Kashmir tourism, and  
Kashmiri cuisine (Kashmiri-wazwaan).

**Proposed Solution:** To know what the other people know about Kashmir, Kashmiri Tourism and Kashmiri wazwaan(cuisine) I have done Sentimental Analysis on Tweets. Twitter is common source for sentiment Analysis.

I have done sentiment analysis in three separate cases:

1. Sentiment Analysis of #kashmirtourism
2. Sentiment Analysis of #kashmir

3. Sentiment Analysis of #kashmiriwazwaan

#### **Data Collection**

Twitter, one of the world’s major social media platforms, with 187 million daily active users as of the third quarter of 2021[1], was chosen as the data source. Twitter is a common source of text for sentiment analysis [2][3]. I used the Python scraping library to access the twitter service and collect tourism–related tweets posted between 2022-03-03 and 2008-01-03. The search terms “kashmir tourism”, “kashmir” “Kashmiri wazwaan” were used to search target tweets.

**Keywords:** #kashmirtourism #Kashmir #kashmiriwazwaan**.**

|  |  |  |
| --- | --- | --- |
| Keywords | Total Number of tweets | **From Date**: 2022-03-03  **End Date:** 2008-01-01 |
| #Kashmir | 6383 |
| #Kashmiriwazwaan | 316 |
| #Kashmir tourism | 6336 |

**Methodology:**

**1.Data preprocessing**:

* In Data cleaning all Hashtags are removed from the Tweets column.
* After removing Hashtags, Punctuation is removed from all the tweets.
* Whitespaces is removed from the tweets by applying the function remove multiple white\_spaces from the tweets column.
* Urls and emojis are removed from the data by using the (neattext.remove\_urls) and (neattext.remove\_emojis) functions respectively.
* ‘\\n’ and all the digits are removed from the data by importing the regular expression.

**2. Sentimental Analysis**

**Polarity & Subjectivity**: The cleaned data thus obtained is subjected for polarity & subjectivity. For the polarity and subjectivity of the data, Text Blob library is imported by which all the data get labelled between the value of -1, to +1. Here -1 stands for negative tweets and 0 for neutral and 1 for positive tweets.

The Total number of all positive and negative and neutral tweets by using different keywords are given below in the table.

Total Number of tweets in different keywords.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **#kashmir** | **#kashmirwazwaan** | **#kashmiritourism** |
| **No. of positive tweets** | 1685 | 110 | 1476 |
| **No. of negative tweets** | 245 | 31 | 219 |
| **No. of neutral tweets** | 4454 | 224 | 4672 |

**3. Keyword extraction:** In Keyword extraction all the most used and most important words are extracted so that we can derive the important conclusions from the text.

**Remove stop words:** Stop words are the very common words like 'if', 'but', 'we', 'he', 'she', and 'they'. These words are removed, since these words do not change the semantics of a text. After removing stop words all the tweets are stored in the List data structure so that the further processing will be easy.

**Tokenization:** Tokenization is a way of separating a piece of text into smaller units called tokens. In Tokenization all the tweets are split into the individuals into small units called the tokens. Positive, Negative and Neutral tokens are stored in their respective variables.

After tokenization the most common words in the positive, negative and neutral tokens are counted so that we will know about the types of words that makes the most impact on the result.

**Tools used in Sentimental Analysis:**

* neattext
* nltk
* Regular expression
* Python
* Google colab
* For polarity and subjectivity, I used TextBlob library.

**Results and Discussions**

Total Number of tweets in different keywords.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **#kashmir** | **#kashmirwazwaan** | **#kashmiritourism** |
| **No. of positive tweets** | 1685 | 110 | 1476 |
| **No. of negative tweets** | 245 | 31 | 219 |
| **No. of neutral tweets** | 4454 | 224 | 4672 |

