MGT561 Text Mining Final Team Project

Analysis of Japanese Whiskey Review Data

Section 01 Group 08

Fiona Liu 466054, Xinyun Liu 467731, Xiaoyao Wang 465943, Yutong Zheng 467035, Julia Huang 466044

Data Description

Our dataset (https://www.kaggle.com/koki25ando/japanese-whisky-review) was scraped from Master of Malt, and online booze store that delivers whiskey to consumers. The dataset contains 1130 customers' reviews for 4 Japanese whiskey brands that Master of Malt is selling. The first column of the dataset is the **numerical label** of the review, and the second column, **Bottle_name**, indicates the name of the whiskey as there might be multiple types of whiskeys under one brand. The third column, **Brand**, shows the Japanese whiskey brands, which includes Hibiki (響), Yamazaki (山崎), Hakushu (白州) and Nikka (ニッカ). This dataset covers four brands in total. The following column is the **title of the review**, which summarizes the overall idea or sentiment about the particular whiskey in several words, such as "overpriced" or "great Japanese whiskey". The last column, **Review_Content**, entails how customer thinks about the whiskey in detail, which is the main target we will analyze in the project.

Objectives for the Analysis

Our first objective is to find out what aspects of Japanese whiskey do customers care through topic modeling of reviews. For example, some reviews are related to the taste of whiskey, some focus on the prices of products, others comment on the appearance of the bottle. For each topic, we aim to find out what characteristics are the most valued by consumers. Companies of Japanese can design and produce their whiskey by considering those different dimensions that customers care about.

The second objective is to compare reviews across different brands. We aim to analyze the characteristics of Japanese whiskey and compare cross each national brand to find out how those brands differ in consumer perception by plotting word cloud. Using sentiment analysis, We can also find the brands with the most positive reputations. We also aim to advise companies on how they can make managerial actions on negative and positive reviews.

Based on the insights gained above, we would further list several suggestions in terms of how businesses can use those insights to make decisions and succeed in the Japanese whisky industry.

Methodology

- **Topic modeling**

We used Latent Dirichlet Allocation (LDA) Model to do topic modeling.

Firstly, the corpus was normalized to get separate tokens. Then we estimated the topics using LDA with different numbers of topics and sparsity parameters. We combined human judgement and metrics like Log-likelihood, Perplexity and Topic Coherence together to select the best topic numbers and sparsity parameters. Finally, the modeling results were visualized and business insights are generated accordingly.

One pro of the model is that the Dirichlet distribution is exponential and conjugate to the multinomial distribution so that the variational inference is tractable. Also the variational parameters of theta (document-specific) could be regarded as the representation of a document so the feature set can be reduced. *z* are sampled randomly within the document so one document can be associated with multiple topics.

One defect of the model is that LDA Model assumes independence, and therefore it is unable to capture the correlation between different topics.

- **Sentiment analysis**

We used the unsupervised machine learning method - Lexicon-based sentiment analysis - for sentiment analysis. VADER was used as the dictionary to assign positive and negative scores.

One advantage of the unsupervised approach is that it does not require annotated data to learn a model. Since the dataset we used is unlabeled, the unsupervised approach is more convenient than the supervised approach and it will save much time and labour.

One defect of the unsupervised approach is that we cannot measure the accuracy, precision, or recall. It will take much time and energy to label the reviews manually if we want to measure them. Another problem is, using general lexicons, unsupervised methods might perform worse on prediction accuracy for the reason that they can't take into account the specific domain & context of our target corpus.Since we uses the existing dictionary to define whether a word or a word bag is positive or negative, the unsupervised approach might not be able to capture words with specific meanings for Japanese whisky. For example, "bitter" does not necessarily mean negative for whisky but according to VADER the vocabulary has a negative score.

Results and discussion

- **What are common words that are used to describe whiskey taste?**

Like what we did in class, we use the review_content column to build the topic modeling. At first, we wanted to use only the title for analyzing because the title can reflect the attitudes of the customers. However, the word sizes of the title are too small, and it may not be that representative. Therefore, review_content turns out to be the best choice.

At first, we normalized the corpus and then we estimated topic models using the Latent Dirichlet Allocation. After the results were shown. We changed different numbers of topics and sparsity parameters. The best number of topics we got was 7 and the best number of words in each topic was 9. This gave us a more satisfying result because the words in each topic are not overlapping as we tried in 10 number of topics. In addition, every topic was included and categorized in a clear division. Also, the topics could tell the better story as the number of words increase. When we tried 6 words in each topic, it was hard for us to see the meaning of each topic. After we increased the number of words, each topic became more human-like language and better understood.

