# Does Culture Proximity Determine Authenticity?

## A Sentiment Analysis of African Restaurants in the DMV & Bay Areas

By

Joel Olujide

An Undergraduate Thesis submitted in partial fulfillment of the requirements for the JOSEPH WHARTON SCHOLARS

Faculty Advisor:

Samir Nurmohamed

Associate Professor of Management

THE WHARTON SCHOOL, UNIVERSITY OF PENNSYLVANIA

APRIL 2023

**Abstract**

The purpose of this study is to analyze the perceived authenticity of African restaurant reviews in the DMV (Washington DC, Maryland, and Virginia) area and compare them to reviews in the San Francisco Bay area. This perceived authenticity would be determined by analyzing the drivers of a sentiment analysis of the reviews. The hypothesis that the DMV reviews sentiments would be based on food quality, while the Bay area reviews would be based on service and ambience. The results showed that the DMV reviews were indeed more skewed towards food than the Bay area reviews, confirming the initial hypothesis. However, the analysis also revealed that other factors were at play for both metro areas. For example, while the DMV area reviews tended to mention the quality and authenticity of the food, aspects of restaurant service were also shown to be important, and in some cases even exceed the importance of food quality when predicting positive sentiments. For the Bay area reviews, while more were focused on the service aspect of restaurants, food quality and safety were significant factors. Overall, the study suggests that positive gastronomic experiences in African restaurants are not only attributed to one factor, although there are positive skews showcased in different metro areas. The findings have implications for African restaurant owners and marketers in both regions, as they highlight the need to tailor their branding and marketing strategies based on the specific preferences and expectations of their target audience.

Keywords: African, Sentiment Analysis, Authenticity, DMV, San Francisco, Food, Service, Restaurant, Reviews

## Table of Contents

## Introduction:

This paper analyzes how authenticity is perceived pertaining to restaurants given different demographics in otherwise similar areas. In this study, the Washington D.C. Metro Area ("DMV") and San Francisco Bay Area ("Bay Area" or "the Bay") will be the cities analyzed, with African restaurants being the specific ethnicity that is analyzed. This paper fills the gap from previous literature, because although authenticity research for restaurants currently has a substantial amount of progress, there is a lack of literature on African restaurants and perceived authenticity for ethnic restaurants in general. Most of the research in this space has also been outside of the United States, and while that research is helpful, the definition of restaurant authenticity in the United States is important for African restaurants in the United States, and for restaurants in other parts of the world who have a significant number of their clientele from the U.S.

## Background

African cuisine is an increasingly popular dining choice in the United States. As the diaspora grows, and African foods become more mainstream, it has recently become 'trendy' to try foods such as Fufu, Ingera, and Jollof Rice, cuisines that were once exclusively enjoyed by the ethnicities of their creators. According to research from Technomic, a foodservice research and consulting firm, around 50% of consumers in the United States said that they would be open to try Ethiopian, West African, and Moroccan foods, but have not had the opportunity to. This is particularly apparent in areas where the majority of Africans tend to immigrate, such as the D.C. – Maryland – Virginia (DMV) area, the New York City metro area, several cities in Texas, and more where tens of thousands of Africans call home. Like many other ethnicities in the United States, members of the African Diaspora have started their own businesses, including restaurants where several traditional meals are cooked and served to the demand of their community members.

Understanding how African restaurants are perceived by consumers is important because this information will help restaurant managers adjust their establishments to be well-received by their consumers and increase their overall reach. It is important for these restaurants to maximize their reach with customers through location and menu creation because it predetermines the success of said restaurants, which are often hard to maintain, and since most are small businesses, they are susceptible to failure (Parsa et. al). Restaurants are often the entry point for many individuals into the working world, and their success and livelihood are not only crucial for the economy, but also for the job progression of people who rely on restaurants for their employment. Restaurant owners and managers could use the information researched in this thesis to increase their customer base and maximize their profits to better contribute to their broader communities. This is especially important in ethnically diverse cities such as New York, Philadelphia, D.C., and Chicago, where local restaurants are responsible for thousands of residents' livelihoods.

It could be used for specific enclaves of restaurants as well. For example, because this thesis focuses on African restaurants in particular, restaurant managers in locations with large enclaves of Africans could particularly benefit from this thesis, and use trends based on others in similar locations to increase their chances of success. Many restaurants like these are family-owned or operated locally, and often have relationships with religious centers and community organizations. Research obtained through this thesis could help those communities as well maximize their own resources and support job creation, financial sustainability, and further opportunities for communities that benefit from African restaurants.

In this thesis, sentiment analysis will take place for African restaurants in the DMV area and the Bay. Through doing this, this thesis will identify the characteristics of DMV and Bay area restaurants that consumers value most in their positive sentiments, what characteristics that are highlighted for negative experiences, and which words are used in reviews to gauge perceived authenticity. The details of how this analysis will take place will be explained in the 'Methodology' section.

**Literature Review**

This thesis tackles the challenge of perceiving customer authenticity, and doing so builds upon literature on consumer sentiment analysis as well as the history of ethnic restaurants' patronage in the United States. Currently, the majority of literature that focuses on ethnic restaurants does so in countries outside of the United States, or on cuisines from ethnicities outside of the African continent. Despite this, the insights they highlight are relevant to the research question and help to guide the hypothesis formulation. Most notably, literature on authenticity perception and sentiment analysis were integral to help build the models that were used in this research. Below, this research will discuss some of the most important pieces of literature used to gain a broader understanding of ethnic restaurant trends and menu creation. Then, this thesis will highlight literature focused on sentiment analysis and perceived authenticity across ethnicities.

A primary challenge for restaurant owners is deciding where to open their establishments and how to market them to consumers. Work from Song et. al describes the impact of location on perceived authenticity of ethnic restaurants. It also looks into the size of the ethnic enclave in which restaurants are located and observes the effects of authenticity based on the enclave size. Ethnic enclaves are defined as the population for a certain ethnicity that help bolster ethnic restaurants in a community, such as a large South Asian population around Indian restaurants, or a prevalent Korean community near Korean BBQ locations. Their studies found that the ethnic enclave's size significantly impacted the perceived authenticity of restaurants, and the history/ownership of restaurants also impacted the gastronomic experiences as well.