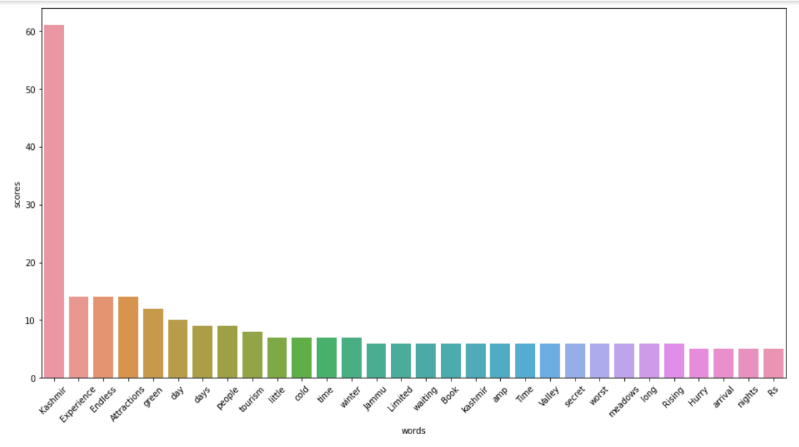
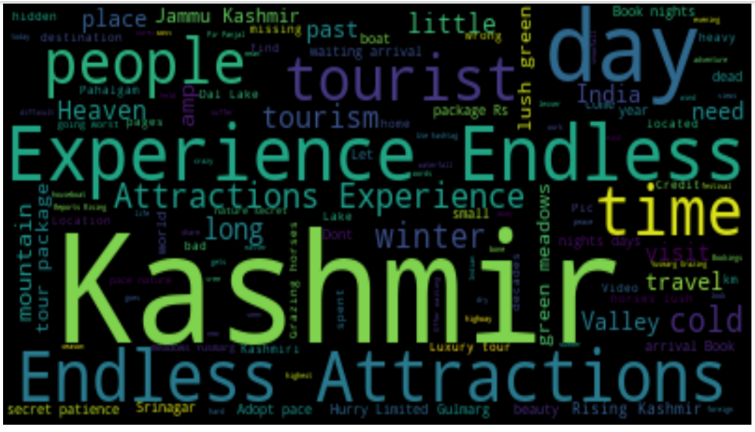
### 

## **Case 1: Sentimental Analysis of #kashmirtourism**

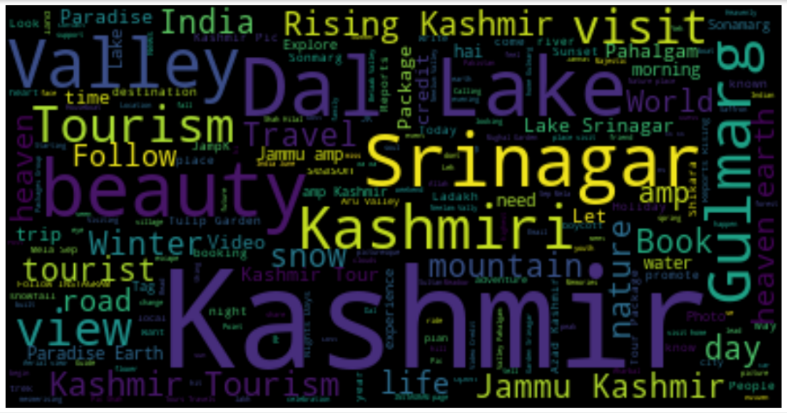
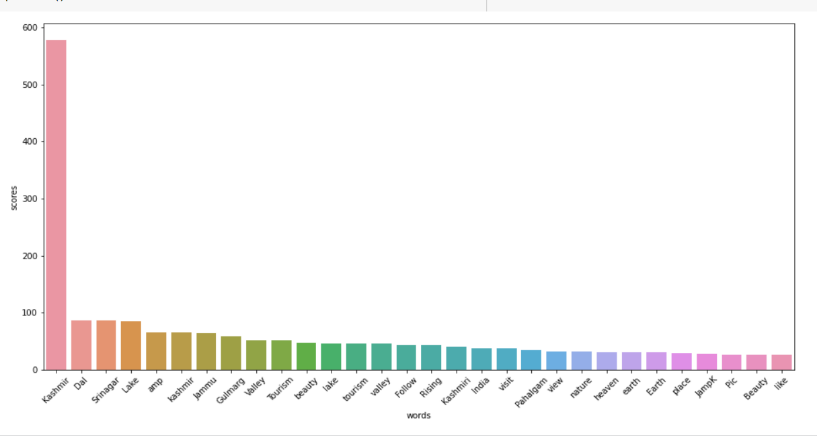
### Positive Tweet Results

