**Code 10 – Text Analysis**

For this text analysis project, I want to take a closer look into the Barbecue scene in Austin, Texas. When it comes to the topic of Barbecue, it is usually insisted that Texas is the place to get the best. Austin, Texas in particular is widely regarded as one of the best, if not the best, place to get Barbecue food. Through the analysis of Yelp reviews given to various barbecue restaurants in the city, I aim to analyze whether this stereotype of Texas barbecue is substantiated and if so, what makes these establishments special in the eyes of customers.

In order to accomplish this task, I first set up my parameters for a business search using Yelp API to compile a sample of Austin barbecue restaurants to consider for analysis. This was done as shown:

search\_term = 'barbecue'

location\_term = 'Austin, TX'

search\_results = yelp\_api\_instance.search\_query(term=search\_term, location=location\_term, sort\_by='review\_count', limit=40)

In effect, the purpose of the above parameters was to specify that I would be searching for barbecue establishments in Austin, Texas based on the review count on the Yelp site. My reasoning for basing my search on review count was to prioritize the longest-standing establishments to base this analysis on. The second half of the above code, in addition to specifying my search criteria, executes the actual API search; from here, the next step was to capture the results. In order to do this, I first turned my query variable ‘search\_results’ into a data frame and then transformed that data frame into a csv file for storage. This was done using the code below:

bbq\_busn\_df = pd.DataFrame.from\_dict(search\_results['businesses'])

bbq\_busn\_df.to\_csv('Austin\_bbq\_businesses.csv')

Now with my CSV of Austin barbecue restaurants established, I had to make my choice on which restaurants to analyze. With the aim of keeping my choice of restaurants unbiased and random, I used the following code to randomize a list of 10 restaurants from the CSV.

busn\_df = pd.read\_csv('Austin\_bbq\_businesses.csv')

analyze\_list = []

for i in range(10):

business = random.choice(busn\_df['alias'])

analyze\_list.append(business)

print(analyze\_list)

The code resulted in the list below:

**analyze\_list** = ['moonshine-patio-bar-and-grill-austin', 'franklin-barbecue-austin', 'stiles-switch-bbq-and-brew-austin', 'coopers-old-time-pit-bar-b-que-austin', 'brothertons-black-iron-barbecue-pflugerville-7', 'terry-blacks-barbecue-austin', 'its-all-good-bbq-spicewood-5', 'stubbs-bar-b-q-austin', 'opies-bbq-spicewood', 'donns-bbq-austin-2']

With my restaurants selected, I now had to collect the reviews for each. To do this, I first assigned each of the 10 chosen restaurants’ alias to a variable:

id1\_for\_reviews = 'moonshine-patio-bar-and-grill-austin'

id2\_for\_reviews = 'franklin-barbecue-austin'

id3\_for\_reviews = 'stiles-switch-bbq-and-brew-austin'

id4\_for\_reviews = 'coopers-old-time-pit-bar-b-que-austin'

id5\_for\_reviews = 'brothertons-black-iron-barbecue-pflugerville-7'

id6\_for\_reviews = 'terry-blacks-barbecue-austin'

id7\_for\_reviews = 'its-all-good-bbq-spicewood-5'

id8\_for\_reviews = 'stubbs-bar-b-q-austin'

id9\_for\_reviews = 'opies-bbq-spicewood'

id10\_for\_reviews = 'donns-bbq-austin-2'

Once this was done, I ran a review query on the API based on these aliases and stored each query into a variable of its own:

review\_response1 = yelp\_api\_instance.reviews\_query(id=id1\_for\_reviews)

review\_response2 = yelp\_api\_instance.reviews\_query(id=id2\_for\_reviews)

review\_response3 = yelp\_api\_instance.reviews\_query(id=id3\_for\_reviews)

review\_response4 = yelp\_api\_instance.reviews\_query(id=id4\_for\_reviews)

review\_response5 = yelp\_api\_instance.reviews\_query(id=id5\_for\_reviews)

review\_response6 = yelp\_api\_instance.reviews\_query(id=id6\_for\_reviews)

review\_response7 = yelp\_api\_instance.reviews\_query(id=id7\_for\_reviews)

review\_response8 = yelp\_api\_instance.reviews\_query(id=id8\_for\_reviews)

review\_response9 = yelp\_api\_instance.reviews\_query(id=id9\_for\_reviews)

review\_response10 = yelp\_api\_instance.reviews\_query(id=id10\_for\_reviews)

