ADA project M1

Feedback:

Strong and realistic proposal overall, great job! 1)Very interesting and unique approach! This is a strong proposal and looks feasible. 2)The idea overall is interesting. I didn’t understand fully what you wanted to do with the price, but it’s a promising idea. For the price, you may need other data sources as I don’t think it’s present in the data. 3)This approach could be interesting. However, it’s not guaranteed that only users with lots of ratings are experts. There could be experts with few reviews. It could be also interesting to see if these users with lots of reviews are really experts or not, but it would be harder to determine who is really an expert or not.

Your project mentor throughout the semester will be Beatriz Borges : beatriz.borges@epfl.ch. For future discussions specific to your P2 and P3 deliverables, you are encouraged to be in touch with your mentor.

1. “Taste same but feel differently”

Since humankind started communicating, diverse kinds of languages have evolved continuously and iteratively. New phrases were invented and became the common usage for younger generations, whereas some vocabularies were considered outdated and thus marginalized. Furthermore, the emergence of social media, short videos, and fast fashion are also changing the way people think about themselves and express their ideas. Such a temporal change of wording should be reflected in every aspect of life, including beer reviews. By studying the wording patterns of beer reviews across categories and time, we might be able to identify the enriched sentiment words of each period (in the scale of years). The alignment of these discoveries with the online corpus such as Google Ngram Viewer might allow us to validate and interpret why the way people describe beers changes as time goes by, although the beers themselves aren’t necessarily so different. Possible analytical methods include correlation analysis, hypothesis testing, and heatmap visualization.

2. “The most controversial beer”

It’s hard to have everyone appreciate the same thing, as the beer. For those types of beers that people have a wide spectrum of opinions on (can be defined based on the distribution of ratings), it would be interesting to know what are the possible factors that contribute to the variance. Is it always because of the aroma? Or is it determined together by appearance and palate? Furthermore, would the variance become smaller as the average price of the beer increases? Perhaps people tend to give compliments to pricy beers because of better quality or partly due to cognitive bias. These phenomena can also be linked to alcohol content, beer style, user nationality (for example, maybe French people tend to complain more, according to what my French friend told me), and many other factors. Possible analytical methods include multi-variable regression, dimensionality reduction, and hypothesis testing.

3. “Does commenting more make you a professional beer taster”

There is no guarantee that every person’s review reflects the real flavor and popularity of the beer. An intuitive idea is that credibility derives from experience. How experienced the person is determines how reliable the comments are. If a person contributes thousands of reviews, doesn’t that prove his or her expertise? Following this logic, it would be quite interesting to study the corpus of people who made significantly more comments about beers. Do they have better beer knowledge in terms of the usage of professional, precise terms for conveying objective opinions that align well with the majority of reviews? For individual commenters, did the wording of the comments change over time? Were their comments dependent on specific beer styles or breweries?