Graph between most common words & number of the common words. Word Cloud of Positive tweets

### Negative Tweet Results:

** **Graph between most common words & number of the common words. Word Cloud of Negative tweets

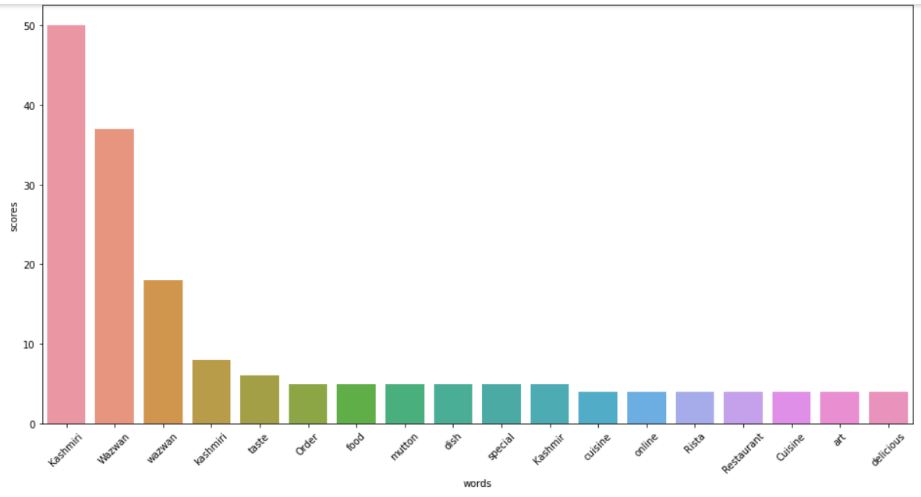
### Neutral Tweet Results:

****Graph between most common words & number of the common words. Word Cloud of Neutral tweet

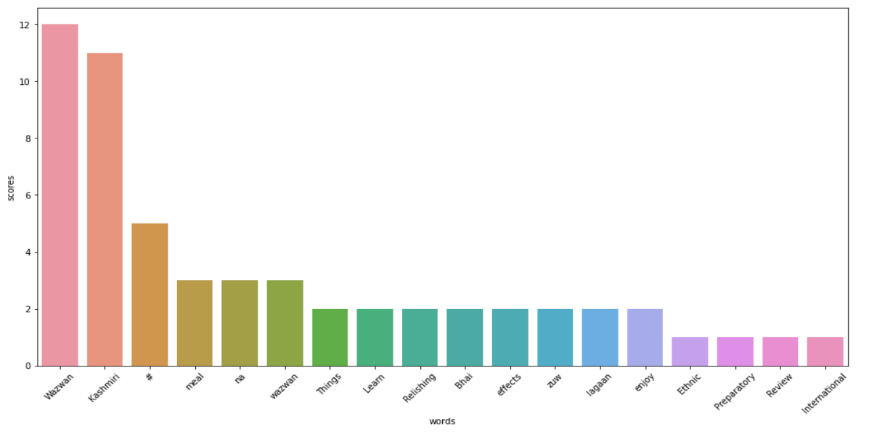
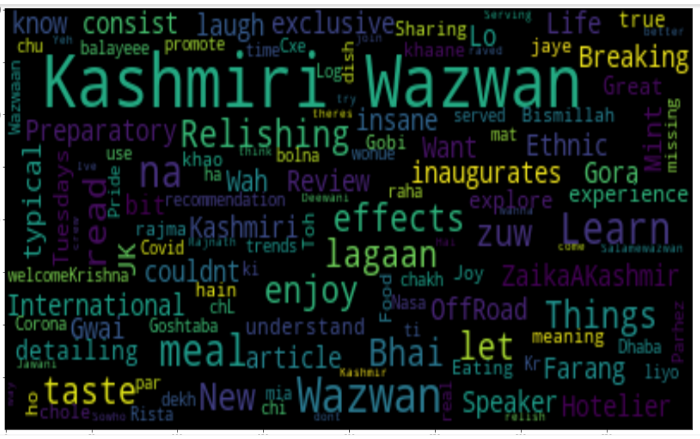
From the above graphs and wordclouds the word Kashmir includes the most positive and neutral comments.The word Kashmir show very less negative comments in the graphs as we have seen the graphs goes only upto the number 60 out of total number of tweets(6336).From the above analyis it is obvious that is kashmir is very beautiful as we have seen in the graph the word Beautiful goes upto the score 200 times.The kashmir Tourist Destinations like Gulmarg,Dal Lake and the Nature of the kashmir valleys has gain the attraction of the tourist.But kindly note this that most of the experience about the tourist destinations such as Pahalgam,Srinagar,Jammu,Dal Lake, Gulmarg are in neutral.

