

## Topic Modeling for Columbia Sportswear Products on Amazon.com

The Amazon Product data used in this project contained about 1.5 million clothing, shoe, and jewelry products and 5.7 million reviews of those products. Of those, 4,988 products were Columbia Sportswear products having a total of 27,278 reviews. Across all 27,278 reviews, Columbia products received an average rating of 4.32 out of 5. Using this data, I was able to create a few different topic models using K-means clustering to better understand strengths and weakness according to reviewers across different product groups.

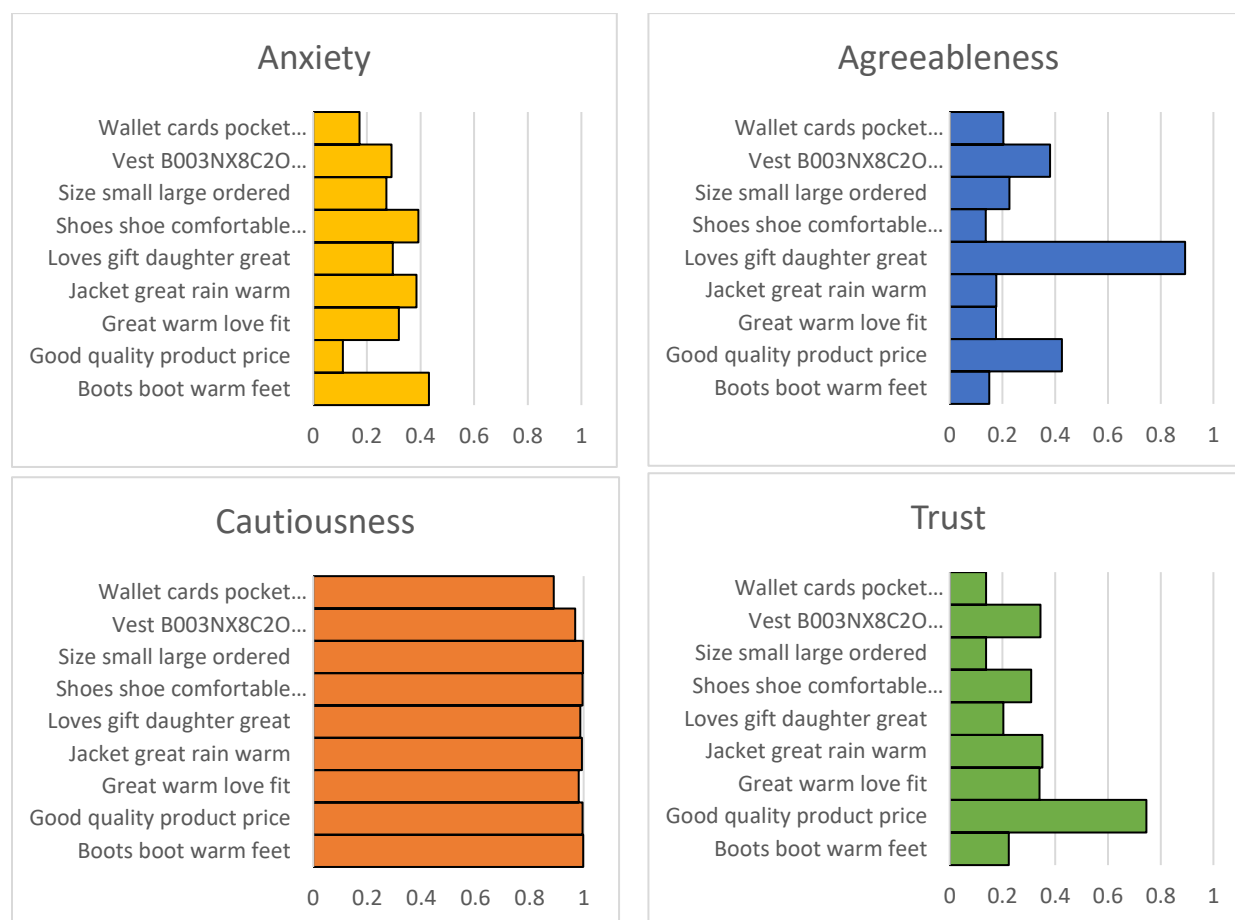
The first step to creating the topic models was to preprocess the data and gather specifically Columbia Sportswear data. I first loaded the product data and iterated through it, creating a dictionary of product categories as well as a dictionary of all products. I then specifically pulled all the Columbia “ASINs” or Amazon Standard Identification Number and matched reviews to those ASINs to ensure that the reviews were for a Columbia product. With this matching, I was able to create a dictionary of Columbia reviews. The next step in preprocessing was to create a list of stop words. I noticed later that there were Spanish reviews as well, so I included both English and Spanish stop words, as well as the word “Columbia” since we knew that would most likely be showing up a lot in the reviews. From there, it was trial and error to get an accurate number of clusters without too much overlap.

