Project 4 Proposal

Amia Graye, Julia Martin, Rosie Quint

The title of your project and the team members names

Food Accessibility Comparison in Baltimore Before and After the COVID-19 Pandemic

Amia Graye, Julia Martin, Rosie Quint

Team Name: APIcons

You should describe a research or data analysis question and explain its importance

You should outline the work you plan to do

How has food accessibility across different neighborhoods in Baltimore changed since 2018? Which groups of people have easier or harder access to healthy food in Baltimore?

The goal of our project is to create a dashboard that builds upon the work conducted in this study conducted by the City of Baltimore Department of Planning and the Johns Hopkins Center for a Livable Future to identify neighborhoods in Baltimore City that are Healthy Food Priority Areas using more current data from Open Baltimore API. The criteria used in their study to classify a neighborhood as a Healthy Food Priority Area is as follows:

1. The average Healthy Food Availability Index (HFAI) score is in the lowest tier

2. The median household income is at or below 185% of the Federal Poverty Level

3. Over 30% of households do not have a vehicle available

4. The distance to a supermarket is more than ¼ mile

Source: (City of Baltimore Department of Planning)

We plan to adopt their criteria but split the neighborhoods in Baltimore into 3 different levels, based on how many of the four criteria they meet. The levels will be as follows:

High Priority: (consistent with the criteria from the Dept of Planning Study) All 4 criteria met.

Medium Priority: 2 or 3 of the criteria met

Low Priority: 0 or 1 of the criteria met.

After dividing up the neighborhoods into the different tiers, we will compare to the 2018 report to see if there are any major differences. We will see if there are any neighborhoods which were previously identified as High Priority that are no longer High Priority or vice versa. We will additionally investigate the possibility of a criteria for the percentage of residents receiving SNAP food stamps, depending on how this affects neighborhood classifications.

Lastly, we will conduct an analysis to compare the demographics of the low, medium, and high priority neighborhoods. We will see if there are any differences based on the age of residents in those neighborhoods (i.e. high proportion of children or elderly people), the race, and level of education. We will use machine learning paradigms such as regression models to explore potential associations between food accessibility and these demographic factors.

This research is important because access to healthy food continues to be an important predictor of health (Franco et al, 2008). Access to healthy food is an ongoing issue in Baltimore, with nearly a quarter of the population living in HFPAs in 2018 (Misiaszek et al, 2018). Moreover, it is likely that access to food has changed since the last study was conducted in 2018, as a result of the pandemic and the cost of living crisis. Grocery stores may have closed, but more importantly, household income and access to vehicles may have changed as a result of the economic hardships and loss of jobs that many people faced during the pandemic. Identifying which areas of Baltimore and which groups are at the highest priority for improving access can hopefully inform city leaders and policy makers of who needs the most help and resources to be able to have access to healthy food, thus creating a healthier Baltimore.

You should summarize work that already exists (if it does)

There is a study done by the City of Baltimore's Department of Planning and Johns Hopkins Center for a Viable Future that outlined priority areas for healthy food insecurity in Baltimore City. (Misiaszek et al, 2018). The identified a "Healthy Food Priority Area" as having: a low Average Healthy Food Availability Index score, median household income at or below 185% of the Federal Poverty Level, over 30% of households have no vehicle available, and the distance to a supermarket is more than ¼ of a mile. We modeled our analysis after this study, but we wanted to look at other criteria for a healthy food priority area, create a scale for healthy food insecurity (not just a binary variable), and showcase how these areas may have changed after the COVID-19 pandemic.

There is also prior research on the association between neighborhood characteristics and food accessibility in Baltimore, which concluded that neighborhoods with higher proportions of black people and low-income people have worse access to healthy food (Franco et al, 2008).

References:

Franco M, Diez Roux AV, Glass TA, Caballero B, Brancati FL. Neighborhood characteristics and availability of healthy foods in Baltimore. Am J Prev Med. 2008 Dec;35(6):561-7. doi: 10.1016/j.amepre.2008.07.003. Epub 2008 Oct 8. PMID: 18842389; PMCID: PMC4348113.

Misiaszek, C et al. Baltimore City's Food Environment: 2018 Report. Available at: https://planning.baltimorecity.gov/sites/default/files/City_FoodEnv_Brief_2018.pdf

You should demonstrate you have access to the data, describe the data, and propose how you will collect the data

Primary Data source: Open Baltimore API

In the data provided by the API, each observation is a different neighborhood, and the variables correspond to the different criteria we are using to categorize the neighborhoods, e.g. % of people with no vehicles.