## **Case 2: Sentimental Analysis of #Kashmiriwazwaan**

### Positive Tweet Results of #kashmiriwazwaan:

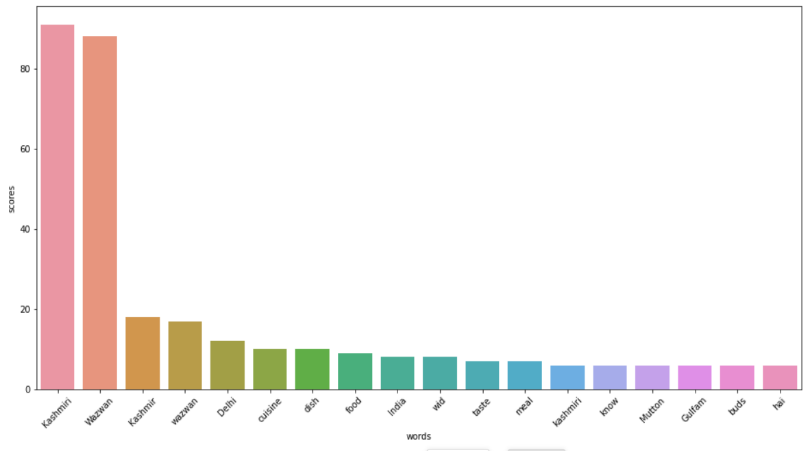
**** 

Graph between most common words & number of the common words. Word Cloud of Positive Tweets #Kashmiri wazwaan

Negative Tweet Results of #kashimiriwazwaan ****

Graph between most common words & number of the common words. Word Cloud of Negative Tweets #kashmiriWazwaan

Neutral Tweet Result of #kashmiriwazwaan

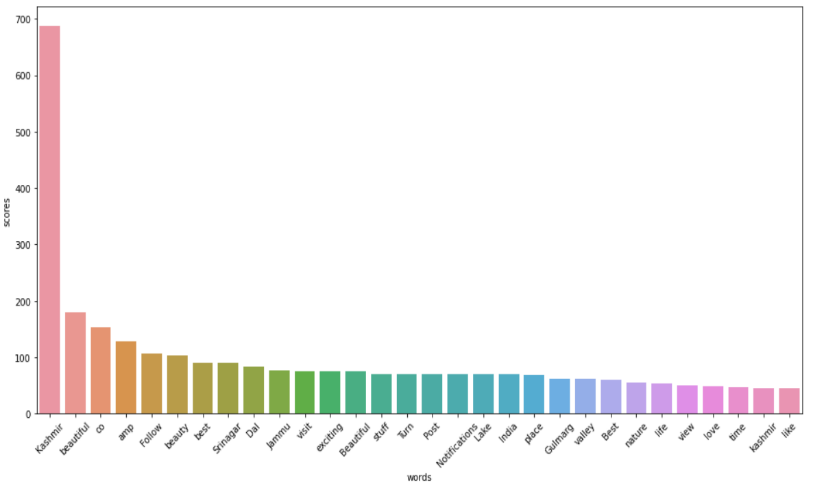
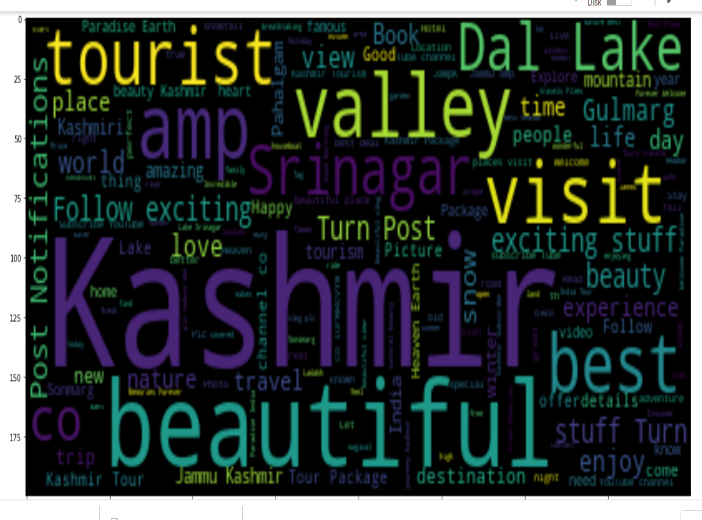
**** ****

