

In our study, we conduct correspondence analysis on four products - Vitafusion gummy bears, Renew cosmetics, Thorne basic supplements, and Thorne prenatal supplements. We started off our project by webscraping reviews for these products off amazon.com. We then identify “stop words”, words that we classify as portraying little to no relevant information. In our analysis, we categorize ratings into three groups. One star out of 5, and two stars, are considered to be “negative reviews” whereas four and five star ratings are “positive reviews”, and finally three star ratings are neutral. We count the number of times keywords appear in positive, negative, and neutral reviews. We store these counts in data frames in a data frame.

In this interpretation, I'll focus on our first biplot (the correspondence analysis plot for Vitafusion). We visualize keywords against their respective classifications, where keywords are rows and the sentiment classifications are the columns. Horizontally, we see positive reviews charted to the left and negative reviews to the right of the chart, where neutral reviews are in between the two. Keywords that are stored near the center of the data tend to be less impactful or relevant on the grand scale of the reviews. For example, the words “diet” and “sugar” might be considered as keywords depicting positive emotions, but they are not as associated with positivity as a word such as “price” or “tastes”, as these keywords are far from the center of the chart. Several keywords are clustered near the center of the chart, so even though they can be classified as one of three sentiments, they do not provide the most useful information when compared to keywords further away. The strongest keywords present in this chart are “stars”, “easy”, “tastes”, “textured”, “ordered”, “rubber”, and “melted”. To no surprise are rubber and melted gummies prevalent in negative reviews for Vitafusion gummy bears. The results of our model seem to match the general notion of what words depict. The biggest outlier in our chart is the word “stars”, which - as expected - corresponds to neutral reviews. “Easy” and “price” were quickly associated with positive sentiment, making us assume that the gummy bears are easy to chew and are sold at low prices. However, here is where things start getting a little uncertain. The phrase “price” does not necessarily give much insight about the product without the context it's used in. For example, I oftentimes regard things to be “of high price” or of “too much price”, which would then be considered as a negative sentiment. What's more is that there's a strong distinction between the term “price” and the term “textured”. The keyword “textured” was classified as a neutral word. Going in with the same evaluation as “price”, this makes a lot of sense. Textured can either be a pro or a con of a food-based product. Maybe some customers like their food to be textured, making it unique and providing distinction. Other customers might consider this to be a defect or a disadvantage of the product, as it doesn't fall with the traditional norms of a standard gummy bear. With two sides to the argument, labeling the term “textured” as a neutral term makes a lot of sense. However, this makes me ponder about the classification of the term “price”. Price is a similar case to texture - it can either represent an advantage of the product or a disadvantage depending on the context of the review. In this case, why is there a large discrepancy between these two keywords? The answer to this question is data.

The data that we are conducting sentiment analysis on would be Amazon reviews regarding Vitafusion gummy bears. With several keywords that represent prominence, as well as consistency, in user reviews, companies and manufacturers can better grasp an understanding of what words and phrases customers often use to describe their products. This particular case

study is vital because it shows that a word such as “price”, which could be strongly thought of as a neutral term without any additional context, is actually a phrase that Vitafusion customers use to show satisfaction with their gummy bears. Vitafusion can use the insights from our model to better understand their customer needs and propel user satisfaction through their products.