**Amazon Music Industry Reviews Assessment**

***Group 15 (FNU Ankita, Monika Soni, Shrooq Feda)***

**Business Context**

Music Industry is blooming with new products with advancement in technology and increased sync with digital platforms like mobile and other computing devices. Customers buy Music Products from Amazon.com across the US and submit reviews and feedback based on their experience. In this scenario, companies are more open towards receiving valuable feedback on their products to compete among each other. It is therefore imperative to study product reviews using customer sentiment analyses and topic modeling to better strategize new product launch and changes in sales and marketing strategy.

**Input Data and Sources**

Amazon Music Product Reviews data was sourced from University of California, San Diego academic study resource server (Link: <http://jmcauley.ucsd.edu/data/amazon/>)

Data consists of the following relevant columns:

* Data ID – Review Identifier
* Reviewer ID – Customer Identifier
* Reviewer Name
* Product Review
* Year/Month/Date of Review
* Customer ZIP Code

**Analysis Approach**

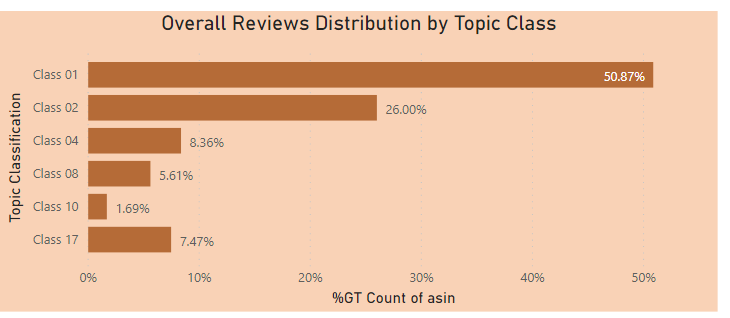
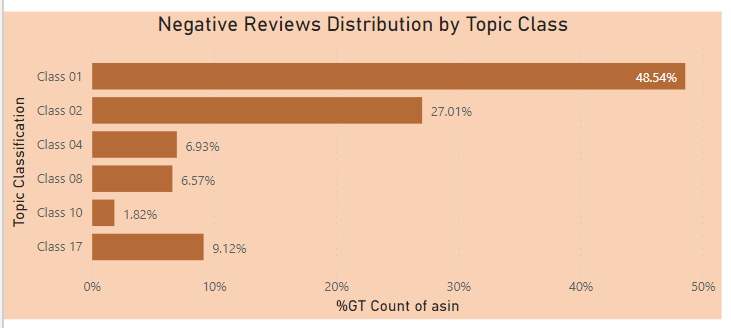
1. Source and Load Data: JSON file was sourced from UCSD website and loaded to RStudio
2. Pre-Process Data: Reviews Column was cleaned to remove
   1. Unwanted characters, symbols or numbers like @, # etc.
   2. Punctuation in text
   3. Unwanted Space
3. Perform Sentiment Analysis:
   1. Sentiment Analysis was performed to calculate sentiment score or polarity based on text reviews
   2. Sentiment scores were used to classify reviews into positive, negative or neutral
4. Perform Topic Modeling:
   1. Topic Modeling was conducted on R post creation of DTM (Document Term Matrix) and running LDA (Latent Dirichlet Allocation) algorithm
   2. Best model was selected based on coherence score
5. Text Word Cloud

**Findings & Recommendations 1/3**

On comparison of sentiment scores (or values) with Topic classes associated with each review, it was found that ‘Topic #17’ has the highest contribution change in negative sentiments compared to average reviews