The sparsity parameters we changed from 0.05 to 0.01. We found that when adjusting alpha (doc_topic_prior) and beta (topic_word_prior) to 0.01, which was close to 0, the words were better classified. Words like “whiskey”, “fruit”, “Yamazaki” were grouped together and “expert”, “commenter”, “guy” were grouped together. From the result, I learned that by introducing sparsity parameters, the story of each topic became more complete.

- **What are the characteristics of Japanese whiskey and each national brand?**

By running a histogram for counts of the four whiskey brands (Appendix A), we found out that Yamazaki got the most reviews and was the most popular brand, and Hakushu was the least popular brand. When we ran word vectors and calculated the word weights, the result indicates that the word “taste” has the most weight of 0.00592 and it shows that consumers were most concerned about taste when they purchased Japanese whiskeys. The visualization plot for the reviews (Appendix B and C) shows the most popular words are “whiskey”, “Yamazaki”, “bottle”, “taste”, “fruit”, “love”, “like”, “friend”, “good”, and “nice”. Among those words we could confirm that the brand of “Yamazaki” was the most discussed and it might be the largest brand in the Japanese whiskey market. Customers paid attention to whiskey’s bottle design, taste, ingredients when deciding what to buy, and they would make recommendations to their friends if the whiskey is good. The general sentiment perceived from the most popular words is positive, as words like “love”, “like”, “good”, and “nice” occurred at a relatively high frequency. According to the graph, the topics were very much concentrated as there was only one large circle, which makes sense because the domestic Japanese whiskey market was relatively small and all the brands were competing on the same dimension.

In order to explore the characteristics of each national brand of Japanese whiskey, we performed a word cloud analysis for each of the brand and for the column of “title”, which summarizes the general idea of the reviews. Referring to the word cloud of Yamazaki (Appendix D), the most popular brand, the most mentioned words are “best”, “great”, “better”, “friend”, and “malt”. It implies that one of Yamazaki’s whiskey is made of malt and consumers viewed this brand as their top choice, and they would be willing to recommend this brand to others or share Yamazaki’s whiskey with their friends. The word cloud of the second brand, Hibiki (Appendix E), shows this brand might have special bottle design and blended whiskey flavors. Consumer also mentioned prices, but we are unable to identify whether the price is too high or too low simply from the word cloud since there is no sign of positive sentiment or negative sentiment. Hakushu, the third brand (Appendix F), probably is also made of malt and has a smooth and sweet taste. Consumers mentioned words such as “Japanese” and “alternative”, which might be interpreted as Hakushu is an alternative for Japanese whiskey if there is no other choices. Lastly is the brand of Nikka (Appendix G), and it is also known for its bottle, and consumers enjoyed the taste very much. The word cloud of the column “title” (Appendix H) showed consumers’ general attitude about Japanese whiskeys, and they gave high comments about whiskey’s taste when they wrote “delicious”, “excellent”, “good”, and “phenomenal”. However, some consumers thought those whiskeys are overpriced. Overall, Yamazaki undoubtedly is the most popular brand in Japanese whiskey market, and consumers were positive about all the four brands regarding their taste and ingredients.

- **Which brands have the most positive reputations?**

Based on the Lexicon method with threshold equal 0.1, we found out the different proportions of positive and negative review cross four brands. Overall, there are 1130 reviews, 930 of which are positive and the rest are negative. (Appendix I)

There seems to be some disparity among those four brands to some extent. Hakushu has the highest ratio of positive sentiment, following by Nikka, Hibiki, and Yamazaki. But throughout four brands, the average proportion of positive words is 18% and differences are insignificant. The following bar chart illustrates the ratios of positive and negative comments from customers for 4 brands.

If we set the threshold at 0, the results could be different. The brand with the most positive reputation turns out to be Nikka, following by Hibiki, Hakusha, and Yamazaki. The difference in the ranking indicates that the results of threshold at 0.1 is not robust. In other words, there is not much difference in the reputation among the four whisky brands.

However, there is at least one lesson that we can learn from. Yamazaki ranks the lowest in both cases, indicating that the brand is perceived relatively negative among customers to some extent.

Conclusion

To sum up, Yamazaki is the most popular brand in terms of number of reviews, assuming that the more reviews a brand has, the more purchases consumers are making. However, there's another possibility that it is the most controversial brand and people simply have the greater incentive to commend on Yamazaki. If this is the case, Yamazaki will get relatively less proportion of positive reviews. But considering that in the sentiment analysis results all 4 brands have similar proportion of positive reviews, it is more likely that Yamazaki has the largest number of reviews because of higher sales.

