**Difficult moments in completing this project.**

1. The ‘review.json’ file was too large to be loaded into notebook. EVERYTHING was TOO LARGE.
2. Eventually solved this problem using itertools (chunksize wouldn’t work for some reason).
3. We were not familiar with how to deal with the hierarchy structure of json file, thus it is hard to sort out the information in lower part of the structure.
4. We were not familiar with how to deal with the hierarchy structure of json file, thus it is hard to sort out the information in lower part of the structure.
5. Text analysis is complicated - how to eliminate stopwords, sarcasm, phrasing confusion? (used module).
6. Choosing which parts of the data to use.
7. Files have some superfluous variables, for example some that only apply to hair salons, that must be eliminated. Other variables, like dietary restrictions, seem like they would apply, but had so little data regarding bakeries that they were useless; these also had to be eliminated.
8. We planned to create a star schema database using MySQL community server, since the results from the python coding gave us enough insights about the guiding questions we were in a huge dilemma to implement the Star Schema database for the Yelp dataset.
9. Despite the opportunities to use the colonial one super computer for running the analyses hassle free, we didn’t utilize that, as the option didn’t become available until we were already well underway analyzing the data in Python.

**Better times in completing this project.**

1. It was a great learning experience.
2. Knowledge sharing was at its peak as group members shared what they had learned in other classes or experiences, as well as coding tricks they had figured out.
3. We had a wonderful team which was proactive and supportive.
4. We did not have any misguiding results and spurious dataset.
5. As the dataset was simple to understand and reflect, we had no issues in framing the guiding questions and working upon the solutions. Each discovery was interesting to us and we could relate it back to our lives on a personal level.
6. We did not run into errors that would have ruined the flow of the project. We used the simple descriptive analysis codes that were learnt in the prior classes.
7. We had the opportunity to learn to create “Word Clouds” in python. This is an effective data visualization tool.