The Effects of Food Deserts on Life Expectancy in Baltimore City, MD

**Abstract**

We aim to understand the effects on life expectancy, if any, for residents living in food-desert-vulnerable neighborhoods in Baltimore City as a means to expand on the topic of food deserts and their public health impacts. We ask: What is the difference between the life expectancies of individuals living in food deserts and those in neighborhoods with adequate food access? To answer this, we source mortality data from the National Center for Health Statistics (CDC) for census tracts in Baltimore, and use data from the USDA Food Access Research Atlas for 2015 to identify tracts where residents were considered to be far from a supermarket. The average life expectancy for individuals living in food adequate areas versus food deserts is 73.489 years and 72.515 years, respectively. To test our findings, we use a z-test statistic calculation and find no statistical significance exists between the two groups.

## Background

Food deserts are a source of structural inequality that burdens marginalized, low-income communities. Studying the impacts of food deserts on health outcomes may help researchers, policymakers, and advocacy groups better understand the situation and develop an effective solution. Baltimore City, MD is home to wide spatial wealth disparities due to its history of racial segregation and redlining (Hilton, 2021). The spatial inequality disparities in Baltimore are so pronounced that they have been given a name- the “White L” and the “Black Butterfly.” This Baltimorian phenomenon is commonly referred to in this way because majority White neighborhoods form an “L” shape on the map while majority Black neighborhoods form a butterfly shape on the map (Brown, 2016). This map can be viewed in the appendix.

Since historically, race, socioeconomic status, and opportunity are strongly linked in Baltimore, the “Black Butterfly” neighborhoods are mostly low-income and food-desert-vulnerable neighborhoods (Hilton, 2021). In Baltimore City, a food desert is identified by the following metrics:

a. An area where the distance to a supermarket or supermarket alternative is more than ¼ mile; and
b. The median household income is at or below 185% of the Federal Poverty Level; and
c. Over 30% of households have no vehicle available; and
d. The average Healthy Food Availability Index (HFAI) score for all food stores is low (Mapping the Food Environment, 2018).

These factors compound to make access to healthy and nutritious food disproportionately difficult, and as a result, residents’ health in food deserts suffers. Areas with few supermarkets and stores that stock very limited healthy food items, such as fruits and vegetables, “are also frequently areas with high rates of obesity and chronic, diet-related diseases” (The Public Health Effects of Food Deserts, 2009). This paper aims to expand upon the academic discussion of food deserts. While the impact of food deserts on health outcomes is generally well-known, its correlation with life expectancy is less publicly understood.

## Research Questions

The purpose of this project is to understand the relationship between living in a food desert and its effects on life expectancy in Baltimore City, MD. We ask the following questions:

- *Does living in a food desert correlate to a decreased life expectancy?*
- *What is the difference between the life expectancies of individuals living in food deserts and those of individuals living in neighborhoods with adequate access to food?*
- *How significant are the discrepancies in the data between the two groups?*

## Methods

The first set of data comes from the National Center for Health Statistics of the CDC, which details the average life expectancy in the United States by census tract for the period between 2010-2015. This is the most recent data on life expectancy by census tract. To produce the data, the U.S. Small-area Life Expectancy Estimates Project identifies all the census tracts in the United States as indicated by the Census in 2010 and the ACS population data from 2011 to 2015. To calculate life expectancy, “abridged life tables” were created for all census tracts with a population of at least 5,000. Abridged tables display the mortality information (such as the probability of death) for all ages in 5 to 10 year intervals, except for the first and second year of life, which are kept individually and not compiled into a larger age category (Arias et. al, 2018).

For the experiment, we will specifically be looking at the life expectancy data produced for the 200 tracts in Baltimore City.