In terms of the characteristic of Japanese whiskey and each brands: the bottle design, taste and ingredients are the most considered factors; Yamazaki is viewed as the top choice of consumers and is for share with friends; Hibiki has special bottle design and blended whiskey flavors; Hakushu has a smooth and sweet taste; Nikka is known for its bottle and taste.

From the point of sentiment of reviews, the Japanese whiskey as a whole and each individual brands all have around 80% positive reviews, indicating that there's no obvious difference among consumers' sentiment towards different brands.

Recommendations

Whisky businesses can use the insights above in several ways.

- Help brand positioning. Whiskey companies that cooperating with Master of Malt can utilize the review analyzing results to design and produce their whiskey by considering different dimensions that customers care about: bottle design, taste, and ingredients. The result of text mining can inform management team about the perception of different brands, and assist in the decision of enhancing existing advantage or change product positioning to occupy market vacancy. Since customers would make recommendations to their friends if the whiskey is good, a company could design its product and marketing strategy from those dimensions to win more customers through the network effect.

- Build recommendation system. Another application is to combine how a specific consumer would perceive different whiskeys in various dimensions into the recommendation system. That's to say, useful and meaningful features can be abstracted from the reviews and then feeded into the recommendation system to enhance personalized recommendation.

- Based on the sentiment analysis, companies can detect dangerous warnings if their reputations decrease a lot among customers, or if they have the lowest proportion of positive reviews in the whole market. Once detecting the sign, a company can further dig into reviews to find out the root of its problem and also find out why its competitors perform better.

- Companies can make managerial actions on negative and positive reviews. Different values of threshold should be used for different purposes. For example, if a company wants to offer additional discounts to consumers who post negative comment about its whisky, it should pay attention to negative reviews and minimize the false positive, and therefore set the threshold above zero. If a company wants to do price discrimination by setting higher price for those who post highly positive reviews, it could minimize the false negative and hence set the threshold below zero.

- Businesses can compare customer reviews on different platforms to gain a more comprehensive of brand perception and also make different marketing strategies for each platform. The perception of customers can differ across various platforms as they may have different target customers. For example, most of customers on Amazon could be regular buyers who buy whiskeys for daily use. Customers on Master of Malt could be large-batch-buyers or whisky-lovers who would register for the website.

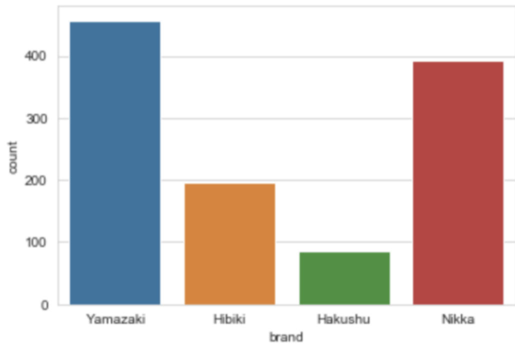
Limitations

- The sentiment analysis results can be affected by different values of threshold. The company need to put extra efforts in setting the optimal threshold value that is good in both enhancing prediction accuracy and assisting business decision making. To do that, the company might need to search for domain specific lexicons or train their own lexicon. Also, the company need to be clear in the expected action out of the sentiment results in order to decide which type of error (false positive & false negative) it cares about the most and adjust threshold accordingly.

