Final Project Report ——

Exploring the Intersection of Income Levels and Culinary Diversity: A Zip Code Level Analysis in the Boston Metropolitan Area

Project Group Member: Yiqian Zheng & Yixin Wang

Research Question:

The culinary landscape in metropolitan areas often mirrors the socio-economic status of residents. In many major cities in the U.S., the unequal disparities between communities of different economic levels are also reflected in the dining environments of these diverse neighborhoods. For example, in higher-income neighborhoods such as West Hollywood, there is typically a greater array of culinary choices, encompassing Italian, Spanish, and Thai cuisines, as opposed to the comparatively limited options in lower-income neighborhoods like East Los Angeles. Thus, members in relatively lower-income areas are limited to sharing the same accessibility to diverse food cultures as higher-income community members, which may hinder further culture integrations.

To take an in-depth look at the potential social disparity issues, we aim to understand the following question by conducting our data analysis:

- **What is the relationship between income level and the number of types of cuisine restaurants in the Boston metropolitan area?**

Data Collection:

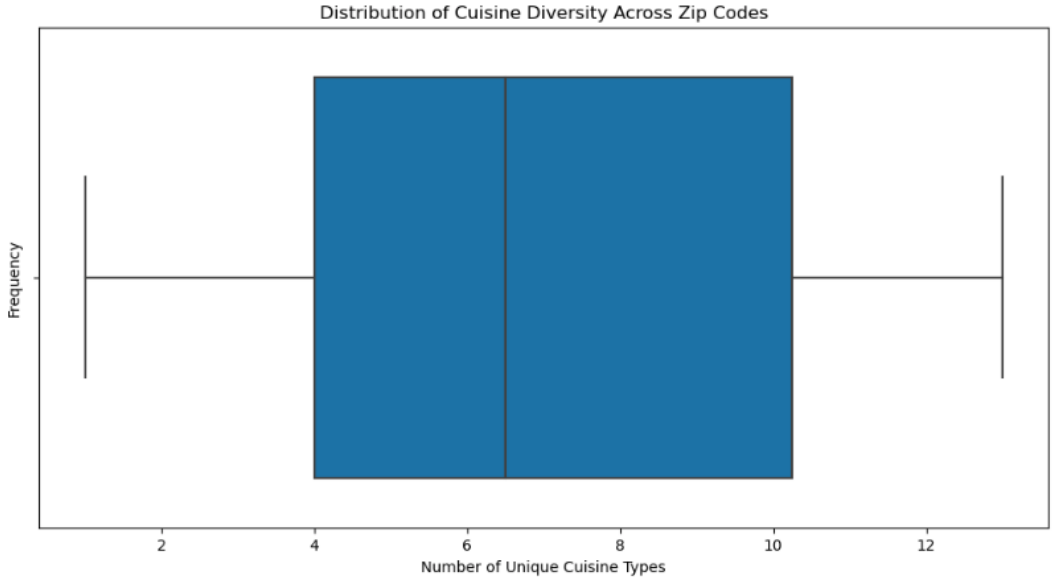
We collected data from two public APIs - Yelp API and the United States Census Bureau's API. We gathered 1,000 sets of restaurant names, cuisine types, and zip code data from the Business Search dataset via Yelp's API. We used the Census Bureau's API to collect Boston residents' demographic data at the zip code level. More specifically, we gathered the population distribution among 10 income levels, categorized by 30 zip codes in Boston.

Initially, we planned to examine if the distribution of cuisines in an area is proportional to its distribution of race and ethnicity. However, it was challenging to find race and ethnicity data down to the exact detailed level as cuisine types have. For example, while there are Chinese, Korean, Japanese, Indian, and other types of Asian food in cuisine datasets, Asians are integrated as one race and ethnicity category in demographic datasets. While cuisine datasets offer granularity by distinguishing between multiple types of Asian cuisines, demographic data lumps these diverse cultures into a singular category. Given these challenges, our focus has shifted to utilizing available income level data, analyzing the dynamics of culinary diversity from an economic perspective.

Data Analysis & Visualization:

We mainly established a correlation analysis to evaluate the relationship between income levels and the number of types of cuisine restaurants.

But firstly, in order to understand the overall cuisine type landscape in the Boston Metropolitan area and to check if there are any outliers within the bigger community, we drew a box plot to show the general distribution of the number of types of restaurants.