With this step done, I converted each of the resulting dictionaries into 10 separate data frames, then combined these data frames into one cumulative data frame. Finally, this cumulative data frame was stored in its own csv file. These last few steps were done using the following code:

reviews\_df1 = pd.DataFrame.from\_dict(review\_response1['reviews'])

reviews\_df2 = pd.DataFrame.from\_dict(review\_response2['reviews'])

reviews\_df3 = pd.DataFrame.from\_dict(review\_response3['reviews'])

reviews\_df4 = pd.DataFrame.from\_dict(review\_response4['reviews'])

reviews\_df5 = pd.DataFrame.from\_dict(review\_response5['reviews'])

reviews\_df6 = pd.DataFrame.from\_dict(review\_response6['reviews'])

reviews\_df7 = pd.DataFrame.from\_dict(review\_response7['reviews'])

reviews\_df8 = pd.DataFrame.from\_dict(review\_response8['reviews'])

reviews\_df9 = pd.DataFrame.from\_dict(review\_response9['reviews'])

reviews\_df10 = pd.DataFrame.from\_dict(review\_response10['reviews'])

culm\_df = pd.concat([reviews\_df1,reviews\_df2,reviews\_df3,reviews\_df4,reviews\_df5,reviews\_df6,reviews\_df7,reviews\_df8,reviews\_df9,reviews\_df10]).sort\_index()

culm\_df['text'].to\_csv('A\_bbq\_review\_text.csv',index=False)

**Text Analysis**

With the data to be analyzed collected, the next step was to actually plan and execute the textual analysis. With the goal of my analysis being to figure out what exactly consumers enjoy about Austin barbecue, my first area of interest is to conduct sentiment analysis on the review data collected. To do this, I first executed code to import my newly saved review csv as a data frame. Once my “reviews\_df” data frame was formed, I implemented “SentimentIntensityAnalyzer” from VADER into the following for loop:

analyzer = SentimentIntensityAnalyzer()

for review in reviews\_df:

sentiment = analyzer.polarity\_scores(review)

print(sentiment)

print(review)

print("\n")

Due to the large number of reviews that were collected from the API, this code yielded the following lengthy output:

*{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}*

*this is my second time coming to moonshine patio bar and Grill and let me tell you it always hits. This restaurant has a variety of southern classics from...*

*{'neg': 0.0, 'neu': 0.77, 'pos': 0.23, 'compound': 0.8302}*

*Y'all, this place!! I mean, the brisket is delicious, the turkey is better than ALL the rest, and their sausage makes my toes wiggle!!! Get in their early!...*

*{'neg': 0.0, 'neu': 0.831, 'pos': 0.169, 'compound': 0.7269}*

*First let me say, the people are friendly. The food however, is not what I expected it to be. Coming from Chicago, I've had way better bbq @ I 57; even in...*

*{'neg': 0.044, 'neu': 0.767, 'pos': 0.189, 'compound': 0.743}*

*Hands down the best BBQ in the state... maybe the country.*

*This place is one of the things I miss most about living in the area. I wish there was one...*

*{'neg': 0.0, 'neu': 0.824, 'pos': 0.176, 'compound': 0.7345}*

*Ten stars for this place! We got here early as once it's sold out, you're outta luck. Good thing because as soon as we ordered, I looked behind us and the...*

*{'neg': 0.0, 'neu': 0.65, 'pos': 0.35, 'compound': 0.9253}*

*DELICIOUS! We asked several locals where they like to go for good, authentic, Texas BBQ. Several people told us that Coopers is where it's at! It was a HUGE...*

*{'neg': 0.0, 'neu': 0.797, 'pos': 0.203, 'compound': 0.784}*

*This restaurant was a stone's throw away from our hotel, which was perfect because we wanted to have some good southern BBQ as soon as we arrived! Three of...*

*{'neg': 0.085, 'neu': 0.698, 'pos': 0.217, 'compound': 0.5267}*

*Delicious food. Worth the wait. No need to come too early. They had plenty left when we got to the counter.*

*{'neg': 0.0, 'neu': 0.726, 'pos': 0.274, 'compound': 0.8271}*

*Everything was delicious and the service was nice. Looking forward to trying their BBQ on my next visit.*

*A little pricey but I enjoyed the breakfast.*

*{'neg': 0.0, 'neu': 0.842, 'pos': 0.158, 'compound': 0.5719}*

*From the moment I walked in, the bustling atmosphere and inviting aroma of barbecue had me excited for my very first experience of Texas barbecue. There was...*