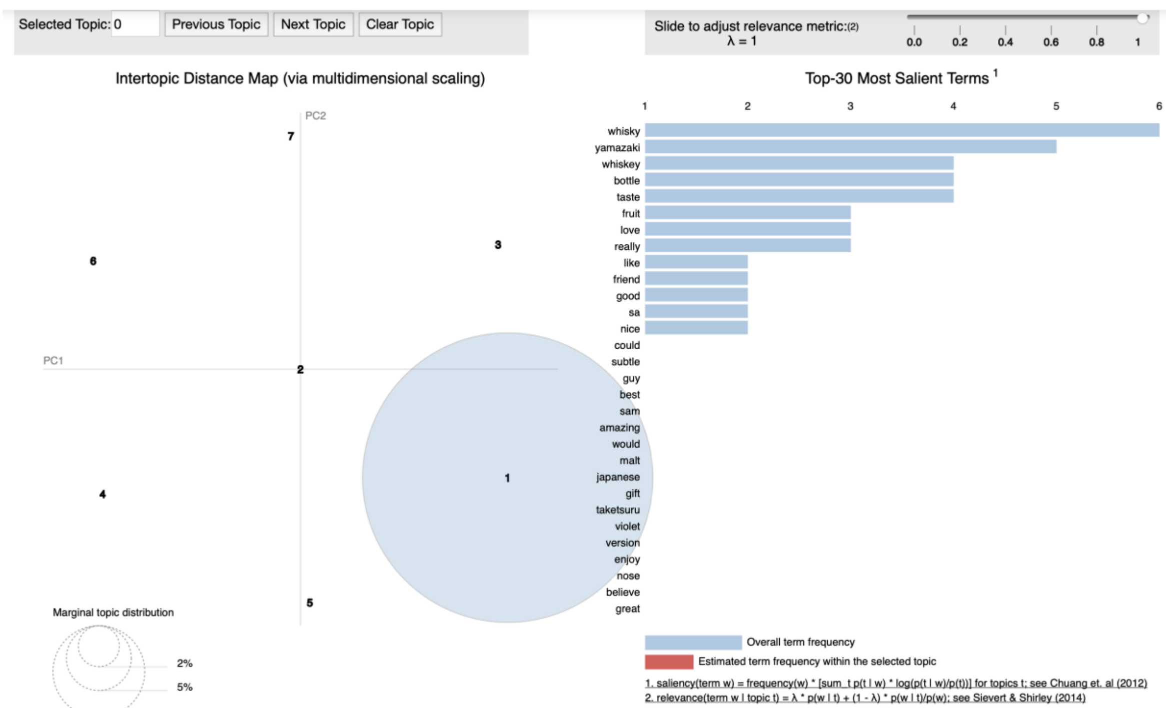
- Our sentiment analysis lacks comparable benchmarks to generate further business insights. Although it is clear that the four Japanese brands sold on Master of Malt all have around 80% positive reviews, we don't know whether this radio indicate good or bad performance -- it depends on the average proportion of positive reviews for the whole category. If most online platforms have around 90% positive reviews for Japanese whiskey, Master of Malt might need to consider how to improve customer satisfaction.