Graph between most common words & number of the common words. Word Cloud of Neutral Tweets # Kahmiri Wazwaan

During the Sentimental Analyis of Kashmiri-Wazwan there were total number of 316 tweets.From the above graphs and word cloud It has been seen that the most of the tweets and are positive and neutral.There are very less number of negative tweets in the kashmir Wazwaan. From the above the discussion it has to be seem that there are only 9% of Negative tweets in kashmri wazwaan. This shows that the most people likes the kashmir wazwaan.and it obvious the Kashmiri Wazwaan is famous,the kashmri wazwaan includes the tasty dishes like Rista, Kabbabs,chicken,Rogan Joosh,Gushtaab,etc.

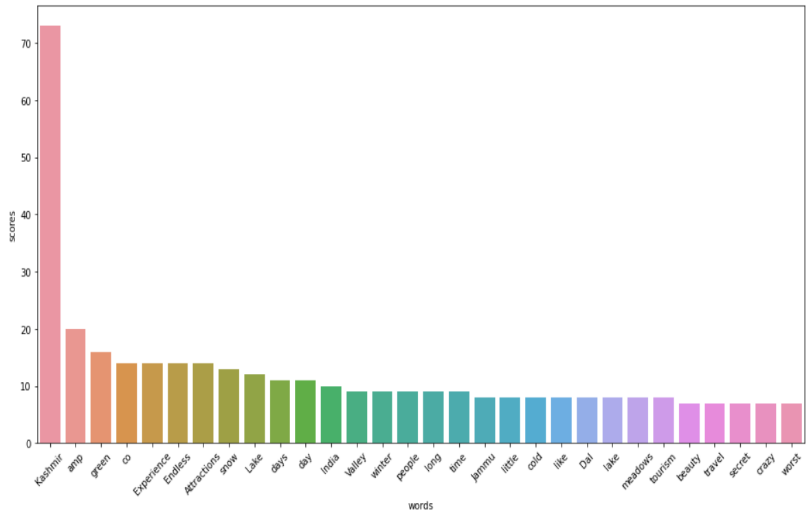
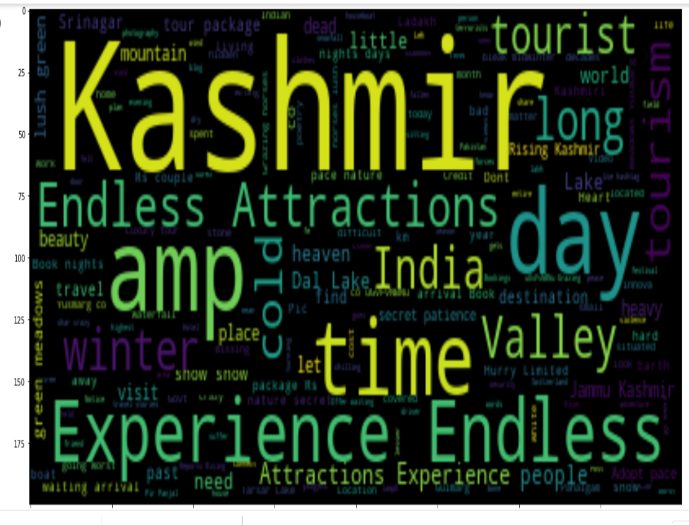
## **Case 3: Sentimental Analysis of #kashmir**