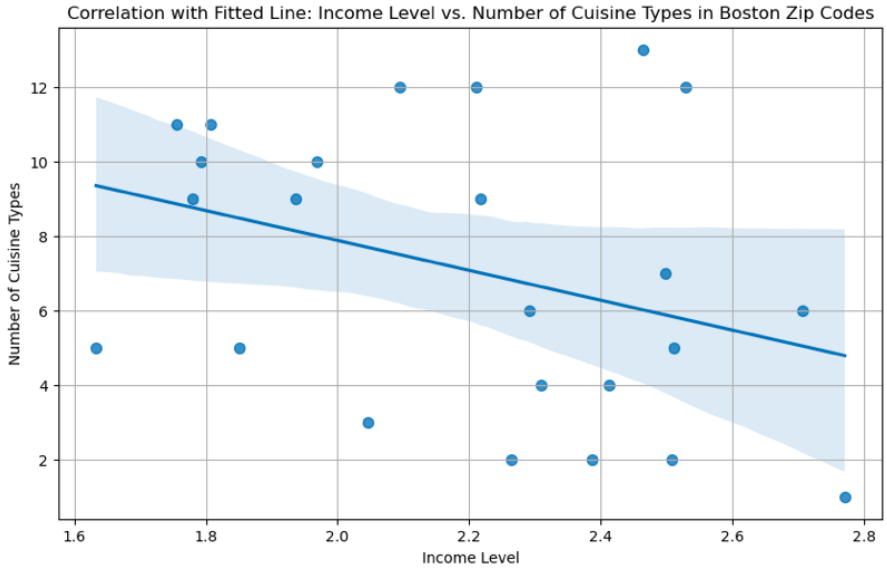
The box plot illustrates the distribution of unique cuisine types across various zip codes within the Boston metropolitan area. With a median value of 6.5, it signifies that, on average, 50% of the metropolitan regions in Boston host more than 6 distinct types of cuisine establishments. The absence of outliers is indicative of the absence of areas with an exceptionally low diversity of restaurant types, reflecting a positive scenario. The Interquartile Range, measuring at 6.25, suggests a relatively consistent distribution in the numbers of cuisine types across the Boston metropolitan community, underscoring an even distribution of cuisine restaurant diversity.

Then, we created a scatter plot to examine the correlation between the number of restaurant types and income levels across different zip codes. We applied the number of cuisine types as the y-axis and a weighted score of income level as the x-axis. Initially, we attempted to use the median income level of each zip code on the x-axis. However, this approach proved ineffective as many zip codes shared the same median income, hindering the visualization of cuisine number distribution.

To differentiate the income level of different zip codes, we assigned a weighted score to each zip code's income level using the following logic:

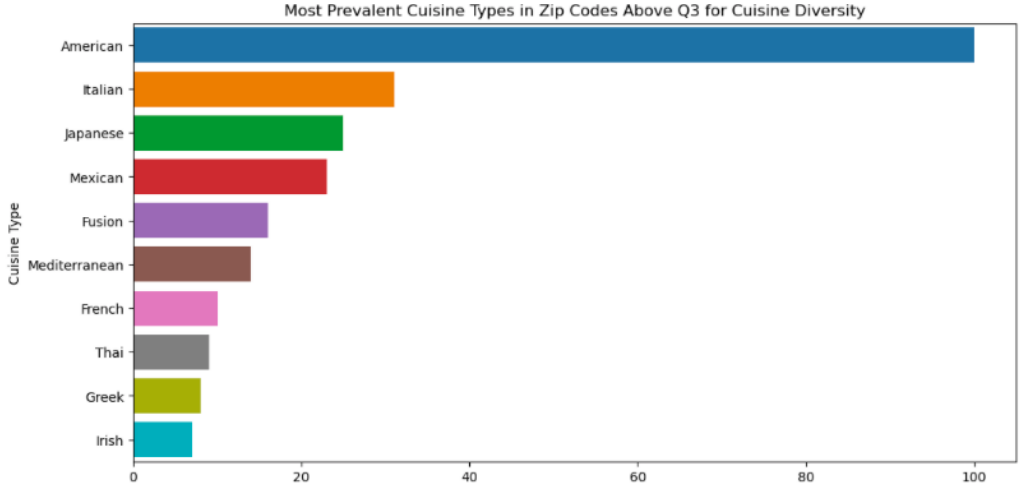
Each income bracket is given a weight that reflects its relative position in the income scale. Lower income brackets receive lower weights, and higher income brackets receive higher weights as higher income brackets represent a greater economic contribution or more affluent

households. We assigned the weight from 1 to 6 to a total of 6 income brackets. The weighted income level of each zip code is calculated by the product of the number of individuals in each income bracket to the weight assigned to that bracket, divided by the total number of individuals. A higher income level indicates a greater proportion of individuals in that zip code fall into higher income brackets and, hence, a relatively wealthier area it is.