Building upon this research, another question this thesis seeks to answer is what aspects of a restaurant are most significant for positive sentiments or perceived authenticity. These aspects include food quality, restaurant atmosphere and customer service. Wen-Qi Ruan and Shu-Ning Zhang pursued this research on restaurants in China, and considered the differences between the gastronomic experiences of local restaurants compared to foreign restaurants. Their studies found that foreign restaurants' service and atmosphere contributed the most to their brand equity, while for local restaurants, cuisine authenticity and culture were the primary drivers of their success. This is in an area where Chinese people were the ethnic majority, so they were more critical of authentic dishes due to their experience with them. The study also showed that a lack of familiarity with dishes resulted in more positive experiences with the foreign restaurants in the study. While the results from this thesis may differ because of the country the study is being conducted, trends around familiarity and the implications of customers' experiences on restaurant reviews will still be very applicable to this research.

One of the most significant studies that this research will build upon is how authenticity is perceived through online reviews, which is the primary source of sentiment analysis that this research will use. Le et. al conducted research on over 33,000 reviews in Australia on Zomato, where they used rigorous analysis to correlate authenticity terms that customers used to certain restaurant attributes determined by the reviews of the restaurants that they visited. Shown in Figure 1, they used 88 restaurant attributes found using machine-learning tools and building upon the research of O'Connor et al (2017) and Kovács et al (2014). Findings of their studies showed that authenticity stemmed from both the reviewers' personal experiences, as well as the actual attributes of the restaurants themselves. The researchers investigated how customers saw themselves through the experiences of the restaurants they dined in, and how restaurants could analyze their personal perceptions and structure their establishments' gastronomic experiences to appeal to them.

**Data**

The DMV metro area and the San Francisco Bay Area were the locations selected for this study because of their similar demographics aside from their Black and foreign-born populations. The DMV area has a population of 6.25 million people, a median age of 37, and a household income of roughly $106,000. 22.8% of its population is foreign-born, and 24.6% of its overall population is Black/African American only. In comparison, the Bay area has a population of 4.71 million, a median age of 39, and median household income of roughly $110,000. Its foreign-born population is 30.6%, but its Black population is about 7%. Of their foreign-born populations, the states within the DMV area have over 170,000 people from West African countries and Ethiopia, while the entire state of California has a little under 64,000 people from these countries (Data USA).

To analyze perceptions of authenticity in the DMV and Bay areas, a key indicator of sentiments from restaurant customers comes from online reviews. Online reviews are a reliable source of information for sentiment analysis because they are voluntary, which means that they do not suffer from user bias that conventional surveys of restaurant sentiment could create. Moreover, they do not suffer from response bias that surveyors may incur when their respondents are offering responses within proximity of the restaurants that they are reviewing. Most reviews are given primarily when customers have a particularly positive or particularly negative experience, so other incentives to give reviews are mostly avoided as well (Matakos & Tsaparas, p.532). This is because participants submit reviews to either encourage others to dine at a restaurant, or to warn others before entering an establishment. Because they have no response time requirement, customers also have no need to feel rushed in giving reviews and would not give false reviews for the sake of time. Overall, these reviews are free to access to researchers and provide powerful insights on customer sentiments, values, and the attributes of restaurants themselves.

In this study, about 10,387 reviews of African restaurants were analyzed between D.C. and SF, consisting of about 10 restaurants per city. These reviews are dated from early April 2023 to April 2013. The primary ethnicities of the restaurants in both cities consisted of West African Cuisine and Ethiopian cuisine. There were about 4,867 reviews collected in the DMV, and they were sourced from OpenTable, a restaurant reservation and reviews website. OpenTable was very useful in this analysis, because in its reviews, it gives every customer the option to rate the restaurant on three attributes outside of the overall rating: *Food, Service,* and *Ambience*. Reviewers rated each attribute from a scale of 1-5, and this helped to provide insights on which attributes contributed the most to the sentiment of individual reviews. Within the 5,520 reviews of Bay area restaurants, the reviews collected were from Yelp rather than OpenTable. Unfortunately, outside of the written reviews, Yelp reviews only provide the overall rating, which removes the further insights gained from OpenTable reviews. While the goal was originally to use OpenTable for both cities, it seems that Yelp was far more popular for African restaurants in the Bay area than OpenTable was. Similarly, there were significantly fewer Yelp reviews of African restaurants in the DMV than OpenTable reviews, so uniformity in the review platform was unattainable for the purposes for this study. While the use of a different platform could open the results to source bias, Yelp was still used for the Bay area reviews because of the limitations of the data available for the purposes of this research. Moreover, there was no significant reason to suspect that Yelp reviewers have significantly different values from OpenTable reviewers, and the differences between the cities will still be reflected despite the differences in the platform used.

## Methodology

### Hypotheses

The goal of this study is to evaluate the different measures of authenticity that exist for African restaurants in different communities. To do this the following hypotheses is proposed:

**H1: Authenticity measures will be primarily based on food quality and taste in places with a higher proportion of members of the African Diaspora (The DMV area), and**

**H2: that these measures will be primarily based on restaurant experience and ambiance in places with a lower proportion of the African Diaspora (The Bay Area).**

For the purposes of this study, "authenticity" will be measured by how significant food and descriptions of the food of African restaurants in each area are to their overall sentiments. Therefore, sentiments that put less emphasis on food will be perceived as "less authentic". This hypotheses is based on the scholarship of Song et. al, with their research suggesting that the size of ethnic enclaves impacts the perceived authenticity of restaurants. According to their research, this authenticity was also influenced by the history and ownership of ethnic restaurants. Wen-Qi Ruan and Shu-Ning Zhang's research also suggests that ethnicities are more critical of cuisine authenticity from restaurants serving their local cuisines, and this hypotheses believes that this phenomenon will be reflected when measuring positive and negative sentiment drivers in the DMV area compared to the Bay area.