**Average Reviews Distribution Negative Reviews Distribution**

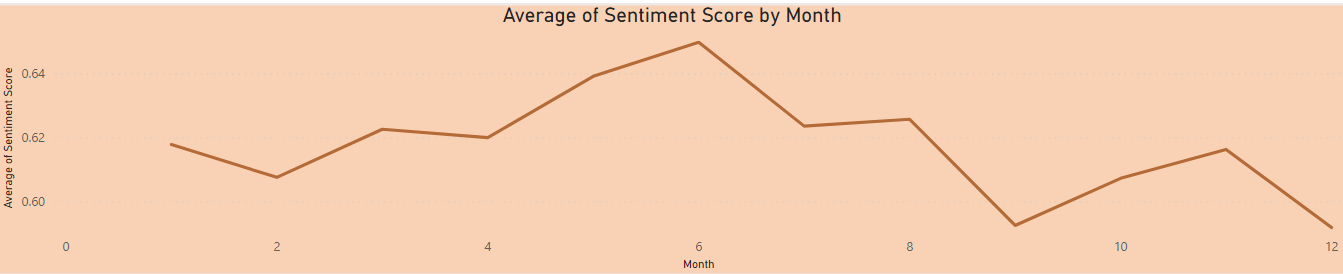
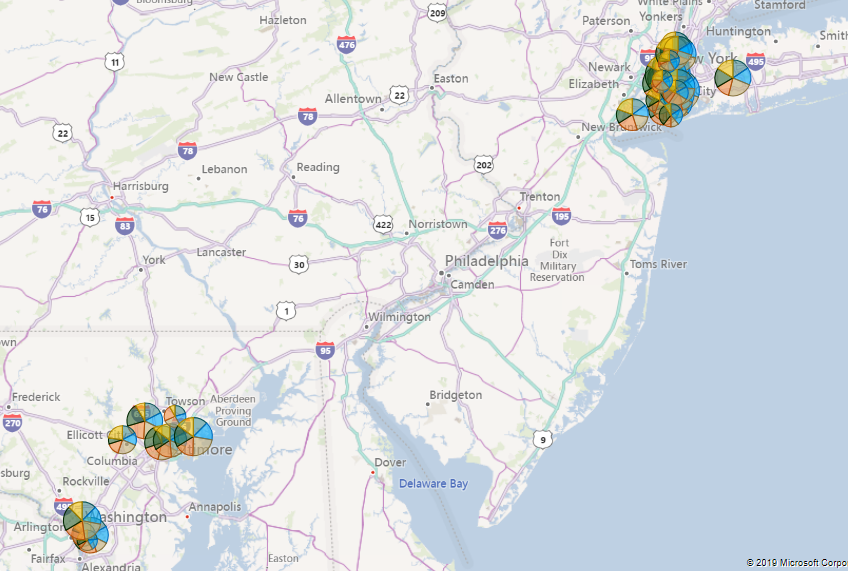
** **

**Recommendation:** Company should pay more emphasis on the following in order of priority:

1. Better Pricing w.r.t. competitors
2. Better Product Quality: Avoid usage of cheap plastic
3. Improve Robustness and ensure longer life of products
4. Resolve customer issues and complaints quickly

**Findings & Recommendations 2/3**

On an average, sentiments recorded by customers are highest in Mid of year compared to other months. This distribution follows a distribution like normal distribution with peak in June month, and with highest sentiment scores in the eastern US

**Recommendation:**

1. Company should pay more attention to reviews or product sales towards end of the year. This might be due to seasonality factor, and customers’ drop-in trust/usage of products towards end of the year.
2. Company should also pay more attention to sales coming in from eastern region of the US esp. NY and MD states as higher customer sensitivity towards the products are observed in these states

**Findings & Recommendations 3/3**

**Text Word Cloud** was used to identify highest frequency of words in product reviews and also the most popular products from the company. It can be observed that the overall product review is positive with highest emphasis on words like ‘good’, ‘sound’, ‘well’, ‘quality’, ‘great’ etc. It is also observed that guitar is among the most popular choice of products



**Recommendation**:

As the overall product reviews is positive through Amazon, company seems to be on the right track. It should however try to maintain guitar sales through quality consistency, and emphasis on sound quality. As next step, it is also recommended to evaluate other products performance and estimate customers’ feedback on those.

