

# hw3\_thirdtry

Julia Bright

2025-03-05

## Setup and data cleaning

```
library(tidyverse)
library(sf)
library(ggspatial)
rm(list=ls())
getwd()

supermarkets <- st_read("plan372_hmks/hw3/hw3-data/triangle_supermarkets_osm.shp")
streets <- st_read("plan372_hmks/hw3/hw3-data/triangle_highways_osm.shp")
blocks <- st_read("plan372_hmks/hw3/hw3-data/orange_durham_wake_block_groups.shp")
census <- read_csv("plan372_hmks/hw3/hw3-data/triangle_census.csv")

blocks <- st_transform(blocks, crs=4326) #match other projections
blocks$GEOID <- as.numeric(blocks$GEOID) #format same as census GEOID

blockscensus <- left_join(blocks, census)
blockscensus <- blockscensus %>%
  filter(COUNTYFP==183) #filter to only wake county
st_crs(blockscensus)

marketsbuffer <- supermarkets %>%
  st_buffer(dist = 1609) #buffer supermarkets to 1 mile (1609 meters)

wake <- st_join(blockscensus, marketsbuffer) #join blocks, census data, and buffered
↳ markets
```

## Map of non-food deserts in Wake county

```
wake$fooddesert <- ifelse(is.na(wake$osmid), "Food Desert", "Not Food Desert") #create
↳ dummy variable to categorize

wake_major_roads <- streets %>%
  filter(highway %in% c("motorway", "trunk", "primary")) #filter to just major roads

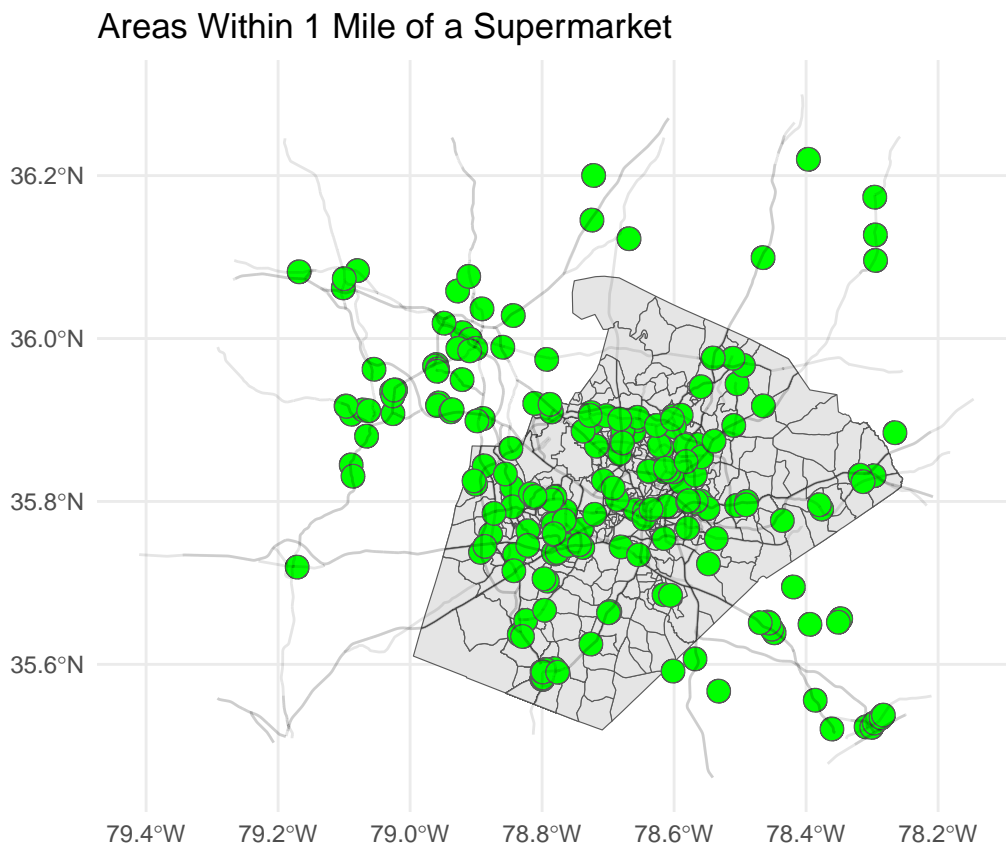
#p <- ggplot() + geom_sf(data=wake, aes(fill=fooddesert)) +
  #geom_sf(data=supermarkets, size=.3) +
  #geom_sf(data=wake_major_roads, alpha=0.1) +
```

```

#theme_minimal() +
#labs(fill = "Supermarket Accessibility")
#This turned out to be more like what's being asked in question 2, using a dummy
  ↳ variable to show which blocks are food deserts

p <- ggplot() + geom_sf(data=wake) +
  geom_sf(data=marketsbuffer, fill = "green") +
  geom_sf(data=wake_major_roads, alpha=0.1) +
  theme_minimal() +
  labs(title = "Areas Within 1 Mile of a Supermarket")
p