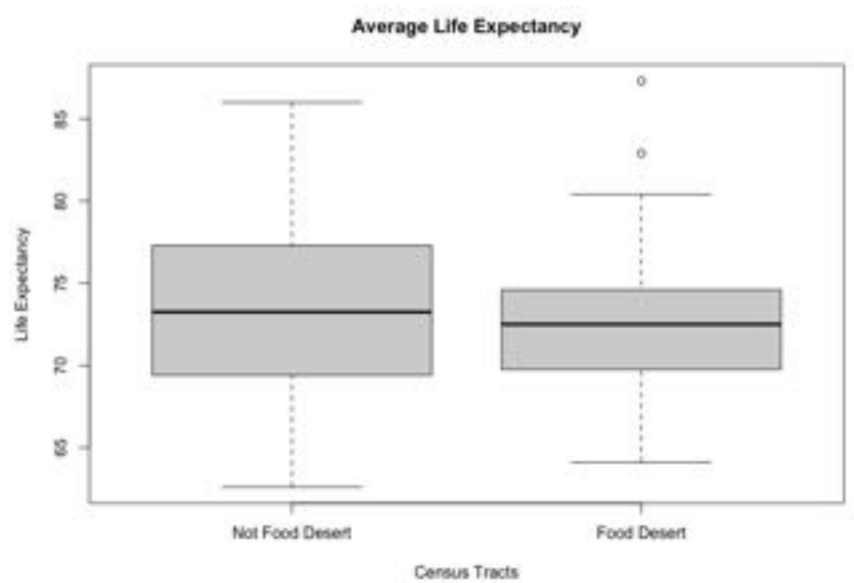
The second set of data is the 2015 Food Access Research Atlas from the Economic Research Service at the USDA. The Atlas was created using data from a list of supermarkets in 2015, the 2010 Census, and the 2010-2014 ACS. With this data, the Atlas highlights census tracts that are low-income and where the residents are considered to be far from a supermarket. Of particular interest to us is the category in the Atlas of “LI and LA at ½ and 10 miles.” This category identifies census tracts where at least 500 people or 33% of the population in the tract are low income and live at least ½ of a mile from a supermarket in an urban area or at least 10 miles in a rural area. The Atlas considers a census tract to be low income if it meets any of the following criteria: there is a poverty rate of 20 percent or larger, the tract’s median household income is no greater than 80% of the statewide income, or the tract is in an urban area and has a median household income of no greater than 80% of the urban area’s median household income (Economic Research Service, 2021).

The first variable in our analysis is life expectancy by census tract. Life expectancy is defined as the average number of years of life per individual in a particular census tract. The second variable in our analysis will be the presence of food deserts. While Baltimore City defines food deserts with the criteria discussed earlier, we chose to use the slightly less restrictive definition of supermarkets ½ mile away from the census tract in order to be able to use the data in the Food Access Research Atlas (as opposed to Baltimore City’s definition of no supermarkets within ¼ of a mile).

To display our data, we use two boxplots. One boxplot shows the data for life expectancy for census tracts in food deserts, and the other shows life expectancy for tracts that are not food deserts. Using boxplots allows us to analyze the various factors that may have been obscured in a different type of graph, such as spread and outliers. The boxplots are generated through R studio.

**Results of the analysis**

The average life expectancy, in years, for people who live in census tracts that are not food deserts is 73.251 years with a standard deviation of 5.157. This group contained 94 samples. The minimum average life expectancy is 62.60 and the highest is 85.8. The 25th percentile life expectancy is 69.35 and the 75th is 76.7.




The average life expectancy for people who live in census tracts that are food deserts is 72.500, with a standard deviation of 3.382. This group contained 82 samples (after removing 2 outliers). The minimum average life expectancy is 64.1 and the highest is 80.4 (after removing an outlier of 87.3). The 25th percentile life expectancy is 69.725 and the 75th is 74.375.

The box plot displaying this data can be seen above. While there were initially two outliers in the food desert census tracts as can be seen in the box plot, these outliers were removed from our data calculations.

Among the census tracts not classified as food deserts, there is a larger spread between the lowest and highest average life expectancies, as can be seen in the box plot. Two census tracts classified as food deserts were outliers. The difference in means between the two groups was 1.043 years.

We want to test if there is a difference in means between the life expectancies of individuals belonging to each sample group. The null and alternate hypotheses are:

$$H_0: u_1 - u_2 = 0$$
$$H_a: u_1 - u_2 \neq 0$$

Our variables are defined as $u_1$ = average life expectancy of census tracts that are food deserts and $u_2$= average life expectancy of census tracts that are not food deserts. To test our hypothesis, a two-tailed z-test statistic calculation for a difference between two sample means was performed (See Appendix). According to this hypothesis test, the difference between the means of census tracts that are and are not food deserts is not statistically significant. Therefore, we fail to reject the null hypothesis, meaning there is not enough statistically significant evidence to suggest a difference between the mean life expectancy of census tracts in food desert versus non-food desert areas.