**Conclusion**

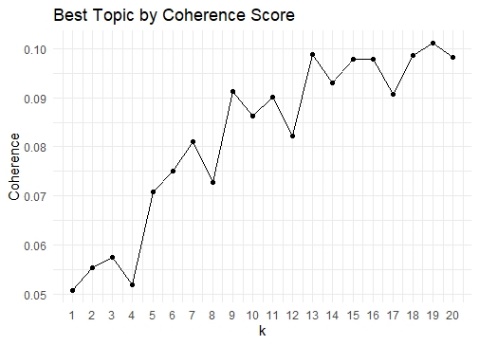
Overall, the music products sold through Amazon.com are positively reviewed by customers. There are however, certain problems reported which impacts the reviews negatively. Some of them are pricing, quality, customer service etc. It was recommended to maintain sales of guitar, which is one among the highest rated products and also focus on increasing other products sales by paying more attention towards above mentioned points.

Link to dashboard: [Link](https://app.powerbi.com/view?r=eyJrIjoiYzc2YWU2YzgtZDYwMi00ZGY3LWFlNWMtMTg5YzMyZDhlY2RiIiwidCI6ImVhODczMzkwLThjMWMtNDIzMS1hNzk5LTZiNWEwMjM1YjJlNiIsImMiOjN9)

**Appendix**

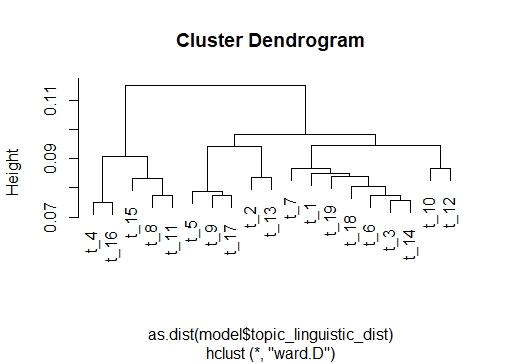
**Topic Modeling Details:**

**Model Selection:** Best LDA model was selected based on Coherence at multiple values of K. It was observed that K value of 19 gives the highest Coherence. Accordingly, 19 different topic classes were formed**.**



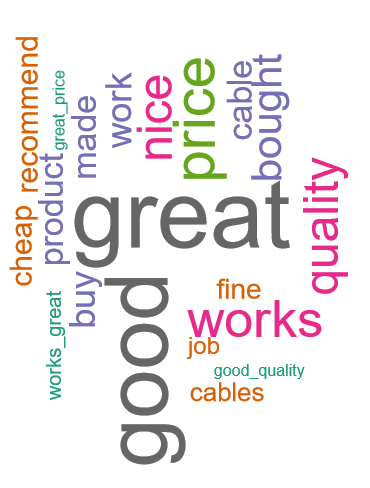
**Dendrogram test,** which uses Hellinger distance was used to collapse similar topic classes to a set of 6 different classes:

1. T4 & T6: **T4 Group**
2. T15 & T8 & T11: **T08 Group**
3. T5 & T9 & T17: **T17 Group**
4. T7 & T1 & T19 & T18 & 16 & T3 & T14: **T01 Group**
5. T2 & T3: **T02 Group**
6. T10 & T12: **T10 Group**



**Topic Classes** were created based on grouping of similar words and frequency

Topic Class #1: **Positive Topic Class #17: Negative**

**Sentiment Analysis:**

What is Sentiment Analysis?

Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations.

Why Sentiment Analysis?

Sentiment analysis can be an essential part of market research and consumer service approach. Not only can we see what people think of our own products or services, we can see what they think about our competitors too.  The overall customer experience of your users can be revealed quickly with sentiment analysis. It provides you the ability to quickly understand your product and react accordingly.

By Performing sentiment analysis on the reviews of customers buying the company’s product through amazon, will provide the company the answers to the following questions which can be used to do a detailed analysis on its market and provide better product and service to its customers.