Data Sourcing

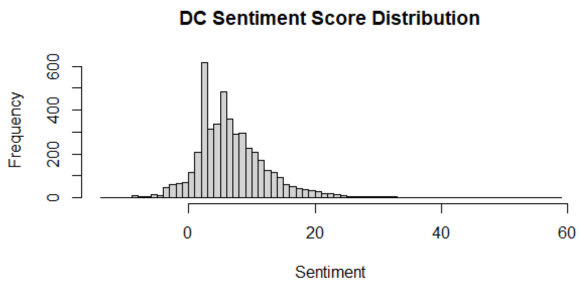
To scrape reviews from OpenTable and Yelp, a free website scraper tool was used, called Octoparse 8. This tool automatically analyzes several attributes of websites through reading the HTML code, and for the purposes of this research, review pages of individual restaurants were imported into the platform. After pulling the review data from each page, unneeded information was pruned from the dataset, resulting in the data needed to be scraped. The tool also has the capability to scrape multiple pages of reviews, so once that capability was set up for each website, a site-wide scrape could take place. Manipulation using this tool was simple because of the similarity of its user-interface to a regular web browser. However, in the case of OpenTable websites, its pagination function would not work because it would not recognize the difference between going forward to the next page, and clicking the back button to go to the previous page. To circumvent this problem, the original website was used to copy the full XPath of the HTML code of the forward button to let the tool know which button to press to scrape more data and to prevent duplicates in the resulting scrape. According to Mozilla, a XPath, which stands for XML Path language, is a flexible method of using non-XML syntax to point to different parts of an XML document. Octopare 8 read the XPaths of different aspects of websites imported into the system in order to perform data scrapes. After solving this problem, Octoparse 8 had the capacity to scrape thousands of reviews in a matter of minutes for both Yelp and OpenTable Reviews.

Sentiment Analysis Methods

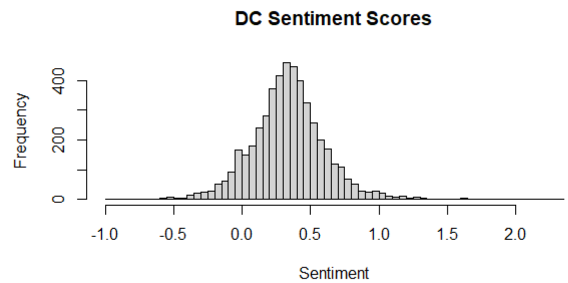
After successfully scraping the review and rating data, it was imported into R for further analysis. R is a powerful statistical and programming language that is often used for data mining and statistical analysis. There are several user-generated packages within the program that are used to perform various functions ranging from logarithmic regressions to plotting violin plots. For the purposes of this research, several packages in R were used, most importantly the "tidytext", "sentimentr", and "text2vec" packages. Initially, the tidytext package was primarily used to perform sentiment analysis on the reviews from the restaurants, using the "AFINN" lexicon as a resource for formulating the sentiment scores. The "AFINN" dataset was created by Finn Årup Nielsen in 2011, and used Twitter data to create a dataset of about 2,477 words which each have an assigned sentiment score ranging from -5 to +5. To analyze the customers' reviews, the tidytext package was used to first tokenize each review by word using the 'unnest_tokens()' function, assign each word a sentiment score from the AFINN lexicon, and then aggregate these scores for each unique review. This resulted in the reviews for each city having a wide range, with DC reviews having a sentiment score range between -14 and 59 across more than 4,600 reviews.

There were two significant challenges with using the AFINN dataset. The first was that the sentiment score recorded was essentially based on the length of the review. For example, if a review said, "The food was good, the lobster was good", it would get a sentiment score of 6 because the word "good" was used twice, and that has an AFINN score of +3. However, if the review simply said, "The food was good", it would get a sentiment score of 3, although it arguably is not a more positive review. This resulted in most reviews with a low number of words having relatively low score, and about 75% of the data having a sentiment score of 10 or lower. The second challenge was that the AFINN lexicon could not recognize negators in sentences. For example, the sentences “The food was good” and “The food was not good” would have the same score of 3 because the word “good” was recognized but would not account the negator “not” within the sentence. This significantly impacted the accuracy of the sentiments that the tidytext sentiment analysis created, and led to the search for a more robust sentiment analysis method that hosts a larger word bank and accounts for more nuanced words in the reviewers’ language.

This is where the ‘sentimentr’ package was discovered and used for the subsequent analysis for the purpose of this thesis. The sentimentr package was developed by Tyler Rinker in 2014, and it is a package that expands upon the tidytext method of sentiment analysis to account for negators in texts, and it calculates the average sentiment score for a length of text. The average sentiment score ensures that longer texts do not get a higher sentiment score simply because of their length. Additionally, the sentimentr package can recognize when amplifiers are used in text, so the phrase “This food is <u>really</u> good” would have a higher sentiment score than the phrase “This food is good”, because it recognizes that the adverb “really” implies a greater positive sentiment for the adjective “good”. This significantly increased the accuracy of the sentiments, and resulted in the data more significantly conforming to a normal bell curve, which made subsequent analysis have increased accuracy in linear regressions. This can be seen below with the DC Reviews:



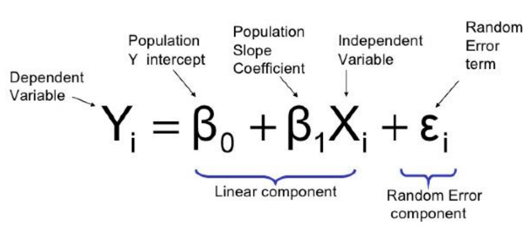

*'tidytext' Sentiment Distribution*




*'sentimentr' Sentiment Distribution*

<u>Linear Regression</u>

After the sentiment analysis from both cities was completed, several statistical methods were used to calculate the sentiment data of each city, and the combined sentiment of both cities. For the DMV area reviews which had the ratings for each reviewer's perception of *Food*, *Service*, and *Ambience*, gaussian linear regressions were run to predict which parameters were most significant in predicting the sentiment score. These formulas were run in R using the 'lm()' function. The gaussian linear regression formula used to calculate predictors for the sentiment score is shown below:




This represents the function used to find the significant predictors of the sentiment score. In the DMV reviews, *Food*, *Service*, and *Ambience* served as the independent variables, and the average sentiment score of each reviewer served as the dependent variable that was fed into the formula. The random error term is used to measure the distance between the observed values and the predicted values of the linear regression model. This shows the model's ability to predict the dependent variable, and the lower this value, the more accurate these predictors could be for the sentiment score.