### **Positive Tweet Results of #kashmir:**

**** ****

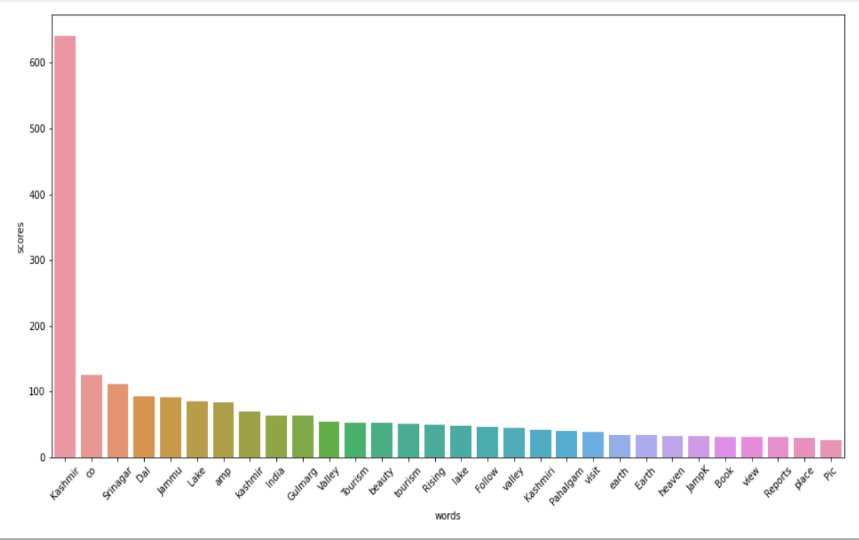
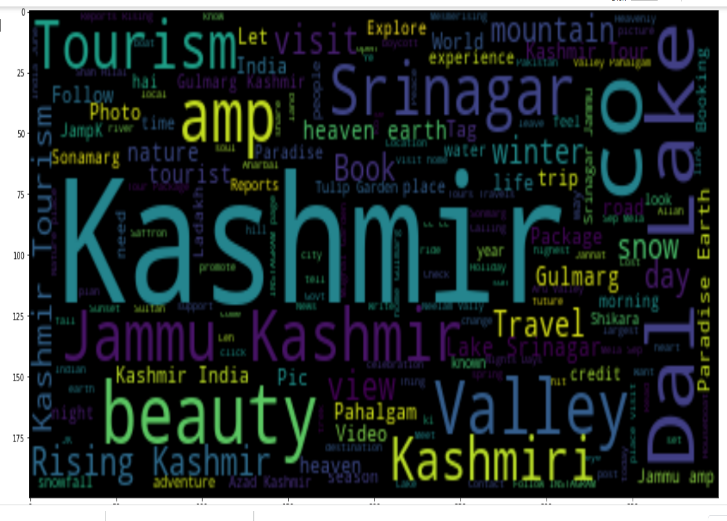
Graph between most common words & number of the common words. Word Cloud of Positive Tweets #Kashmiri

### Negative Tweet Results of #kashimir:

** **

Graph between most common words & number of the common words. Word Cloud of Negative Tweets #kashmiri

### Neutral Tweet Results of #kashmiriWazwaan:

 ****

Graph between most common words & number of the common words. Word Cloud of Neutral Tweets # Kahmiri

In the setimental anlysis of the Kashmir it has been seen from above graphs and wordclouds that the word Kashmir includes most positive and neutral comments.The word Kashmir show very less negative comments in the graphs as we have seen the graphs goes only upto the number 60 out of total number of tweets(6336).In other words there are only 4.0% of the Negative tweets.From the above analyis it is obvious that is kashmir is very beautiful as we have seen in the graph the word Beautiful goes upto the score 180 times.From the above analysis it is has to be seen that Kashmir is the place of peace.However in winters there are some negative tweets which shows that some people feel very bad during the chilly winters.

References:

1. Investor fact sheet. Twitter. 2020.   URL: <https://s22.q4cdn.com/826641620/files/doc_financials/2019/q4/Q4_19_InvestorFactSheet.pdf> [accessed 2021-05-06]
2. Khanna P. An approach to opinion mining from Twitter data using R. IJARCS 2017 Aug 30;8(8):252-256. [[CrossRef](https://dx.doi.org/10.26483/ijarcs.v8i8.4716" \t "_blank)]
3. Zimbra D, Abbasi A, Zeng D, Chen H. The State-of-the-Art in Twitter Sentiment Analysis. ACM Trans Manage Inf Syst 2018 Sep 05;9(2):1-29. [[CrossRef](https://dx.doi.org/10.1145/3185045" \t "_blank)]