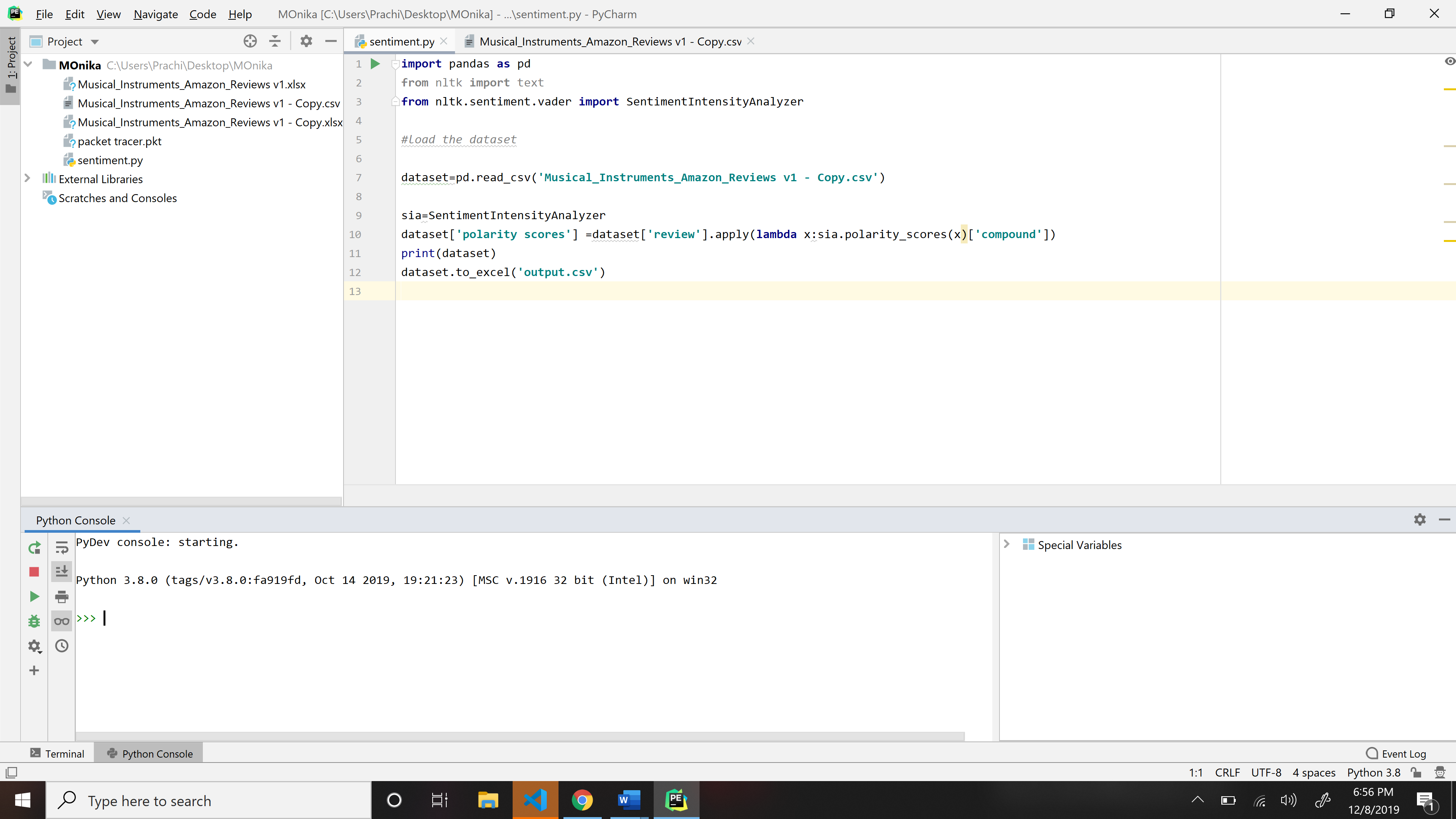
* What is the overall sentiment on the reviews? Is it positive, negative or neutral?
* Is amazon a good source to sell its product?
* During which period the company got negative/positive/neutral reviews and why?
* In which region the customers are more sensitive and why?
* What feature of a product got popular? Etc.

This analysis will help the company to develop effective **marketing and business development strategy.**

**Analysis**

To analyze the dataset and get sentiment/polarity score of each review we used python script using Vader which assign scores to each review based on the words used by the customer in reviews.