Logistic Regression

While this method of analysis was particularly helpful with the DMV data, because the Bay area reviews were not scraped from OpenTable, insights on their customers' sentiments on food quality, service, and restaurant ambience were not easily available to analyze. To create a powerful workaround, the 'text2vec' package was implemented to analyze the Bay area reviews, as well as the entire dataset. This package essentially allows users to create a Document-Term-Matrix (DTM) of the reviews, fit a model into that DTM, and then apply that model to new data. A DTM is a matrix that describes the number of times that a word or group of words occurs in a document, or in this case, a review. To fine-tune the review data that would be included in the DTM, groups of words found in many reviews were combined to form 'n-grams', and the vocabulary was pruned to exclude 'stopwords' such as "not", "can't" and "don't" from the analysis, as these words do not have meaning when looking for word significance. This package allowed for the tokenization of each review to be fitted in a logistic regression model, which not only creates a highly accurate model to predict the response variable but can also provide insights into which words hold the most weight in positive or negative sentiments using the 'coef()' function.

To fit the sentiment data into the text2vec package, specifically in a logistic regression, the dependent variable must be a binomial one, meaning that it must have only two options. This was a challenge because the sentiment scores were initially calculated as a range of scores, but to fit within the logistic regression, they had to be converted into a binomial variable. To do this, the reviews that had an associated overall rating of 4 and above were assigned to have a "positive" sentiment, and the reviews associated with an overall rating of 3 and below were assigned to have a "negative" sentiment. With a binomial response variable, the logistic regression was possible, and it was achieved using the 'glmnet' package. Within this package, the 'cv.glmnet()' function was used, which allowed the model to run with an L1 penalty as well as 4 folds of cross-validation that helped to solidify the model's accuracy when used on test data. The data was also split into 'train' and 'test' datasets, where about 80% of reviews were trained in the model, and their validity was tested amongst the remaining 20% of the test dataset. The L1 penalty, also known as Lasso Regularization, encourages the model to minimize the sum of the absolute value of the model's coefficients, as well as minimizing the error between the predicted and actual outcomes. It is denoted as 'alpha' in the 'cv.glmnet()' function. The 4 folds of cross-validation were used to repeatedly run the model on four roughly equal-sized chunks of data to strengthen the model's accuracy. As a performance metric, the AUC curve was used, which is a performance metric to see the strength of a binary regression model. It ranges from 0-1, with scores between 0.5-0.7 considered poor, scores between 0.7-0.8 average, scores between 0.8-0.9 good, and scores above 0.9 considered excellent. Finally, after the model was run, its validity was compared to the test dataset. Code to the cv.glmnet() function can be seen below:

```
glmnet_classifier = cv.glmnet(x = dtm_train, y = train[['Binary_Rating']],
                              family = 'binomial',
                              alpha = 1,
                              type.measure = "auc",
                              nfolds = NFOLDS,
                              thresh = 1e-3,
                              maxit = 1e3)
```

**Results**

Sentiment Data by Area

The DMV area appeared to generally have a higher average sentiment score across their reviews of African restaurants than the Bay area did. As shown in Figure 2, the DMV area has an average sentiment score of 0.33, while the average sentiment score of the Bay area is about 0.26. When looking at sentiment score per overall rating level, the DMV area outperforms the Bay area for all individuals who rated a 4 or 5, but the Bay area had more positive sentiments for individuals who rated a 2 or 3. Interestingly, as shown by Figure 3, amongst individuals who rated a 1 overall, the DMV area had significantly more negative sentiments than the Bay area did. These results suggest that African restaurant dining experiences are more favorable in the DMV area than in the Bay area, and sentiments overall are more extreme both positively and negatively in the DMV.

Linear Regression Results

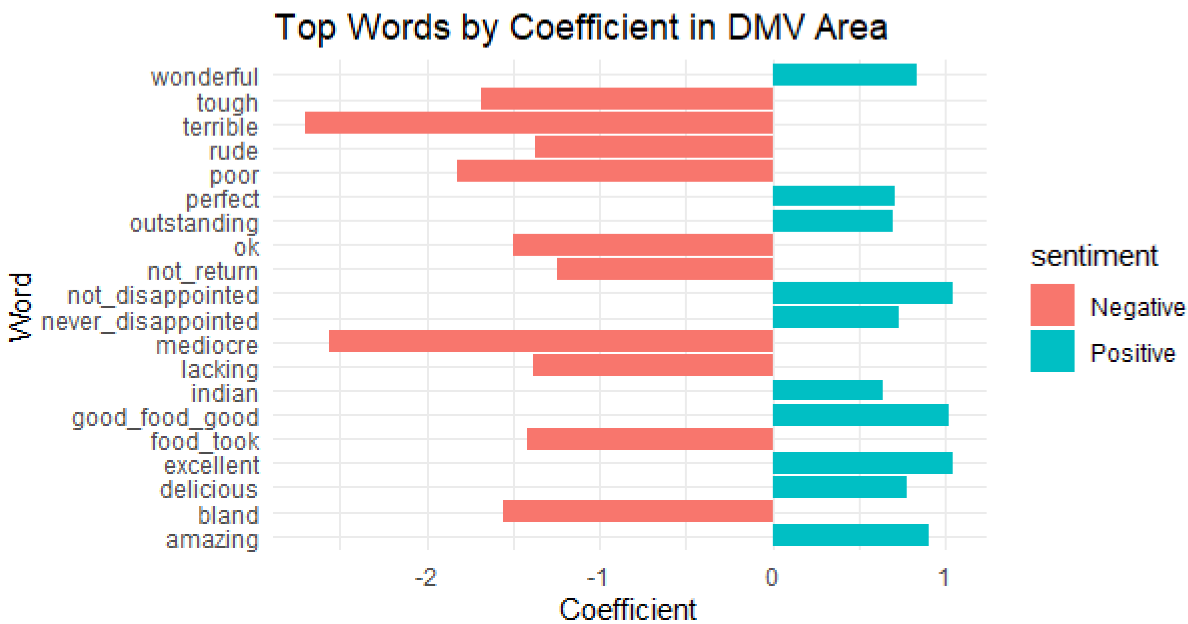
For the reviews in the DMV area, a linear regression was run to see which of the 3 characteristics held the most weight in predicting the average sentiment score. Because each characteristic was a categorical variable with five levels each, results showed which of the 5 levels for each characteristic most significantly predicted the average sentiment score. Results of this data can be seen in the table below:

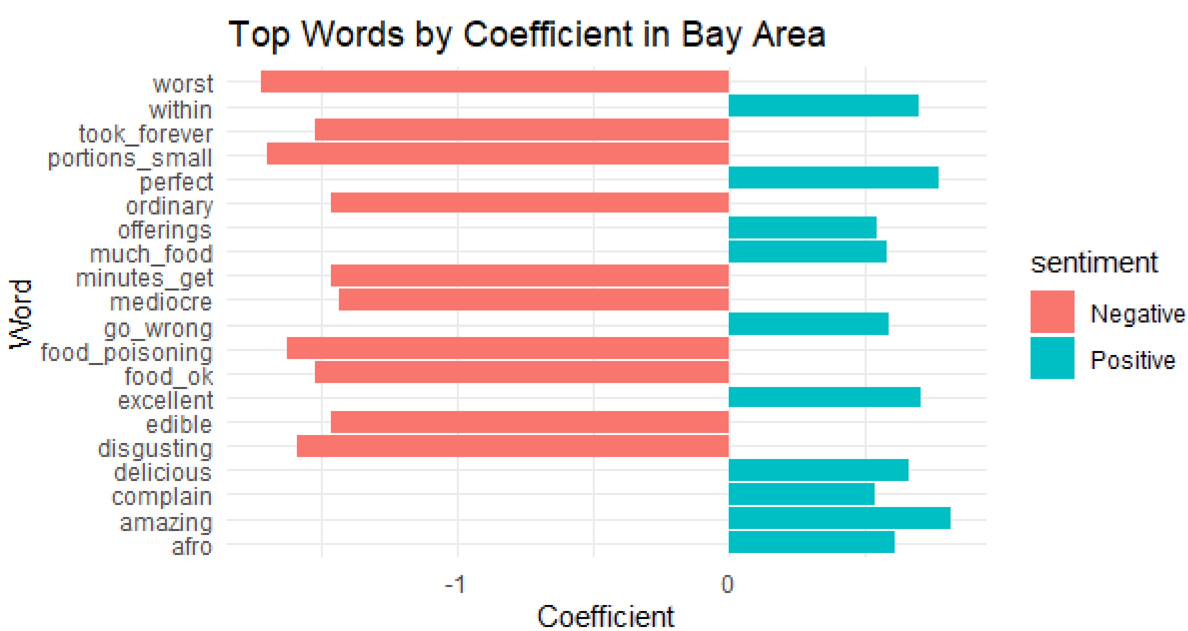
```
Residuals:
     Min       1Q   Median       3Q      Max 
-1.19185 -0.13747 -0.02239  0.11800  1.89129 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)     -0.143226   0.031929  -4.486 7.43e-06 ***
Score_Overall2  -0.007988   0.031128  -0.257 0.797494    
Score_Overall3   0.039301   0.033494   1.173 0.240707    
Score_Overall4   0.168308   0.036872   4.565 5.13e-06 ***
Score_Overall5   0.241806   0.038451   6.289 3.48e-10 ***
Score_Food2      0.030646   0.030487   1.005 0.314842    
Score_Food3      0.059722   0.029709   2.010 0.044461 *  
Score_Food4      0.121685   0.030093   4.044 5.34e-05 ***
Score_Food5      0.104425   0.030589   3.414 0.000646 ***
Score_Service2   0.011497   0.024266   0.474 0.635679    
Score_Service3   0.056321   0.023545   2.392 0.016792 *  
Score_Service4   0.130315   0.024074   5.413 6.50e-08 ***
Score_Service5   0.155108   0.024297   6.384 1.89e-10 ***
Score_Ambience2  0.043040   0.035544   1.211 0.225993    
Score_Ambience3  0.068093   0.032309   2.108 0.035119 *  
Score_Ambience4  0.077002   0.032338   2.381 0.017298 *  
Score_Ambience5  0.062840   0.032705   1.921 0.054738 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2286 on 4849 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.2712,	Adjusted R-squared:  0.2688 
F-statistic: 112.8 on 16 and 4849 DF,  p-value: < 2.2e-16
```

This table shows which characteristic per rating has the highest beta for predicting the average sentiment level. As seen in the categories highlighted, the *Overall Score*, *Service,* and *Food* scores are seen to be the most important in predicting the average sentiment level. These categories are seen to be significant in predicting sentiment only when customers ranked a 4 or 5 in said category. These categories are all statistically significant, meaning that they each have p-values under 0.05. Out of the categories, the *Overall Score* of 5 appears to be the largest predictor in the model, with a beta of about 0.242. Outside of the *Overall Score*, the *Service* score seems to be the next largest predictor, with betas of 0.15 and 0.13 for ranks 5 and 4 respectively. These are slightly higher than the *Food* score's betas which are 0.1 and 0.12 for ranks 5 and 4 respectively. This suggests that H1 may be challenged with *Food* not having the most significant beta to predict sentiment scores. Regardless, this linear regression's adjusted r-squared is about 0.268, meaning that only about 27% of the data can be predicted using the model.

## Logistic Regression Results

The logistic regression showed very impressive validity compared to the linear regression. The AUC of the model of both the DMV and the Bay area was about 0.9079, indicating that the model does an excellent job at predicting the response variable. When the train model was compared to the test data, the AUC was about 0.887 for the DMV, and 0.885 for the Bay area. This means that the model was close at predicting the response variable on new data, showing the validity of the model. The significant insights that were taken from the model were the words and phrases that most significantly predicted a positive or negative sentiment in each metro area. Overall, the coefficients in the DMV area were stronger than the Bay area model, reflecting the higher intensity of sentiments described earlier. The size of a coefficient is determined by the number of times a word or group of words is found in relation to a positive or negative sentiment score. The more times a group of words is found, the higher its coefficient will be. This could suggest the intensity of positive and negative sentiments felt by both areas. A visual representation of the top 10 positive and negative sentiments of each city are shown below:




The top words of both cities provide interesting insights on what contributes to positive sentiments, and what contributes to negative sentiments. In the DMV area, phrases such as "good_food_good" having one of the highest coefficients in model suggest that food quality is a strong predictor in showing a positive sentiment. Moreover, phrases like "not_dissapointed" and “never_dissapointed” suggest expectations that customers had in the DMV which were met by their positive restaurant experiences. While top negative coefficients included “poor” and “terrible”, which suggest negative experiences in the restaurant overall, rather than a specific negative experience with the cuisine. For the Bay area, many of the positive sentiments, while having lower coefficients, are general positive attributes such as “amazing”, “perfect”, and “excellent”. Its negative coefficients are most interesting, with phrases such as “took_forever” and “portions_small” suggesting that negative service experiences contributed the most towards negative sentiments. Both DMV and Bay area positive sentiments included words like “delicious” and “amazing”, showing how there are similarities in what customers value in both areas.