**Discussion**

The objective of this paper is to determine whether or not there exists a significant difference in the life expectancy of individuals who reside in food deserts versus those who do not. The results determine that there is a much larger range in life expectancy in areas categorized as “Not in Food Desert” compared to the range in “Food Desert.” The hypothesis test also determines that there is not enough evidence to suggest that there exists a statistically significant difference between life expectancy for the two census tracts. Thus, the data from the results is unable to prove that living in a food desert correlates to a decreased life expectancy. This was surprising given how, as discussed earlier Baltimore City has a long history of racial segregation that we anticipated would affect public health otucomes.

One strength of this study is its scope. In order to conduct our test, we used information from every census tract possible in Baltimore City. Census tracts in Baltimore City were only excluded if there were not enough people living in them for the National Center for Health Statistics to determine average life expectancy. One weakness of the study is its lack of generalizability. This paper is limited in determining the correlation between life expectancy and food deserts in a specific area in the United States. The results of this project are likely not broadly applicable towards other cities in the country that are more demographically dense than Baltimore, and therefore likely to have people living closer to grocery stores, or those that are less demographically dense and would require city residents to have a car to travel to grocery stores.

In order to expand upon the discussion of the project, the data on life expectancy may be collected across other major cities. Then by conducting a hypothesis test using the mean life expectancies of “Not in Food Deserts” and “Food Deserts” across the cities, we may determine the relationship between the two variables (life expectancy and living in a food desert) that is applicable to the United States as a whole. This would help federal agencies determine which policies should be implemented or changed to address this public health issue. Other discussions on the impacts of living in a food desert may examine not only life expectancy, but also other public health issues, such as obesity using BMI or the rate of type II diabetes.

In conclusion, we did not find a statistically significant connection between life expectancy and living in a food desert. This suggests that researchers and policymakers focusing on public health and increasing life expectancy may want to consider other aspects besides food deserts in order to understand and improve life expectancy.

**References**


Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski J. U.S. Small-area Life Expectancy Estimates Project: Methodology and Results Summary pdf icon[PDF – 8 MB]. National Center for Health Statistics. Vital Health Stat 2(181). 2018.

*Bringing Economic Justice to Baltimore: How Civil Justice's Economic Justice Project is Fighting Back Against Companies that Prey on Under-Resourced Communities*. Civil Justice, Inc. (2018). Retrieved October 14, 2021, from https://civiljusticenetwork.org/NewsEvents/CJinthenews/EconomicJusticeforBaltimore.aspx.

Brown, L. (2016, June 28). *Two Baltimores: The White L vs. the Black Butterfly*. baltimoresun.com. Retrieved October 14, 2021, from https://www.baltimoresun.com/citypaper/bcpnews-two-baltimores-the-white-l-vs-the-black-butterfly-20160628-htmlstory.html.

Economic Research Service (ERS), U.S. Department of Agriculture (USDA). Food Access Research Atlas, https://www.ers.usda.gov/data-products/food-access-research-atlas/

Hilton, J. (2021, January 25). *Baltimore, a Starving City*. ArcGIS StoryMaps. Retrieved October 14, 2021, from https://storymaps.arcgis.com/stories/1b0e7912e3974c1d9f0a1d671c62d6c5.

*Mapping the Food Environment*. Baltimore City Health Department. (2018, December 5). Retrieved October 14, 2021, from https://health.baltimorecity.gov/food-access/mapping-food-environment.

US National Research Council. (2009). *The Public Health Effects of Food Deserts: Workshop Summary*. National Center for Biotechnology Information. Retrieved October 14, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK208018/.