*{'neg': 0.039, 'neu': 0.689, 'pos': 0.273, 'compound': 0.8591}*

*First thing that caught my attention is that the staff are super nice and friendly.*

*You enter, get in the lineup, pay, take your cutlery and find a table...*

*{'neg': 0.0, 'neu': 0.804, 'pos': 0.196, 'compound': 0.765}*

*This is such a great spot to get some barbecue, especially if you're coming from Krause Springs, which is like 5 min away.*

*We got here on a Saturday...*

*{'neg': 0.0, 'neu': 0.821, 'pos': 0.179, 'compound': 0.7089}*

*The food nor service were very good. The brisket was on the dry side and though I ordered pulled pork, I was clearly delivered dry turkey or some other...*

*{'neg': 0.0, 'neu': 0.925, 'pos': 0.075, 'compound': 0.2942}*

*Come early & come prepared!*

*We got to Franklin 3.5 hours before they opened. We brought chairs, blankets, drinks, and games, and we had a blast hanging out...*

*{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}*

*This place is tucked behind a gas station in a strip center which is not visible from the road. It is located in the Pflugerville Plaza on Pecan St. they...*

*{'neg': 0.043, 'neu': 0.881, 'pos': 0.076, 'compound': 0.2846}*

*The Waffle Breakfast plate might be the best deal on the menu. I forget how much it cost, but it was less than the Single Meat Plate. It will fill you up...*

*{'neg': 0.0, 'neu': 0.587, 'pos': 0.413, 'compound': 0.9531}*

*Awesome service and cleanliness in the place and in the restrooms. The brisket and beef rib were amazing. The sausage was great too. The best brisket in...*

*{'neg': 0.0, 'neu': 0.923, 'pos': 0.077, 'compound': 0.3291}*

*Perfect spot to bring out-of-town guests. It's a large restaurant but yet feels cozy. We went on a Thursday night and the wait was over an hour. We dipped...*

*{'neg': 0.0, 'neu': 0.798, 'pos': 0.202, 'compound': 0.7955}*

*What was very good - The two sides I had, tater tot casserole and pinto beans, were fantastic!*

*Brisket is usually my go to at BBQ places. I had brisket...*

*{'neg': 0.0, 'neu': 0.818, 'pos': 0.182, 'compound': 0.7712}*

*I enjoyed my dinner with my friends at this Texas style BBQ!*

*Since I know too less for BBQ and only lived in one midwest city, Kansas City , I learned that...*

*{'neg': 0.0, 'neu': 0.75, 'pos': 0.25, 'compound': 0.8402}*

*It was a good hill country destination, and some things were pretty good. Of the meats, the sausage is ok, the pork ribs were on the dry side, and the...*

*{'neg': 0.0, 'neu': 0.932, 'pos': 0.068, 'compound': 0.2263}*

*Food was served cold, beer was warm and the meats don't have much flavor except for their very own sausage which is tasty. Cornbread was cold and...*

*{'neg': 0.0, 'neu': 0.734, 'pos': 0.266, 'compound': 0.8786}*

*I got their TX Rueben to go. I have to say it was excellent. The brisket was juicy. The sauerkraut was tasty and the sauce was really good.*

*It was just...*

*{'neg': 0.0, 'neu': 0.862, 'pos': 0.138, 'compound': 0.5411}*

*This place is a yes! We went mid afternoon on Friday and there was still a line. Parking is tight in the lot but across the street there is adequate...*

*{'neg': 0.062, 'neu': 0.615, 'pos': 0.323, 'compound': 0.8805}*

*Service was good. Parking was crazy expensive 25.00 to park.*

*Staff was very nice and food was good. The beans are excellent.*

*I am more of a thick sauce...*

*{'neg': 0.0, 'neu': 0.877, 'pos': 0.123, 'compound': 0.5859}*

*Bbq fanatic from Florida here. Went with my daughter and partner.*

*Food was amazing. Got the brisket sandwich and the meat was just out of this world. It's...*

*{'neg': 0.0, 'neu': 0.918, 'pos': 0.082, 'compound': 0.3612}*

*Court might find me GRILL-ty for being the minority in not finding this BEEF-tastic like other people do. FRanks a lot for hearing me out.*

*.:: FOOD ::.*

*We...*

*{'neg': 0.0, 'neu': 0.824, 'pos': 0.176, 'compound': 0.7177}*

*Delicious brunch! Solo traveler which means i had the advantage of being able to grab a seat the bar on a saturday am & forgo the wait. The team were...*