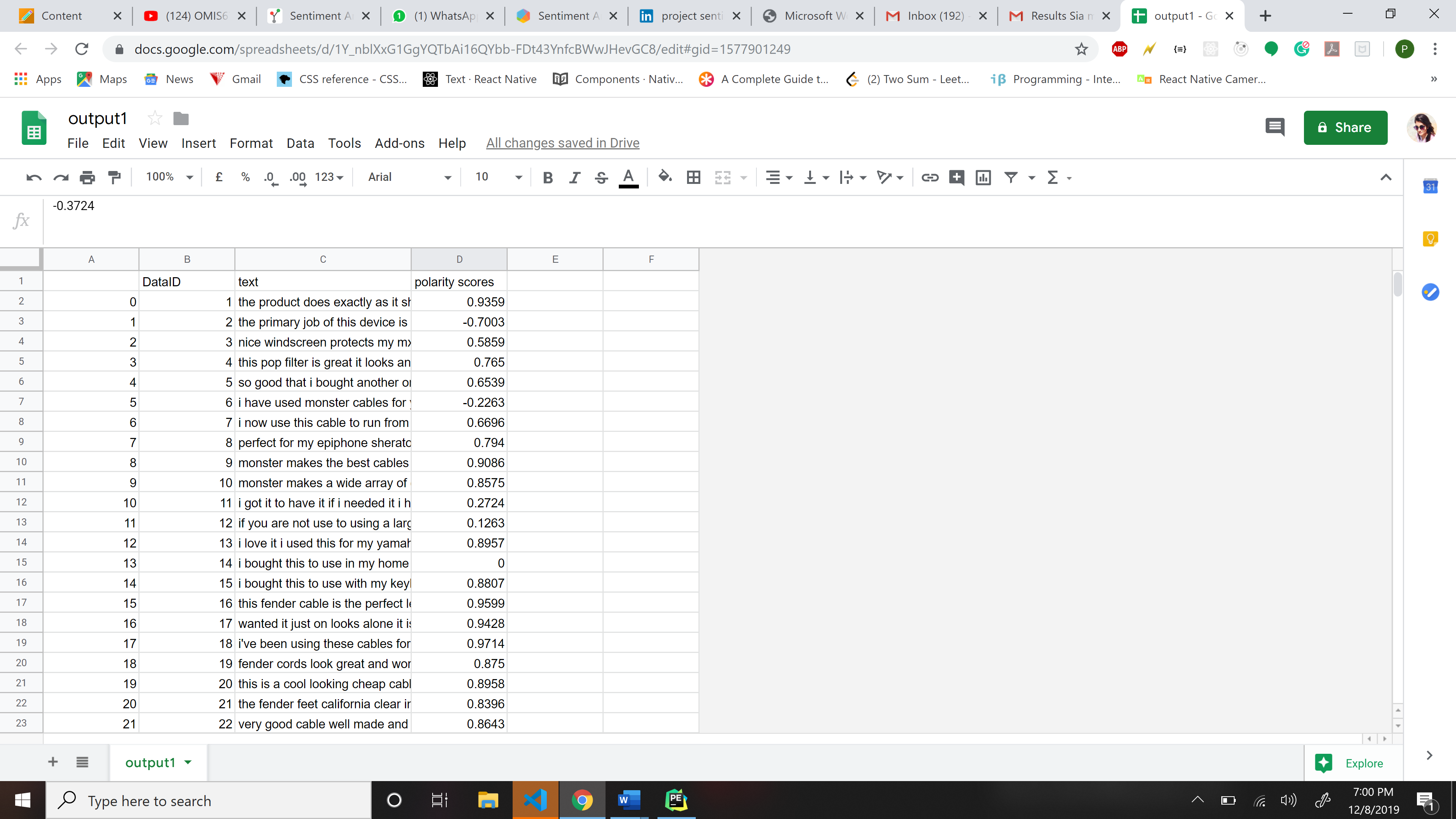
The python script is as follow:



**Packages used:**

* Sentiment Intensity analyzer from NLTK
* Pandas to import text data

Output file:



To differentiate the scores in positive, negative and neutral we used Power Bi tool by adding the additional conditional column named “sentiment”. We have considered the positive reviews which are greater than 0.5 sentiment score, negative reviews which are less than 0 sentiment score and lastly, neutral which are less than 0.5 sentiment score.