## Discussion

### Sentiment Analysis

As seen in Figure 2, the sentiment analysis data suggests that African restaurants have more positive sentiments in the DMV area than in the Bay area, despite the slightly larger dataset that the Bay area reviews were pulled from. These results could support H1 that authenticity in the DMV area is primarily based on customers’ perceptions of food quality being more positive in the DMV area, but this data does not truly show the driving forces behind the positive sentiments outside of their polarity. Another interesting part of the Figure 3 data is how Bay area sentiments are more positive when customers rated the restaurant a 2 or 3 overall than the DMV area sentiments are. This could suggest a lower tolerance level in the DMV area when it comes to African restaurants. DMV sentiments overall are either very positive or very negative, while Bay area sentiments tend to be relatively positive even after giving a lower rating. This suggests that there are other factors at play in the Bay area that would cause reviewers to still have relatively positive sentiments despite lower overall ratings.

<u>Linear Regression</u>

This data provides the opportunity to dive deeper into the DMV sentiments to understand what factors drive the positive sentiments showcased in Figures 2 & 3. While *Food* is a significant driver according to the regression, its beta is still slightly lower than *Service* as a predictor for average sentiments for ratings above a 4. Lower ratings for all categories were not shown to be a significant driver towards sentiment score, which makes sense because the mean overall rating for DMV restaurants was around a 4.3/5, so most data was in the 4-5 range. This data suggests that H1 could be correct, but it overestimated how food could be secondary to other important factors within a restaurant, most notably restaurant service. The linear regression does suggest that Ambience is not a significant factor for positive sentiments as expected. With an adjusted r-squared of about 0.268, there is a significant amount of variability that is not addressed in this model. This could be due to the polarization of reviews as earlier described, and a linear model not being the best-fitting model to accommodate for that polarization. Regardless, the factors discussed are very significant, which suggests that there are some effects from *Service* and *Food* ratings when predicting the average sentiment score. logistic regression provides deeper insights into the words being used for both areas to understand truly where sentiments and authenticity lie.

<u>Logistic Regression</u>

This is perhaps the most interesting data to analyze to test the validity of the hypotheses. One way to examine the positive and negative coefficients of each city is to compare the terms found in these reviews and compare them to the 88 authenticity terms found by Le et. al. Based on the terms that the researchers found through their Word2Vec models, these terms showcase the perceived authenticity of restaurants to customers when reviewing these establishments online. Beginning with the DMV area, the top three positive coefficients are “good_food_good”, “not_dissapointed”, and “excellent”. The phrase “not_dissapointed” once again suggests that the expectations that diners of African restaurants in the DMV were met by their positive experiences with the food. Having expectations suggests that there is a certain level of experience with African food that diners are entering restaurants with, and then comparing the quality of the restaurants’ foods to their own expectations. This phrase could be matched with the authenticity term “conventional” found by the Word2Vec models. According to the Oxford Languages, ‘conventional’ is an adjective that describes when something is “based on or in accordance with what is generally done or believed”. DMV diners not being disappointed with their restaurant experiences based on their current experiences with the conventions of African foods they are already familiar with could be a strong case towards showing where authenticity truly lies in DMV area sentiments.

Still observing the DMV word coefficients, the top two negative coefficients were “terrible” and “mediocre”. While the word “terrible” is a generally negative sentiment, “mediocre” is interesting because it suggests that diners have enough experience to make judgements on what good African food consists of. As shown in the research by Wen-Qi Ruan and Shu-Ning Zhang, people are more critical of food from their own ethnicity due to experiences with them. It is difficult to give a criticism of a cuisine being ‘mediocre’ without understanding how a cuisine is supposed to taste. Therefore, words like ‘mediocre’ that have very high negative coefficients within the DMV reviews, and phrases like “not_dissapointed” with high positive coefficients suggest that many diners in the DMV area are eating at African restaurants with their own experiences and expectations of African foods, and this either causes them to approve of a restaurant's authentic cuisine, or be critical of a restaurant's culinary choices based on their past experience.

Words pertaining to restaurant service were also found in the top negative coefficients of the DMV reviews. While they had lower coefficients than "mediocre" and "lacking", phrases including "food_took" and "rude" also show how service was a significant factor in the negative experiences of customers rather than in the positive experiences. This suggests that diners visiting African restaurants in the DMV may not have very high expectations of restaurant service, but they still do value positive restaurant experiences and will express negative sentiments when treated rudely or after long wait times for their food.

The Bay area has other interesting words that suggest what many diners in that region value and look out for when eating at African restaurants. While the coefficients are smaller, many of their positive sentiments are general positive sentiments such as "perfect", "excellent", and "amazing". Two in particular that provide interesting insights into what diners in the Bay area value are "much_food" and "offerings", which essentially mean that diners were pleased with the portion sizes offered for their meals or the options available on restaurant menus. This suggests that Bay area diners who had positive experiences at African restaurants were impressed with the portion sizes and menu offerings when compared to other restaurants. This may be an attribute of restaurants that is highly valued in the Bay area but is not mentioned in the DMV area reviews at all.

Valuing service, offerings, and portion sizes are especially shown in the negative coefficients of the Bay area reviews. Some of the highest negative coefficients of include the phrases “took_forever” and “portions_small”, meaning that diners in the Bay area may come with expectations of portion sizes and service times for African restaurants, and it is a pretty big factor into their positive or negative sentiments. Restaurant service overall seems to be a significant factor in the positive and negative sentiments in the Bay area reviews. Terms like “took_forever” suggests that it took a long time for many Bay area diners with negative sentiments to receive their food, be attended to, and more. With “minutes_get” being a phrase receiving a positive coefficient, it also suggests that Bay area diners value the efficiency of time spent in African restaurants as part of their positive or negative experiences.