HFAI per neighborhood: https://data.baltimorecity.gov/datasets/bniajfi::average-healthy-food-availability-index/explore?location=39.295306%2C-76.593379%2C11.11

Scores calculated by the Johns Hopkins Center for a Liveable Future that scores all food stores in Baltimore based on the healthy food they have available. This scoring system was adapted from the Nutrition Environment Measures Survey in Stores.

Percent with no vehicles: https://data.baltimorecity.gov/maps/264376862b824ecfb7cddadc9b265f08/about

"The percentage of households that do not have a personal vehicle available for use out of all households in an area." This was updated in May of 2024.

Grocery store location: https://data.baltimorecity.gov/datasets/baltimore::grocery-stores/explore?location=39.285547%2C-76.608618%2C11.91

Shows the locations of all the grocery stores in Baltimore, including their address, latitude, longitude, and status.

Median Household Income: https://data.baltimorecity.gov/maps/8613366cfbc7447a9efd9123604c65c1/about

Age:

https://www.census.gov/data/developers/data-sets/acs-1year.2019.html#list-tab-843855098

Census API that divides median income variable by four age groups

Race:

Racial Diversity Index: https://data.baltimorecity.gov/datasets/bniajfi::racial-diversity-index-community-statistical-area/explore

"The percent chance that two people picked at random within an area will be of a different race/ethnicity. This number does not reflect which race/ethnicity is predominant within an area."

Percent of Residents All Other Races: https://data.baltimorecity.gov/maps/530645965fa94511833c78ca1e0dfb3a/about

Percent of Residents Black/African-American: https://data.baltimorecity.gov/maps/3b27c89864714b109bde250c628d73e5/about

Percent of Residents White (Caucasian): https://data.baltimorecity.gov/maps/8d70556ecfec4e53b8738b44eceb8/about

Percent of Residents Hispanic: https://data.baltimorecity.gov/maps/bc346d573ee74963beaa8a8b69eb7dfb/about

Percentage of people receiving SNAP:

https://data-bniajfi.opendata.arcgis.com/datasets/bniajfi::percent-of-persons-receiving-snap/explore?layer=0&location=39.284538%2C-76.620524%2C10.84

You should describe the programming paradigms you plan to use and why it makes sense to combine them for your project

We plan to use object oriented programming paradigms to manipulate our data in order to classify each neighborhood into low, medium and high priority neighborhoods.

We plan to use machine learning paradigms such as regression models to explore potential associations between age, race, and food accessibility.

Together, these paradigms make sense for our project because they will allow us to manipulate the data, make inferences using the data, and also will make it easy for us or another researcher to repeat this analysis in the future.

You should describe any packages and/or software you plan to use

httr2 - for extracting data from the Open Baltimore API

Map Baltimore - an R package with neighborhood boundaries that we will use to create maps

Shiny and shinydashboard packages - for creating the dashboard in R

dplyr, tidyr, ggplot2

Caret and tidymodels for machine learning aspect

You should briefly describe the data analytic product you plan to build

We plan to create an interactive RShiny dashboard in which users can select a neighborhood in Baltimore to learn about its food accessibility characteristics and demographics. When users load the dashboard, they will see an interactive map of Baltimore by neighborhood and they will be able to select a neighborhood to learn about its food accessibility and whether or not it is a healthy food priority area. Underneath the map, we will include tables and figures for our comparison of the demographics of the healthy food priority areas and the lower priority areas.

We would also like to include a feature in which users can input their address and our program will return a list of the nearest supermarkets with nutritious food.

4

You should describe a tentative timeline for the proposal

With around 5 weeks to complete this project, this our tentative timeline:

Week 1: Gather all APIs and data sources and extract data so that it is ready to use. This includes any wrangling necessary to join datasets.

Week 2: Classify all neighborhoods into the different priority tiers.

Week 3: Conduct analysis of demographics of different tiers.

Weeks 4-5: Build maps, plots and construct RShiny dashboard

You should describe how the tasks will be split amongst the team members

We will each work on writing code to classify one of the elements of the criteria for determining healthy food priority areas. Then, Rosie will work on the further data analysis of the demographic makeup of the healthy food priority area and Julia and Amia will work on creating and deploying the dashboard, including making necessary maps to put in the dashboard.