Based on the weighting process, we analyzed the correlation between income level and the number of cuisine types of data on zip levels. The negative correlation -0.344 indicates that there is a weak negative correlation between income levels and the number of cuisine types. Different from our expectation that lower-income area has less diverse cuisine options, the result suggests that the lower-income areas have more types of restaurants to access.

In order to understand the potential reason behind this result, we conduct a descriptive analysis of the cuisine categories in the areas with the most and least diverse food options. The x-axis shows frequency of the cuisine types in the top 25% (above Q3) areas with the highest number of cuisines, while y-axis represents the type of cuisine. ‘



In areas with a higher number of types of cuisines, cuisine types vary from different racial origins, including American, Italian, Japanese, Mexican and more. The wide range of cuisine types, spanning various races and ethnicities, suggests a potentially greater racial and ethnic diversity among the residents of these areas.
In areas with limited cuisine diversity, the most commonly found cuisines are American.
However, we think this result is not statistically significant due to a small sample size (only 14 data points), limiting the ability to draw broad conclusions.

Observations & Conclusions:
The analysis indicates a modest negative association between income levels and the number of distinct cuisine restaurants within the Boston Metropolitan area, reflecting that regions with higher incomes tend to exhibit a lower diversity of cuisine types. Overall, the variation in cuisine diversity across different areas in Boston is minor. This observation suggests that diverse economic classes in Boston possess a comparable level of access to varied food cultures.

Considering the reasons behind our observation, the following three factors might explain the negative correlation:
1. **Racial Homogeneity in High-Income Area**: According to a 2021 research conducted by the Federal Reserve System (Federal Reserve System, 2021), White and Asian populations demonstrate higher household incomes compared to other racial groups. This observation implies that areas with higher income levels may be dominated by White and Asian populations, potentially resulting in a relatively homogeneous culinary preference and consequently, a narrower range of cuisine types being offered.
2. **Competitive Landscape of High-End Dining**: Wealthier areas often attract high-end or specialty restaurants, resulting in a fiercely competitive environment. This competition might favor established or mainstream cuisines, thereby limiting the introduction and survival of diverse culinary options.
3. **Economic Factors:** The costs associated with operating in high-income areas (like rent and wages) might also limit the variety of restaurants that can afford to operate there, especially restaurants from a minority culture.

Impacts:
While our findings are partially limited due to constrained data sources and the exclusion of influential factors like racial distribution, the negative correlation implies a potential need for increasing cultural diversity within the wealthier segment of Boston's community. Societal cultural integration is a pivotal factor that fortifies community bonds and stimulates economic growth. Possessing greater resources, financial capabilities, and social influence, the affluent class could play a more impactful role in enhancing our society's cultural integration. Consequently, local governments should advocate for increased participation from wealthier

communities in events related to diverse food cultures, such as cuisine markets, food festivals, and the introduction of a broader range of restaurants in those areas.

Future Improvements:

Due to restrictions in the Yelp API, we could only gather details on 1000 Boston area restaurants, and this didn't cover restaurants in all the city's zip code areas. Given the limits, the current data might not accurately represent the variety of cuisine types in each area, leading to biased results. To address this in the future, we aim to collect more restaurant data using better APIs or web scraping techniques for a more objective view.

Another way to enhance our approach is to consider more factors reflecting accessibility to diverse food cultures. Currently, we only looked at the number of cuisine types to gauge food culture accessibility. Yet, factors like different grocery stores, culinary events, prices, and more also influence accessibility in each community. Including these indicators in our analysis would provide a more well-rounded perspective and more comprehensive results.

Reference:

Board of Governors of the Federal Reserve System. (2021, October 22). Wealth Inequality and the Racial Wealth Gap. https://www.federalreserve.gov/econres/notes/feds-notes/wealth-inequality-and-the-racial-wealth-gap-20211022.htm