Appendix: table and plots

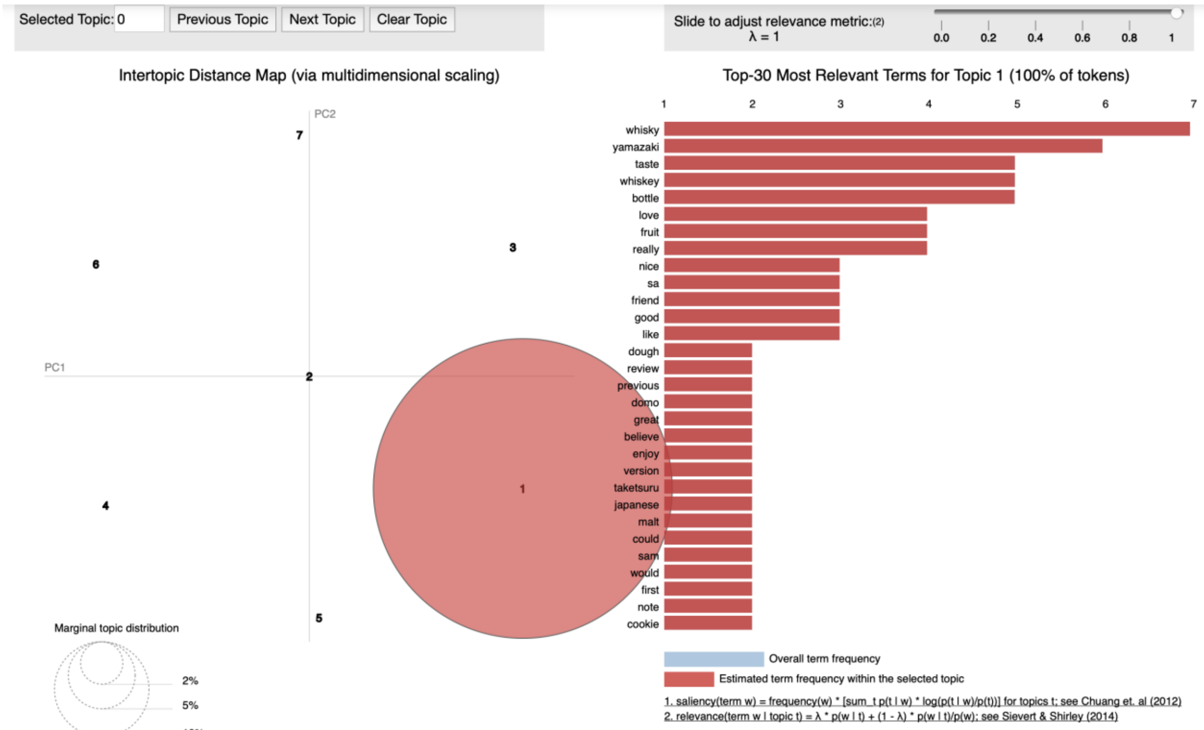
Appendix A



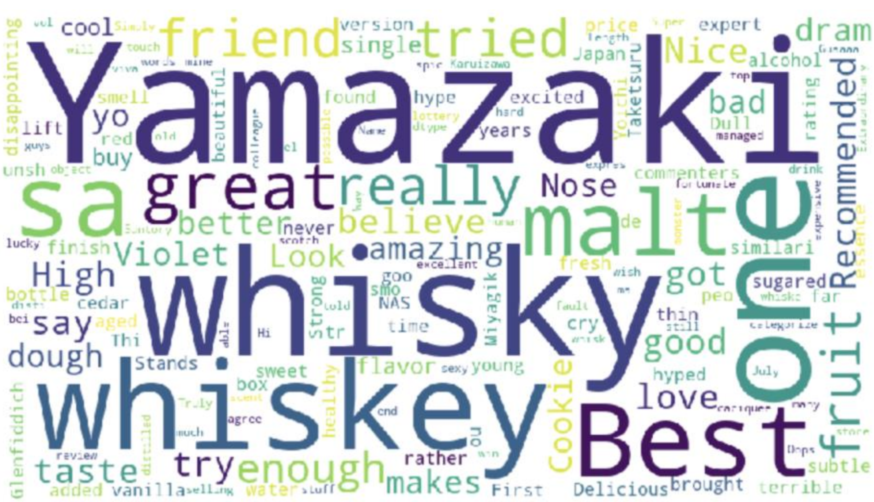
Appendix B



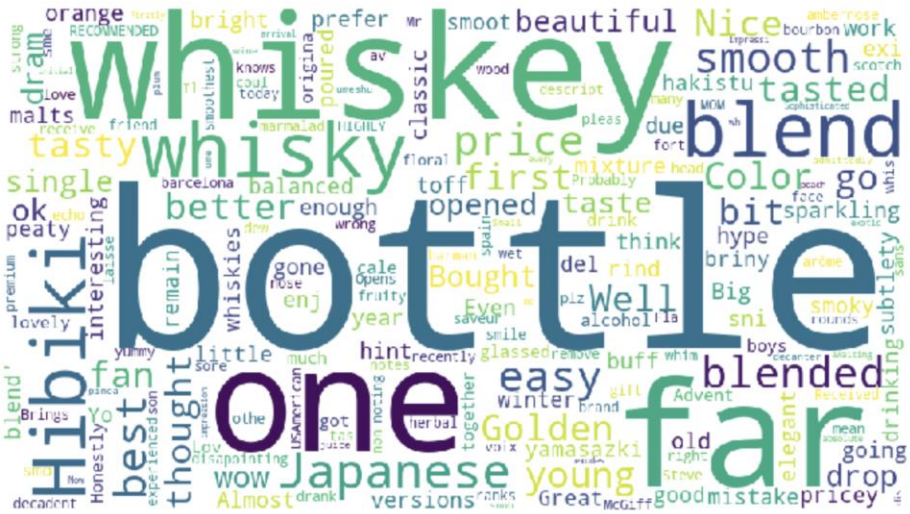
Appendix C



Appendix D Yamazaki

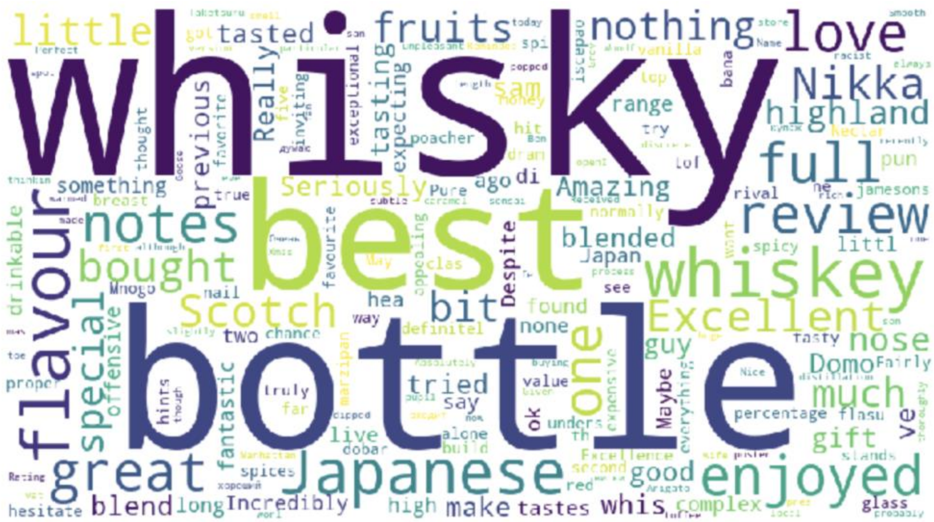


