Feedback for Project Number 01

**What the topic on?**

* Web Scraping Yelp reviews on the 24th Street Pizzeria: Understanding why this restaurant has a 1.5-star rating.

**Is the title consistent with the topic?**

* Yes, the title is consistent with the topic. The presentation revolves entirely around the title and research question. The title is a perfect match in my opinion.

**Are the objectives of the project clearly identifiable?**

Yes, the objectives are clear.

* ***What are they wanting to study?***
  + Reviews posted on Yelp regarding the 24th Street Pizzeria.
* **What is the motivation of this project?**
  + Use text and sentiment analysis to analyze why the pizzeria receives poor reviews.
* **What does the presenter hope to accomplish with this project?**
  + By analyzing the text and sentiment the presenter hopes to gain valuable insight on how the owners of the pizzeria can improve their business and provide better customer service and food.

**What data are used?**

* 170 reviews total
  + 1 Star: 154
  + 2 Star: 5
  + 3 Star: 1
  + 4 Star: 2
  + 5 Star: 8
* **What website(s) are scraped?**
  + <https://www.yelp.com/biz/24th-st-pizzeria-san-antonio?osq=Worst+Restaurant>
* **Are other data brought in from outside sources?**
  + No, the project uses data only from the scraped Yelp website.

**What is your overall impression of the project?**

* Overall, I think this project was simple and effective. The research question is easy to follow, and the insight gained by the data analysis offers valuable clues on how the Pizzeria owners can improve their ratings. Of course, the desire to improve must come from the owners. It would be interesting interview them and gain knowledge on why they think they receive such poor reviews. I can understand if the Pizzeria is super popular and wait times are excessively long. However, burnt or cold pizza should be easily avoidable and should not be left unremedied.

The “Most Used Words: Plot” is simple yet elegant and easy to read. There are no sentence construction words such as “the”, “and”, and “or.” This shows me that the data cleaning and preparation process was effective. Furthermore, I like how you created two-word clouds: one for 5-star revies and one for the 1-star reviews. One thing I want to point out is that there were 154 one-star reviews and only 8 five-star reviews. This is an imbalance, but I don’t think it takes to much away from the comparison and the point you are making.

Regarding the results from the various sentiment analysis tools, I would not worry to much about it. Maybe average them all together or pick one that best fits your conclusion and recommendations. The results from the “Distribution of emotions categories” is puzzling. As you concluded, the reviews are overwhelming negative and in no way positive. I wonder if the IBM Watson sentiment analysis tool would conclude any differently?

In conclusion, excellent job. I enjoyed watching the presentation. It was easy to follow and flowed well from one slide to the next.