Food quality is also a significant factor in the negative coefficients of the Bay area reviews. Words such as “disgusting” and “worst” suggest overall negative experiences with African food in the Bay area. However, similar to the DMV area, words including “mediocre”, “food_ok”, “ordinary”, and “edible” suggest that Bay area diners also dine at African restaurants with expectations around the cuisine taste and quality. Another very interesting and slightly concerning phrase that had a very high negative coefficient relative to other words in the Bay area dataset was “food_poisoning”. This brings into question the quality of ingredients, staff, and upkeep of restaurants that Bay area diners are going to, which results in a significant amount of reviewers with negative sentiments experiencing food poisoning. African restaurant patrons having to be worried for their physical health as a result of dining at African restaurants was not a factor considered when pursuing this research. Regardless it is interesting that cuisine authenticity is not the only consideration of food quality for Bay area reviewers compared to DMV reviewers, who did not address this nearly to the same degree as Bay area reviewers.

## Conclusion

According to the sentiment analysis, the linear regression, and most importantly the coefficient analysis from the logistic regression, H1 was generally supported that DMV area African restaurant diners value food quality more than San Francisco Bay area African restaurant diners. This is shown by the coefficient size of food-based phrases attributed to positive and negative sentiments in the DMV area being significantly larger than the coefficients in the Bay area. This is also despite the larger sample size of Bay area reviews that were sampled. Moreover, the evidence of the H1 is shown by the difference in coefficients of the term “mediocre” from other coefficients in the DMV dataset. The negative coefficient for “Mediocre” is almost twice as large as other negative coefficients relating to service-related attributes such as “rude” and “food_took”. This contrasts with the Bay area that has much less variability amongst its coefficients. While food is valued more in the DMV area than the Bay area, there are other very significant restaurant factors, most notably restaurant service, that play a large role in the overall sentiments of DMV restaurants as well. This is shown by the betas of *Service* scores from the linear regression being larger than the betas for the *Food* scores in the model.

H2 was also generally supported that Bay area reviews would most likely value restaurant service over food quality due to a lack of experience via a smaller proportion of the population from West Africa and Ethiopia. However, H2 was incorrect that restaurant ambiance would be a significant factor towards positive or negative experiences, as this was not suggested in the largest coefficients for the Bay area reviews. Like the DMV area, while service-related phrases had some of the most significant coefficients, there was a high emphasis on the importance of food quality in the Bay area as well. In particular, the quality and safety of Bay area African restaurants cuisine was an important factor in determining the potential negative sentiments that diners experienced. There was also less variability between the negative and positive coefficients in the Bay area dataset, suggesting a more level treatment of food and service attributes in the Bay area, with a slight skew towards service attributes.

Overall, patrons of African restaurants in both the DMV and the Bay area show clear priorities when it comes to positive sentiments, and there is considerable overlap in the priorities of their sentiments. As African restaurants and cuisines are becoming more mainstream, understanding the current factors towards positive and negative sentiments of restaurant goers will be very significant in the future as restaurants find more ways to grow their business and effectively serve their customers based on what they value. Positive and negative restaurant experiences are not black-and-white, and restaurants must find ways to balance their strengths and weaknesses to provide the best gastronomic experience possible to their customers and to their community. Applying this knowledge is critical for the survival of current and future African restaurants in these categories.

**Limitations & Further Improvements**

This research performed a robust sentiment analysis of restaurant reviews to gain an understanding of the drivers behind positive or negative sentiments in African restaurants in two different metropolitan areas. As mentioned earlier, one significant limitation was the use of two different platforms for collecting the review data, which could result in some unaccounted bias in the results of my data. This could be improved upon by using other platforms to collect reviews of restaurants, such as Twitter data of individuals' comments on African restaurants, which could be filtered based on location to compare two different areas. Another limitation of this research is that although authenticity was measured by the weight of food quality for positive or negative sentiments, it is very difficult to define a universal term for authenticity in general. Authenticity by nature is a relative definition based on an individual's experiences and preferences. It could be measured in a different way by surveying individuals' perceptions of African food authenticity, and then giving them the opportunity to compare their authenticity measures to restaurant experiences in their respective areas. This could be done in a separate research study with a smaller sample size to show if the effects found in this study are reflected in individuals' definitions of African restaurant authenticity.

There are also potential interaction terms that could be analyzed when seeing the factors that contribute towards average sentiment. These could be pursued in more detail to understand how the different factors towards positive sentiments in African restaurants overlap. A final limitation to this study is the measurement of sentiment scores used. While the 'sentimentr' package was a significant improvement from the tidytext "AFINN" sentiment analysis package, there are more advanced adaptive machine-learning models that could more accurately measure the true sentiment of individual reviews, which could impact the results in research such as this. This would most likely need to be done with computing power that was not available during the research period of this study.

Appendix

Figure 1: 88 Restaurant Attributes found by Le et al.

| | | | | |
|---|---|---|---|---|
| **amateurish** | deceptive | hoax | mistakable | replicate |
| **ambitious** | dinky-di | home-made | modern | ridgy-didge |
| **artful** | ersatz | home-style | moral | scam |
| **artificial** | ethical | homey | native | sham |
| **artisan** | ethnic | honest | offbeat | sincere |
| **assuming** | exotic | house-made | old-fashioned | skilful |
| **authentic** | expert | humbug | original | skilled |
| **bogus** | extroverted | iconic | orthodox | specialty |
| **careful** | fair-dinkum | idiosyncratic | outlandish | traditional |
| **caring** | faithful | imitation | peculiar | true-blue |
| **cheat** | faked | impostor | phony | truthful |
| **classic** | FALSE | inspiring | pretentious | typical |
| **conventional** | feigned | inspired | professional | unique |
| **crafted** | forgery | integrity | pure | usual |
| **craftsmanship** | fusion | inventive | quack | wholesome |
| **creative** | genuine | legitimate | quirky | workmanship |
| **deadset** | heartfelt | master chef | quintessential | |
| **deceitful** | historical | misleading | real | |