In the end, I chose nine clusters. This was the highest number of clusters I could keep before things started to overlap, particularly jackets and coats. This overlap does make sense, as Columbia Sportswear predominately sells outdoor clothing and jackets are an item that would have many variations of construction and use case, especially for outdoor activities. Additionally, as clusters increased, there seemed to be multiple boot clusters as well. I expect that this is the case due to the same reason as jackets; many different use cases and types of boots creates huge clusters that are easy to break apart. However, to simplify the model, I only included enough clusters to have two boot/shoe clusters and one jacket cluster. These nine clusters are: Boots, Quality and Pricing, Warmth and Fit, Jackets, Gifts, Shoes, Sizing, Vests, and Wallets. These nine clusters represent a wide variety of products and with the slight exception of Warmth and Fit, they seem to be fairly distinct in their topics. Many of the clusters are very highly ranked (near 4.5 stars) however, ‘Sizing’ is a topic to dive further into, as the score for that cluster is low compared to the others. These clusters, their scores, and a short description can be found in the table on the next page.

## K-Means Clustering Topics

Cluster	Top 4 Words	Description	Score
01	Boots boot warm feet	This cluster contains reviews about boots and at a glance seems positive, especially about their warmth. Hiking is a very common activity in this cluster. This cluster is extremely positive.	4.483
02	Good quality product price	This cluster contains reviews not necessarily about any one product but about the quality and construction of Columbia products. This category is extremely positive.	4.493
03	Great warm love fit	This cluster is not as well formed as the other clusters and seems to overlap cluster 02 and cluster 07. It seems to cover comfort, warmth and fit.	4.207
04	Jacket great rain warm	Though rain is a very common word in this cluster, this cluster also contains winter jackets like ski jackets and down/fleece jackets. This cluster is extremely positive.	4.537
05	Loves gift daughter great	This category predominately contains reviews of gifts people bought for friends and family. Though daughter is a top word, son, mother and husband are also incredibly common in this cluster. This cluster is extremely positive.	4.851
06	Shoes shoe comfortable great	Different from the boot cluster, this cluster seems to include variations of sandals and shoes. This would make sense as “warmth” is not as common of a word, and sandals have no warmth.	4.294
07	Size small large ordered	This is the lowest rated cluster out of all nine clusters, and it was consistently the lowest rated cluster even with more clusters. I expect part of this is due to buying clothing online, unable to try it on. However, many of the sizing issues in this cluster seem to be about how Columbia sizing differs from other brands.	3.777
08	Vest B003NX8C2O B00062NNLK great	This cluster is unique in that it contains specific products directly in the top four words. Each cluster of reviews contains the product ASINs, but this is the only one that contains the ASINs in the top four words. The products are “Women’s Benton Fleece Vest” and “Men’s Cathedral Peak Fleece Vest”.	4.442
09	Wallet cards pocket leather	This cluster contains all reviews about wallets. Many reviewers speak about the smell and the durability of each wallet. This is the second lowest rating of the clusters after sizing.	4.155

Looking at these clusters and their average ratings, there are clear winners and clear topics that will need to be improved. Most notably, the 'Wallet' and the 'Sizing' clusters need to be improved. To get a better idea of the content of these clusters, I proceeded to extract "personality insights" from the text files using IBM's Watson API. This showed some interesting divergences between categories as one can see in the charts below.