```



## Map of census blocks considered to be food deserts

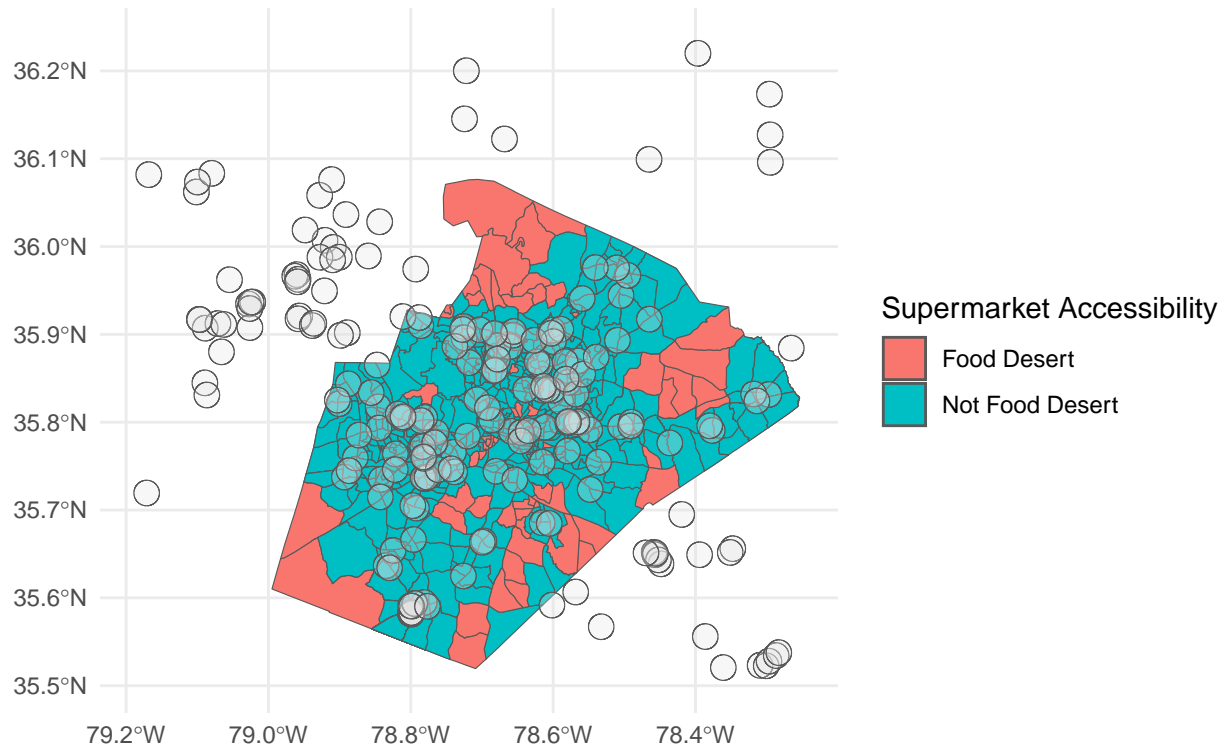
I have chosen to include blocks with partial supermarket coverage in the “Not Food Deserts” category because it indicates at least that those in the census block are still relatively closer to supermarkets than those blocks without any. While there are a couple example of small overlaps but a majority of the block being empty, this decision allows us to focus on the areas that may be most in need. This may also suggest that census block areas are not an effective way to measure food deserts.