Source for Figure 1:

American Community Survey; Analysis: Baltimore Neighborhood Indicators Alliance – Jacob France Institute (2018) via Civil Justice, Inc. (Bringing Economic Justice to Baltimore, 2018)

## Appendix

### A.

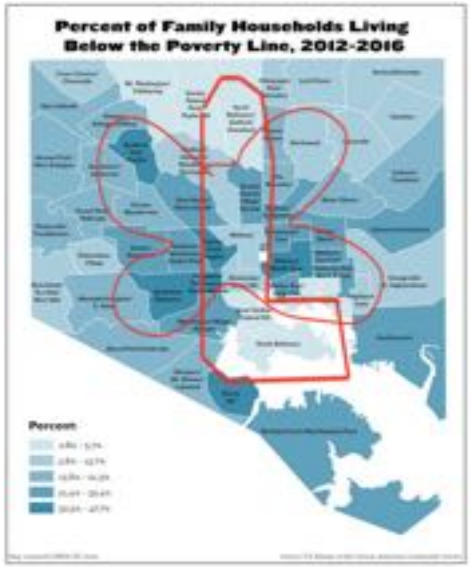



Figure 1. Income inequality in Baltimore City is distributed spatially as illustrated by “the White L and the Black Butterfly” shapes (in red) which are superimposed onto a map showing the Percent of Family Households Living Below the Poverty Line.

**B.** R Code For Box Plot:

a. FDLE<- data.frame(value=c(70.3, 73.4, 68.1, 75.7, 68.8, 77.1, 74, 75.2, 72.5, 68.7, 66.2, 69.7, 75, 71.7, 73.1, 67.5, 71.9, 73.7, 70.2, 70.8, 73.7, 72.7, 68.6, 67.8, 64.1, 71.8, 71.4, 69.5, 69.2, 70.8, 71.3, 74.3, 67.9, 69.9, 74.7, 75.9, 74, 71.2, 65.6, 71.2, 74.3, 73.5, 69.3, 72.1, 68.1, 70, 74.2, 72.7, 72.7, 69.5, 69.4, 69.8, 70.7, 73.1, 72.7, 76.9, 65, 72.3, 73.2, 76.9, 74.9, 74.3, 77, 75.7, 77.7, 72.5, 72.9, 76.3, 72.5, 74.5, 79.4, 67.7, 71.6, 87.3, 82.9, 80.4, 75.7, 78.8, 71.8, 75.3, 74.4, 76.7, 68.4, 67))

b. NFDLE<- data.frame(value=c(77.4, 71.4, 73.4, 76.5, 79, 76, 80.6, 76.7, 74.5, 70.6, 71.2, 65.9, 62.6, 73.1, 66.4, 69.6, 62.8, 65.1, 68.3, 64.6, 74.6, 65.2, 73, 85.8, 78.9, 67.9, 73.5, 76.7, 68.1, 76.7, 69.1, 72.2, 80.5, 74.3, 79.9, 78.3, 71.7, 69.2, 71.7, 67.9, 69.5, 71.4, 66.1, 70.4, 69.3, 73.1, 71.1, 65.6, 76.4, 65.3, 68.6, 64.3, 63.6, 72.1, 68.7, 69.7, 72.1, 66.7, 67.4, 80.1, 72.5, 72.4, 78.7, 80.7, 76.3, 72.6, 70.4, 78.7, 71.2, 75.4, 74.3, 80, 77.2, 75.7, 73.5, 76.7, 75, 72.2, 73.1, 76.1, 76, 74.5, 80.5, 80.8, 82.2, 81.3, 80.1, 73.8, 69.3, 77.6, 81.4, 78.5, 78.6, 75.9))

c. boxplot(NFDLE$value, FDLE$value, ylab="Life Expectancy", main="Average Life Expectancy", xlab="Census Tracts", names = c("Not Food Desert", "Food Desert"))