Figure 2: Average Sentiment Score per Metro Area

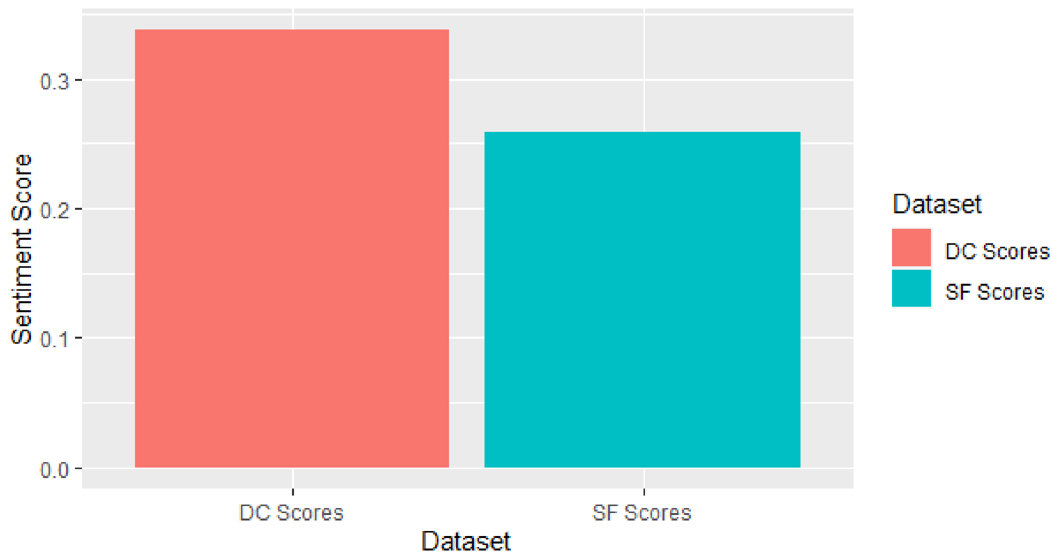
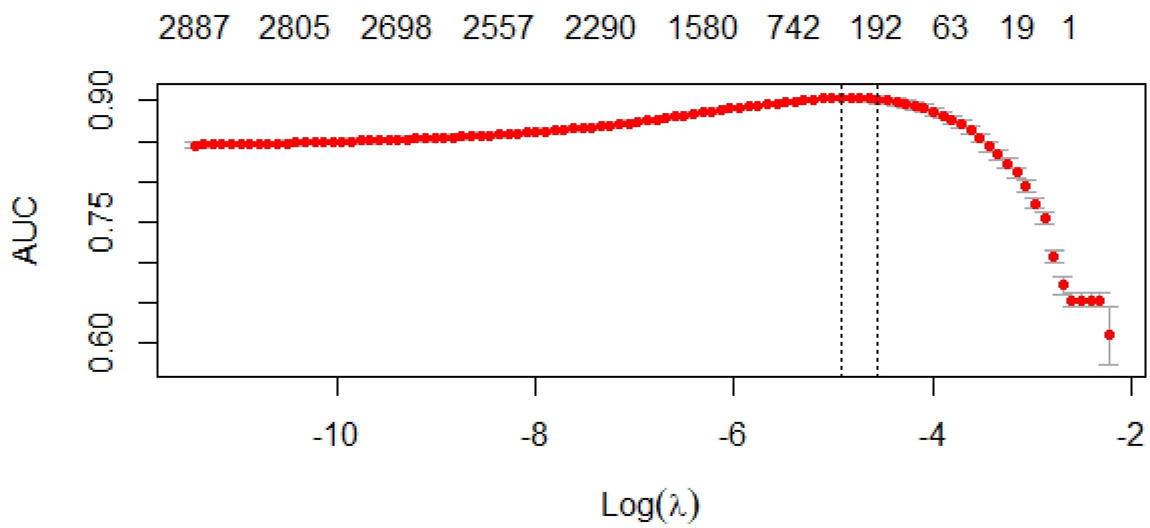



Figure 3: Average Sentiment Score per Metro Area by Overall Rating

Figure 4: ROC Curve of Logistic Regression



## Works Cited


Garwood, Grace. “Is America Ready to Embrace African Cuisine?” *The Food Institute*, The Food Institute, 18 Aug. 2021, https://foodinstitute.com/focus/is-america-ready-to-embrace-african-cuisine/.

Le, Truc H., et al. “How Consumers Perceive Authenticity in Restaurants: A Study of Online Reviews.” *International Journal of Hospitality Management*, vol. 100, Jan. 2022, p. 103102., https://doi.org/10.1016/j.ijhm.2021.103102.

Matakos, Antonis, and Panayiotis Tsaparas. “Temporal Mechanisms of Polarization in Online Reviews.” *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2016, https://doi.org/10.1109/asonam.2016.7752286.

Nielsen, Finn Årup. “A New Anew: Evaluation of a Word List for Sentiment Analysis in Microblogs.” *ArXiv.org*, ArXiv, 15 Mar. 2011, https://arxiv.org/abs/1103.2903.

Rinker, Tyler. “Dictionary-Based Sentiment Analysis [R Package SentimentAnalysis Version 1.3-4].” *The Comprehensive R Archive Network*, Comprehensive R Archive Network (CRAN), 18 Feb. 2021, https://cran.r-project.org/web/packages/SentimentAnalysis/index.html.

Ruan, Wen-Qi, and Shu-Ning Zhang. "Understand the differences in the brand equity construction process between local and foreign restaurants." *Service Business 16*, (2022): 681–719. Accessed September 25, 2022. https://link.springer.com/content/pdf/10.1007/s11628-022-00486-9.pdf.

"San Francisco-Oakland-Hayward, CA." *Data USA*, Datawheel, 2022, https://datausa.io/profile/geo/san-francisco-oakland-hayward-ca#:~:text=Population%20%26%20Diversity,-San%20Francisco%2DOakland&text=In%202020%2C%20there%20were%201.46,third%20most%20common%20ethnic%20groups.

Song, Hanquin, and Jong-Hyeong Kim. "Effects of history, location and size of ethnic enclaves and ethnic restaurants on authentic cultural gastronomic experiences." *International Journal of Contemporary Hospitality Management 34*, (2022): 3332-3352. Accessed September 25, 2022. https://www.emerald.com/insight/content/doi/10.1108/IJCHM-09-2021-1126/full/pdf?title=effects-of-history-location-and-size-of-ethnic-enclaves-and-ethnic-restaurants-on-authentic-cultural-gastronomic-experiences

"Washington-Arlington-Alexandria, DC-VA-MD-WV." *Data USA*, Datawheel, 2022, https://datausa.io/profile/geo/washington-arlington-alexandria-dc-va-md-wv.

"Xpath." *MDN*, Mozilla Corporation, https://developer.mozilla.org/en-US/docs/Web/XPath.