Appendix E Hibiki

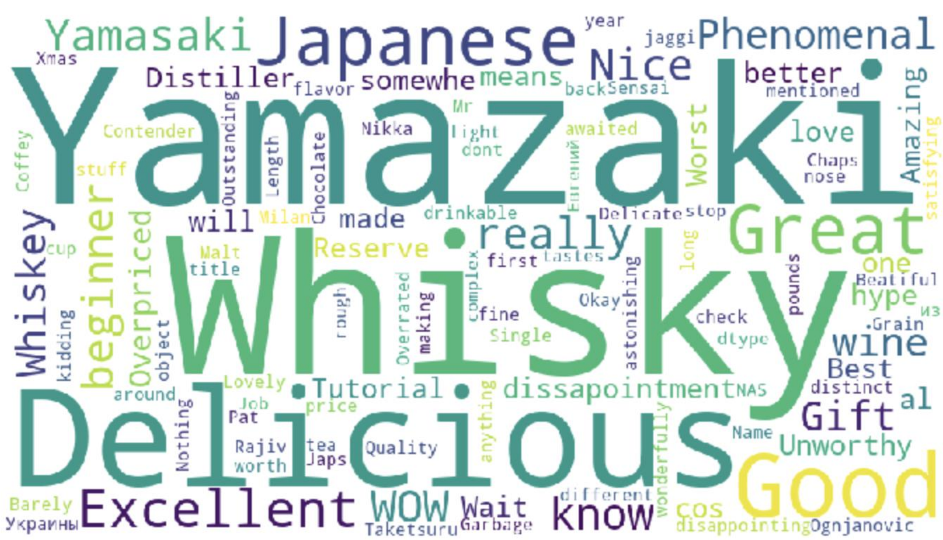


Appendix F Hakushu

Appendix G Nikka

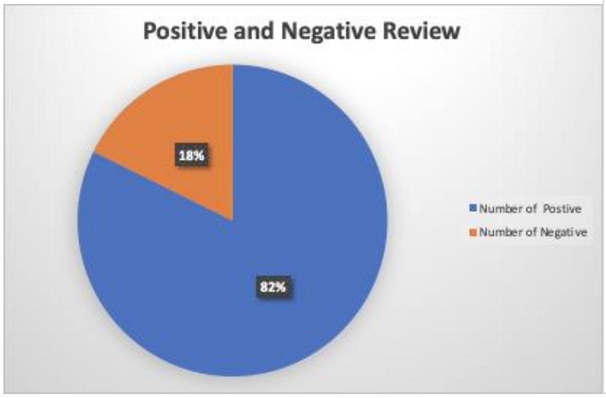


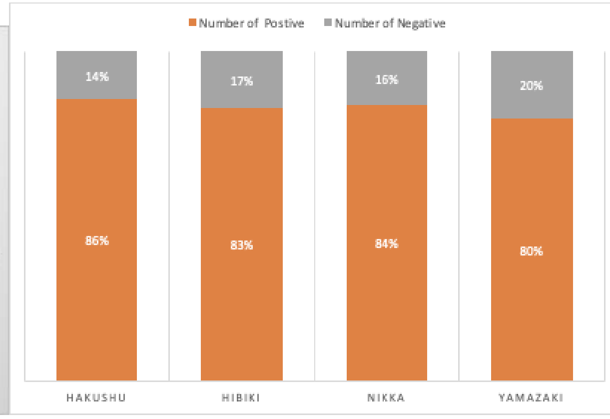
Appendix H Title



Appendix I

When threshold = 0.1





When threshold = 0