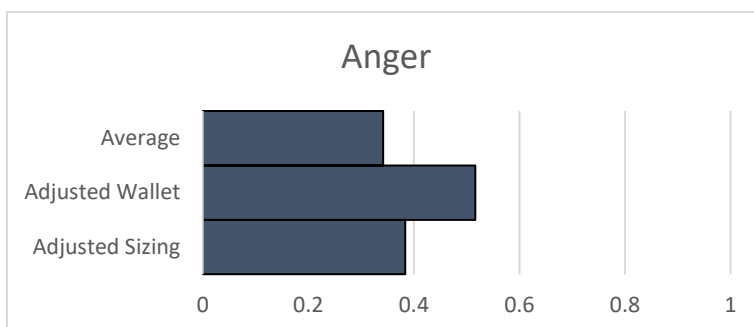
As we can see from these charts there are some extremely consistent trends between clusters and some divergences. Overall, the reviews seem to be incredibly cautious. This would seem to make sense, as someone writing a review is stating an opinion publicly and where the company can see it (theoretically). Anxiety overall is low, however the 'Quality and Price' category is particularly low. In fact, this category stands out in Trust, Anxiety and Agreeableness. The low anxiety in this cluster combined with high trust and agreeableness and a high star rating suggests that this category is extremely positive. The categories with the lowest trust score were 'Wallet' and 'Sizing'. Trust is incredibly important for a brand's image and for continued support of a brand and these are the two categories that require further work into quality control and product design. I believe that the 'Gift' cluster is possibly showing low trust because the reviewers are buying a gift and are unsure if the intended recipient will like it or not, displaying wariness in their review.

Because ‘Wallet’ and ‘Sizing’ are the two lowest rated categories and because they each score very low on the Trust category put forth by IBM Watson, I decided to extract all the low ratings from each topic cluster and re-run IBM Watson Personality Insights on these particularly poor ratings. The ‘Adjusted Wallet’ cluster had 66 ratings below 3.0 stars (not inclusive) with an average rating of 1.530 and the ‘Adjusted Sizing’ cluster had 433 ratings below 3.0 stars (not inclusive) with an average rating of 1.643. This is a significantly higher proportion of poor reviews in the ‘Adjusted Sizing’ cluster than the ‘Adjusted Wallet’ cluster. Simply reading through the reviews in the ‘Adjusted Sizing’ cluster, one gets an idea of the problem in this cluster: the sizing is unexpected by customers. Many customers state that the product is smaller than what they expect and that the products “run small”. In fact, the word “small” shows up in over half of the ‘Adjusted Sizing’ cluster reviews that have a rating below 3.0 (not inclusive), or 280 times in just 433 reviews. Other reviewers state that the color or the thickness is not what they expect – most likely a common problem across all clothing sales online. Reading through the reviews in the ‘Adjusted Wallet’ cluster, one gets a sense that customers see the products within as ‘poor quality’ and that the wallet is not durable. Two reviews state that the wallets smell bad.



Using IBM Watson Personality Insights, we were able to see the same insights as before. Again, though Watson categorizes many different types of insights, I believe these four are most relevant to this topic. As one can see above, these adjusted topics are plotted against the average of all Columbia Sportswear reviews in the dataset. Immediately obvious is that these two adjusted categories have extremely low trust compared to average. Curiously however, though all nine topics rated extremely high in cautiousness, ‘Adjusted Wallet’ does not. This perhaps means that poor reviewers of wallets are much confident or forceful in their opinions of the product. I suspect that this is because ‘Adjusted Wallet’ contains many reflections of quality, whereas ‘Adjusted Sizing’ seems to contain confusion about sizing, rather than outright anger or

aggression. This interpretation of anger about wallet quality seems to prove correct, as one can see in the chart below. Though not extremely high, ‘Adjusted Wallet’ anger is quite a bit higher than average and ‘Adjusted Sizing’, which are relatively similar.



In conclusion, based on the topic models above, Columbia Sportswear is doing a good job with customer satisfaction on Amazon sales. Overall, each cluster had a high rating and even the ratings that were poor compared to the Columbia average were still above 3 out of 5. However, it is clear that in particular, two topics need work: wallets and clothing sizes. Based on the research above, bolstered by IBM Watson Personality Insights, it appears that poor reviews of Columbia wallets and money clips focus on poor quality and durability. To keep trust in the brand, the perceived or actual quality of wallets must increase. As for sizing, customers are fairly confused about Columbia sizing in relation to other brands. Many reviews state that Columbia clothing and shoes ‘run small’. An investigation into whether Columbia clothes are marketed as bigger than they are relative to other products on the market may be needed to determine if Columbia clothing and shoe sizing is accurate and if it needs to change. Accomplishing these two tasks would hopefully result in an increase in brand trust, and bolster the brand.