```

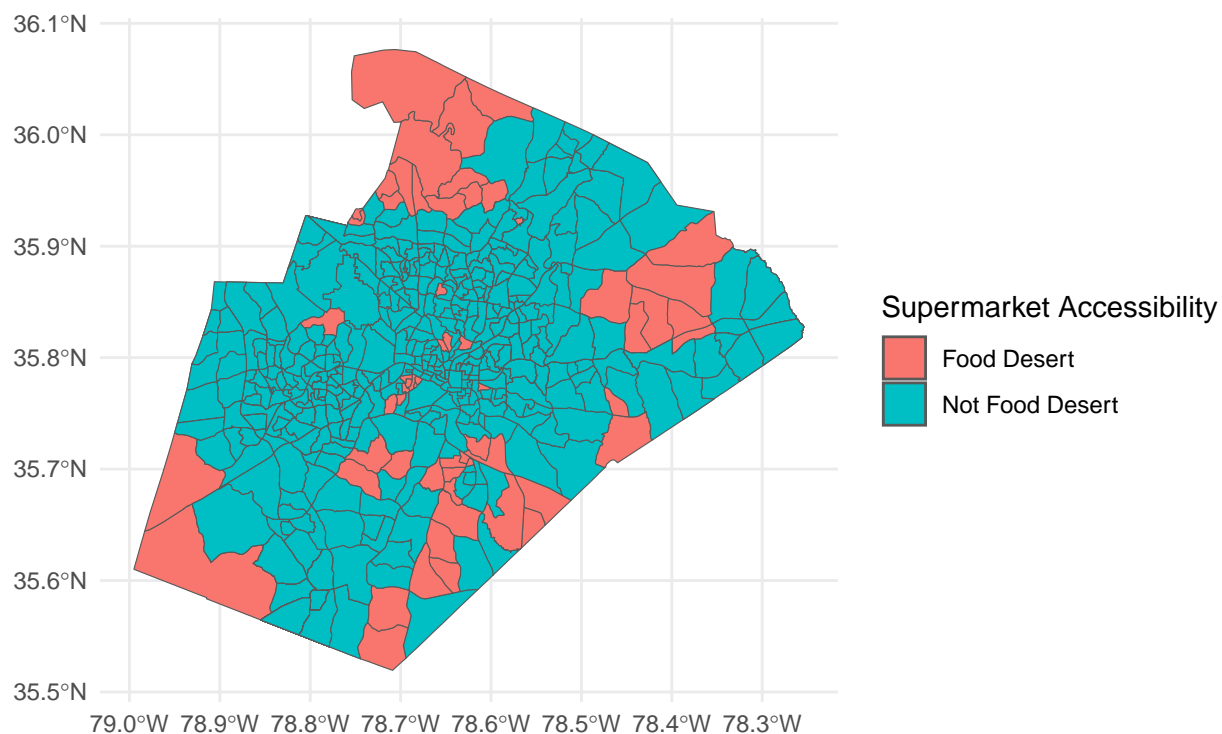
ggplot() + geom_sf(data=wake, aes(fill=fooddesert)) +
  geom_sf(data=marketsbuffer, alpha=0.3) + #Overlay the market buffers to see the extent
  ↳ of overlap with blocks considered not food deserts vs the area within those blocks
  ↳ not in range

```

```
theme_minimal() +  
labs(fill = "Supermarket Accessibility")
```



```
p2 <- ggplot() + geom_sf(data=wake, aes(fill=fooddesert)) +  
  theme_minimal() +  
  labs(fill = "Supermarket Accessibility")  
p2
```