*{'neg': 0.0, 'neu': 0.83, 'pos': 0.17, 'compound': 0.7297}*

*It was Memorial Day weekend and back on our BBQ Texas tour where we drove out to Spicewood, TX to check out It's All Good BBQ. I needed to see if this...*

*{'neg': 0.0, 'neu': 0.49, 'pos': 0.51, 'compound': 0.9623}*

*We live our Austin barbeque restaurant. 3 locations to choose from. Always wonderful experience. Great food, great service. Great folks. Definitely 5 stars....*

While VADER does not always yield perfect results, whether due to the interesting grammatical choices of Yelp reviewers or other factors, enough can be extrapolated from the output to infer how individuals feel about the Barbecue scene in Austin. Generally speaking, and unsurprisingly, people greatly enjoy the barbecue scene in Austin! By taking a simple arithmetic average of the compound scores for each review, I have found the average compound score to be about 0.64. Considering the relatively large sample of reviews and the fact it is recommended that a VADER compound score above 0.05 be interpreted as positive overall, I would say that this average tells us that generally, people do indeed enjoy Austin barbecue at a high level. Of course, it is difficult to say how this stacks up relative to other barbecue scenes without comparative data, but that is outside of the scope of this project.

While sentiment analysis can shed light on how customers feel overall about Austin barbecue, we also want to try to extract what aspects of Austin barbecue are most cited in these generally positive reviews. To do this, I aim to find the most popular words used in the review data that has been gathered. While this task sounds simple, it took a bit of trial and error on the coding front to get it right. In order to get the output that I desired, I first had to clean the review data to ensure that all noise that would inhibit this process was extracted. To do this, I tokenized each review, and filtered the review content to remove stop words as well as all non-alphanumeric tokens since both of these types of tokens would register when measuring token frequency. All words that made it through this filtering process would then be appended into the newly created “new\_rev” list for further use. This was done with the following code:

new\_rev = []

for review in reviews\_df:

words = nltk.word\_tokenize(review.lower())

#print(words)

for word in words:

if word not in stop\_words and word.isalnum():

new\_rev.append(word)

Once this was executed, I finally had a clean body of text to analyze for meaningful word frequency. To do this I used “FreqDist” from the nltk package to count and display the 20 most common words found in the review text with the following code:

frequency = FreqDist(new\_rev)

most\_common\_words = frequency.most\_common(20)

print(f"Most Common Words in the Review Data:{most\_common\_words}")

This code yielded the following output:

*Most Common Words in the Review Data:[('bbq', 11), ('good', 11), ('brisket', 8), ('food', 8), ('place', 6), ('got', 6), ('delicious', 5), ('service', 5), ('great', 5), ('restaurant', 4), ('sausage', 4), ('early', 4), ('texas', 4), ('coming', 3), ('get', 3), ('first', 3), ('people', 3), ('best', 3), ('one', 3), ('like', 3)]*

Some of the most common words in the review data are more telling than others. Words such as “bbq” and “food” are some of the more obvious results; however, others such as “service”, “place”, and “people” shed light on the possibility that people enjoy not only the food that is offered in these restaurants but also aspects of the setting itself.

Based on both the sentiment analysis and word frequency analysis of this sample of 30 reviews from various barbecue restaurants in the Austin area, it is possible to infer that people generally enjoy Austin barbecue, not only because of the food itself, but also because of the setting, welcoming atmosphere, and service that these restaurants have to offer. This can not only be reasonably inferred from the previously mentioned data but also directly from reviews such as “*…the bustling atmosphere and inviting aroma of barbecue had me excited for my very first experience of Texas barbecue…”,* “*Perfect spot to bring out-of-town guests. It's a large restaurant but yet feels cozy.”*, and “*I enjoyed my dinner with my friends at this Texas style BBQ!*

*Since I know too less for BBQ and only lived in one midwest city, Kansas City”.* Among other similar reviews, these may suggest that customers experience or at least perceive a difference in atmosphere/service at these restaurants in line with the commonly touted “Texas Hospitality” stereotype. While I believe that more research with more data would have to be done in order to flesh out and further substantiate some of these inferences, I do think that a barbecue restaurant anywhere could benefit from cultivating the image of having “Classic Texas Barbecue” with a setting and staff training to match. Even just the perception of being an authentic Texas barbecue restaurant could be enough to bring customers wanting to experience the real thing.