**October 2021**

Datasets

1. Nyc\_decennialcensusdata\_2020\_2010

The question/problem

An entrepreneur has hired us as a marketing research firm to find the best place within the city to establish a restaurant to serve the Hispanic/Latino community.

Option A: Where should a restaurant entrepreneur establish a new restaurant to serve the most Hispanic / Latino community members?

Option B: Where should a restaurant entrepreneur establish a new restaurant to earn the highest profits when serving the Hispanic / Latino community in New York City?

Dependent variable

1. Which neighborhood is expected to lead to highest profit for entrepreneur

Potential independent variables

1. Hispanic Population size
2. Neighborhood income

Objectives

1. Find the top 10 neighborhoods with highest hispanic population in each borough
2. Find the Average income by ethnic group

Background about the problem: what phenomena do we observe and what are the key factors to influencing it.

What data we have

Hispanic population by neighborhood

Top 10

Hispanic population by borough

Comparing Hispanic population to other populations within the relevant neighborhoods and boroughs

What data were still looking for:

1. Find the Average income by ethnic group
2. Average income by neighborhood and borough
3. Crime rate by neighborhood and borough
4. Average income by crime rate
5. Infer which socioeconomic groups would we be serving by considering or not considering establishing a restaurant in this area. Also, what are the consequences of that decision?
6. Most popular Hispanic dishes
7. Most expensive Hispanic dishes
8. Least expensive Hispanic dishes
9. Latino community members that eat Hispanic dishes.

Since the data combines the populations of both communities, we need to be able to estimate how this affects our results and analyses.

Predictions (Cause and effect)

Most popular Hispanic dishes

**November 2021**

1. Data:
   1. New York City 2020 Decennial Census Data from: <https://www1.nyc.gov/site/planning/planning-level/nyc-population/2020-census.page>
   2. New York City Restaurant Data last updated October 31, 2021:

<https://data.cityofnewyork.us/Health/restaurant-data-set-2/f6tk-2b7a>

* 1. New York City Real Estate Median Sales data from January 2010 to September 2020 from StreetEasy:

<https://streeteasy.com/blog/data-dashboard/?agg=Total&metric=Inventory&type=Sales&bedrooms=Any%20Bedrooms&property=Any%20Property%20Type&minDate=2010-01&maxDate=2021-09&area=Flatiron,Brooklyn%20Heights#>

* 1. New York City Median Asking Rent Price from January 2010 to September 2020 from StreetEasy:

<https://streeteasy.com/blog/data-dashboard/?agg=Total&metric=Inventory&type=Sales&bedrooms=Any%20Bedrooms&property=Any%20Property%20Type&minDate=2010-01&maxDate=2021-09&area=Flatiron,Brooklyn%20Heights#>

* 1. Fruit and Vegetable Consumption Among Adults in United States, 2015-2018:

<https://www.cdc.gov/nchs/data/databriefs/db397-H.pdf>

* 1. Fast Food Consumption Among Adults in the United States, 2013–2016: <https://www.cdc.gov/nchs/data/databriefs/db322-h.pdf>
  2. Fast Food Intake Among Children and Adolescents in the United States, 2015–2018:

<https://www.cdc.gov/nchs/data/databriefs/db375-h.pdf>

* 1. SOI Tax Stats - Individual Income Tax Statistics - 2018 ZIP Code Data (SOI) from IRS:

<https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2018-zip-code-data-soi>

1. Background about the Problem: An individual wants to open a restaurant to serve the Hispanic community, but does not know which is the best location to achieve his/her goal.
2. Problem Statement: A prospective entrepreneur needs help deciding on a new restaurant location in NYC that would earn the $100,000.00 / year, while serving the Hispanic community.
3. Dependent Variables
   1. Customer income
   2. Demographics of potential consumers
   3. Sales
4. Potential Independent Variables
   1. Borough and zip code of potential consumers with varied amounts of income (assumption: the higher the income, more sales) --- Income of potential customers and Restaurant Sales to these customers
   2. Dishes with Vegetables
   3. Fastfood dishes
   4. Hispanic demographic (assumption: the higher the pop, more likely for more people to consume) --- Restaurant Sales to Hispanic customers. ✓
   5. Similar restaurants in the area (assumption: similar restaurant performance informs our restaurant) --- Hispanic Restaurant Sales
5. Objectives
   1. Which NYC locations have people with higher incomes?
      1. What kind of cuisine would appeal to adults and children?
   2. What type of cuisine would appeal to Hispanic consumers?
   3. What are the top 10 NYC zip codes with the highest hispanic population?
      1. Is there a correlation between high Hispanic demographics and restaurant sales?
   4. How do restaurants that serve Hispanic cuisine perform in the top 10 NYC zip codes?

- Documentation

* Estimated restaurant sales in top 3 neighborhoods or NYC in general.
  + Did not find, but ok.
* What age group and income group buys the most fast food?
  + Approximately 36.6 % of adults (20+) consume fast food according to “Fast Food Consumption Among Adults in the United States, 2013–2016.”
  + Approximately 35.5 % of Hispanic adults (20+) consume fast food according to “Fast Food Consumption Among Adults in the United States, 2013–2016.”
  + Approximately 36.3% of children and adolescents (from 2 -19 years old) consume fast food according to “Fast Food Intake Among Children and Adolescents in the United States, 2015–2018.”
* How much does each income group spend on fast food per transaction.
  + Did not find.

**December 2021:**

Note: FPL = Federal Poverty Level

Tasks:

1. Bar charts top 3 zip codes with Highest number of Taxpayers
   1. Overall and for each income bracket.
   2. For people above FPL.
2. Pie charts of income breakdown of top 3 zipcodes with highest number of taxpayers
   1. Overall and for people above FPL.
3. Top 3 Most populated neighborhoods.
4. Pie charts (income breakdown) - Washington Heights, Jackson Heights, and Corona
   1. Overall and Above FPL.
5. Calculate Mean and Standard Deviation of Tax Payers Earning $75,000+ for the top three neighborhoods.
6. Calculate Restaurant density for top 3 neighborhoods.
7. Prepare Presentation
   1. Discard unneeded data
   2. Fix it
   3. Add supporting documentation.
8. Prepare Report
9. Meet to practice presentation.
10. Submit all documents to professor.