## Demographic Analysis

9.8% percent of Wake County residents, using our definition of food deserts by census block boundaries, live in food deserts. In all of Wake County, 3.96% of households do not own a vehicle, and 18.8% of households are designated as “low income.” Within food deserts specifically, 3.9% of households do not own a vehicle, and 20.8% of households are designated as “low income.”

```
wakeFD_pop <- wake %>%
  distinct(GEOID, .keep_all = TRUE) %>%
  group_by(fooddesert) %>%
  summarize(total_pop <- sum(total_population), na.rm=T)
104925/1069079 #9.8% of residents
```

```
## [1] 0.09814523
```

```
wake_pop <- wake %>% #filter out duplicates to get true household/population size
  distinct(GEOID, .keep_all = TRUE)

zerocars <- wake_pop %>% #number of zero vehicle households
  group_by(fooddesert) %>%
  summarize(sum(zero_vehicle_households))

#Overall percentage of households that don't own a vehicle: 3.96%
(sum(zerocars$`sum(zero_vehicle_households)`)) / (sum(wake_pop$total_households))
```

```
## [1] 0.03967794
```

```
#Percentage of households in food deserts that don't own a vehicle: 3.9%
(zerocars$`sum(zero_vehicle_households)`[zerocars$fooddesert=="Food Desert"]) /
↪ (sum(wake_pop$total_households[wake_pop$fooddesert=="Food Desert"]))
```

```
## [1] 0.03900837
```

```
lowincome <- wake_pop %>%
  group_by(fooddesert) %>%
  summarize(sum(households_income_less_than_35k))

#Overall percentage of low income households: 18.8%
sum(wake_pop$households_income_less_than_35k) / sum(wake_pop$total_households)
```

```
## [1] 0.1880017
```

```
#Percentage of households in food deserts that are low income households: 20.8%
sum(lowincome$`sum(households_income_less_than_35k)`[lowincome$fooddesert=="Food
↪ Desert"]) / sum(wake_pop$total_households)
```

```
## [1] 0.02083604
```

## Location of subsidized supermarket

The map below provides us with a spatial distribution of the poverty rate within census blocks we have designated as food deserts. The poverty threshold in this analysis are those making \$35,000 or less, as determined by the variable, 'households\_income\_less\_than\_35k'. The percentage of households in poverty within each block, or poverty rate, is displayed on a scale of cold to hot colors, with hotter colors being a higher poverty rate. We can reasonably assume that each census block holds a roughly equivalent population, meaning that the very small census blocks are much denser areas than the much larger census blocks. The small census blocks in the center of the map with greater poverty rates are located close to or within downtown Raleigh.

People in these dense, urban areas travel by public transit, biking, and walking at much higher rates than those living in suburban or rural areas. Considering this difference in transportation modes, as well as the increased population when taking into account several adjacent dense census groups in the city compared to a single rural census group, means that placing the subsidized supermarket in an urban food desert would be more effective at improving food access. This is firstly because placing the supermarket within the cluster of counties with high poverty rates will ameliorate the issue of food access for a much greater number of people, and secondly because a higher percentage of those people will be walking, biking, and using public transit to get around.

To follow up on this decision and get a better idea of the impact in regard to modes of transit, I would recommend an analysis of car ownership and transit modes within the population surrounding the proposed site for the subsidized supermarket. It is tempting to jump to the conclusion that the people in this denser urban area have lower rates of car ownership and would benefit even greater from a local supermarket, but we have not performed an analysis here to justify that conclusion. While the rate of car ownership in food deserts was basically the same as outside of food deserts, I would suspect a deviation within this specific area which is dense and urban, experiencing high rates of poverty, and designated as a food desert.

```

fooddeserts <- wake_pop %>%
  filter(fooddesert=="Food Desert")

fooddeserts$poverty <- fooddeserts$households_income_less_than_35k /
  ↪ fooddeserts$total_households

p3 <- ggplot() + geom_sf(data=fooddeserts, aes(fill=poverty)) +
  theme_minimal() +
  scale_fill_viridis_c(option="turbo") +
  labs(title = "Percentage of Households in Poverty within Food Deserts", fill="Poverty
  ↪ Rate")
p3

```