**C.** R Code For Mean, Standard Deviation, and Quartile Values:

a. FDLE<- c(70.3, 73.4, 68.1, 75.7, 68.8, 77.1, 74, 75.2, 72.5, 68.7, 66.2, 69.7, 75, 71.7, 73.1, 67.5, 71.9, 73.7, 70.2, 70.8, 73.7, 72.7, 68.6, 67.8, 64.1, 71.8, 71.4, 69.5, 69.2, 70.8, 71.3, 74.3, 67.9, 69.9, 74.7, 75.9, 74, 71.2, 65.6, 71.2, 74.3, 73.5, 69.3, 72.1, 68.1, 70, 74.2, 72.7, 72.7, 69.5, 69.4, 69.8, 70.7, 73.1, 72.7, 76.9, 65, 72.3, 73.2, 76.9, 74.9, 74.3, 77, 75.7, 77.7, 72.5, 72.9, 76.3, 72.5, 74.5, 79.4, 67.7, 71.6,80.4, 75.7, 78.8, 71.8, 75.3, 74.4, 76.7, 68.4, 67)

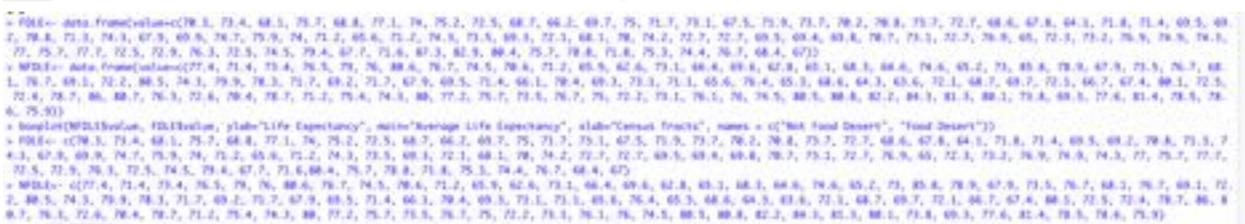
b. NFDLE<- c(77.4, 71.4, 73.4, 76.5, 79, 76, 80.6, 76.7, 74.5, 70.6, 71.2, 65.9, 62.6, 73.1, 66.4, 69.6, 62.8, 65.1, 68.3, 64.6, 74.6, 65.2, 73, 85.8, 78.9, 67.9, 73.5, 76.7, 68.1, 76.7, 69.1, 72.2, 80.5, 74.3, 79.9, 78.3, 71.7, 69.2, 71.7, 67.9, 69.5, 71.4, 66.1, 70.4, 69.3, 73.1, 71.1, 65.6, 76.4, 65.3, 68.6, 64.3, 63.6, 72.1, 68.7, 69.7, 72.1, 66.7, 67.4, 80.1, 72.5, 72.4, 78.7, 80.7, 76.3, 72.6, 70.4, 78.7, 71.2, 75.4, 74.3, 80, 77.2, 75.7, 73.5, 76.7, 75, 72.2, 73.1, 76.1, 76, 74.5, 80.5, 80.8, 82.2, 81.3, 80.1, 73.8, 69.3, 77.6, 81.4, 78.5, 78.6, 75.9)
c. quantile(NFDLE)
d. quantile(FDLE)
e. mean(FDLE)
f. mean(NFDLE)
g. sd(NFDLE)
h. sd(FDLE)

**D.** R Code for Z test:

a. ztest <- (mean(NFDLE)-mean(FDLE)-delta)/sqrt(var(NFDLE)/94 +var(FDLE)/82)

**E.** Screenshots of R Code:


a.


b.


c.

d.

e.

| | |
|---|---|
| zscore | -0.375652742008357 |
| ztest | 1.60387965693205 |

**F.** Math behind the difference of means hypothesis test:

1. Null Hypothesis $H_0$: $u_1 - u_2 = 0$, Alternate Hypothesis $H_a$: $u_1 - u_2 \neq 0$
2. *Distribution: normal (population standard deviation is known and n>30)*
3. *Z-Test Statistic:*

   $((u1) - (u2)) - 0/\sqrt{}((sd1)^2/n1 + (sd2/n2)$ = *z-score*

   $((73.251) - (72.208)) - 0/\sqrt{}((5.157)^2/94 + (3.382)^2/82)$ = *1.6038*
4. *P-value = P(z<-1.6038) + P(z>1.6038) = .0548 + (1-.9452) = .1096*
5. *a = 0.05*
6. *Do not reject null hypothesis*
7. *We do not reject the null hypothesis because p > a*
8. *There is not enough evidence to suggest that there is a statistically significant difference between the means of census tracts that are and are not food deserts*