

Association of Food Environment Variables and Covid 19 Outcomes Driven Locally and Nationally by Population and Median Household Income

Benjamin Romanjenko

Advisor- Rachel Watson

12 May 2021

Contents

Abstract	1
Acknowledgments	2
Background	2
Introduction	3
Project Summary	3
Project Objectives	4
Methods	4
Nutritional Advantage: Wyoming	6
National	11
National vs Wyoming Trends	19
	19
	19
	20
	20
	22
	22
The IRB Process	24
Conclusion	25
References	25

Abstract

Initiated in the Microbiology Capstone class, research began as a team of undergrads interested in the relationships between COVID-19 outcomes and food environment variables. Through the gathering of open source COVID-19 and food environment data, we developed data literacy through preliminary work. Focusing on Wyoming county data, we determined positive correlations between COVID-19 outcomes and farmers markets, Supplemental Nutrition Assistance Program Stores (SNAPS), Women, Infants and Children (WIC) Stores, and percent diabetes. I continued the analytic work, expanding data to include county level data for the United States. Using structural equation modeling I identified population and median household income to be drivers of both COVID-19 outcomes and food environment variables. Understanding the national trends, I reassess the findings at the Wyoming county level.

Acknowledgments

My appreciation goes to the Nutritional Advantage Team for the collaborative environment in which this project originated and giving feedback and input to where our groups work and my own personal work could be directed. I would like to thank Rachel Watson, my professor, adviser, ski coach, for giving us this challenge of redefining what a course could look like. Most of all I want to thank Ella DeWolf for the many hours spent walking me and my team through how to code in R, in addition to teaching yourself and us the basics of R shiny such that we could all share our works publicly.

Background

The COVID-19 pandemic has left no community unaffected, but the effect has been disproportionately directed towards communities of lower socioeconomic class and increased racial diversity. The populations strongly affected by COVID-19, historically, are also affected by food insecurity, and continue to suffer from failing food systems throughout the current pandemic (Alcendor, 2020). Neighborhoods already residing in food deserts, continued to see a drop in access; store shelves became bare of non-perishables as hoarding increased with the notion of home quarantining being imminent (Coleman et al., 2019). These same communities are also associated with poorer health conditions that have been linked to diets based in poor food system environments (CDC, 2020). Diet continues to be an important factor for health in regard to the COVID-19 infection, with it having an indirect effect on the concentration of ACE2, the receptor the viral spike protein binds to, thus affecting the ability to contract COVID-19 and the severity of said infection (Alcendor, 2020, Kassir, 2020). Wyoming, being a rural state is no exception to failing food systems. With the majority of the state being considered a food desert, the citizens have limited access to short supply chain, and healthy foods (Love et al., 2018). Often with heavy reliance on convenience stores, rural populations have limited to no access to fresh produce, and healthy food options. Considering these food environment aspects, and their effect on health, one could expect where there is poorer access to nutritional food there would be an increase in number and severity of the COVID-19 outcomes.

```
#Reading in food insecurity data
insecurity <- read.csv("The Impact of the Coronavirus on Food Insecurity v1.csv")
hist(insecurity$X.C._2020_Food_Insecurity_., breaks = seq(0.06,0.36,0.005),
     xlab = "Percentage of Food Insecurity 2020", ylab="Counts", main = "")
abline(v = mean(insecurity$X.C._2020_Food_Insecurity_.), col = "red", lwd=4)
abline(v= mean(insecurity$X.C._2020_Food_Insecurity_.[insecurity$State == "WY"]),
      lwd = 4, col = "darkblue")
legend("topright", c("Mean", "Wyoming", "United States"),
      col= c(NA,"darkblue", "red"), lwd=10)
abline(v= quantile(insecurity$X.C._2020_Food_Insecurity_., probs = c(0.05, 0.95)))
text(0.232, 170, "0.95*")
```

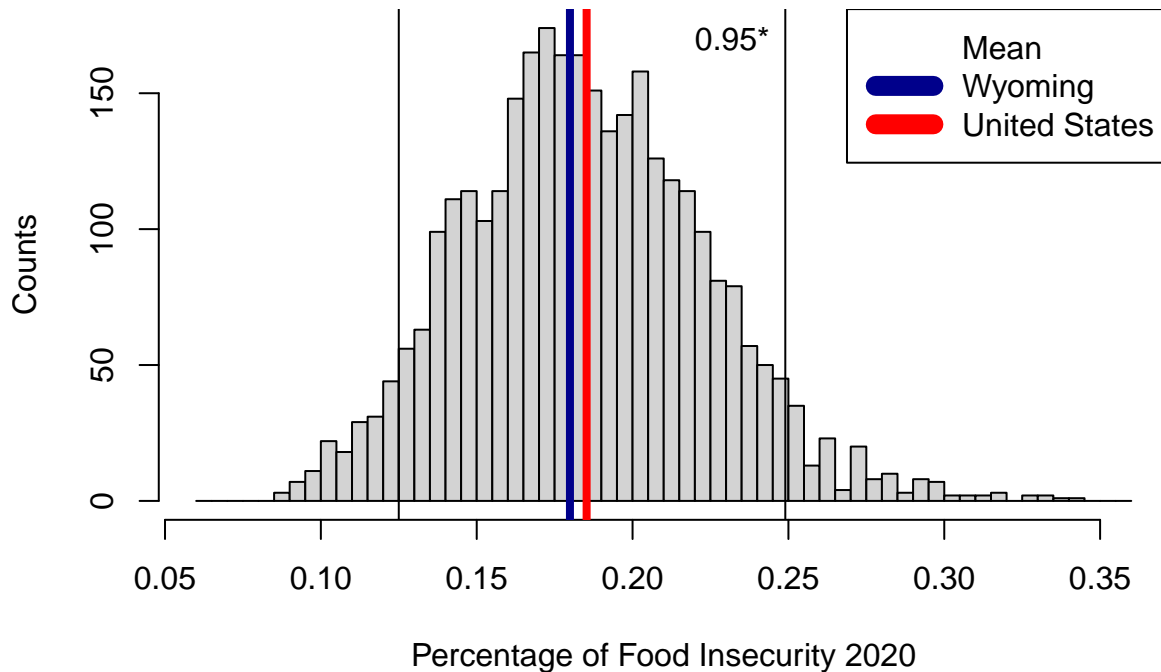


Fig 1 Histogram of 2020 National food insecurity percentages with Wyoming and National means in blue and red respectfully. Figure made by Ben Romanjenko using USDA Food Security Data.

Despite the common belief that rurality represents agricultural abundance thus increased access to food, we can see that Wyoming being quite rural, is not significantly different compared to the national average for percent food insecurity.

Introduction

Project Summary

In fall 2020 I was enrolled in the Microbiology Capstone class. With a shift to an online class structure, due to the COVID-19 pandemic, we, the students, were challenged by our professor, Rachel Watson, to develop a research proposal centered around publicly available COVID-19 data, and our own personal interests. Interested in diet, and food systems the Team Nutritional Advantage comprised of Mercedes Fermelia, Tom Hafner, Gabby Flores, and I, was formed. In early semester we wrote personal proposals and team proposals, to better refine our personal and team interests to determine a header. After much time was spent looking for publicly available data, we had been able to identify the John Hopkins Covid database with data to the county resolution. Additionally, many food security, food environment variables, and metabolic disorder data were publicly available also at the county level. Diet was one data set that we could not find at the county resolution from a reputable source, such that it was left out of our analysis.

With our data identified we were set to begin our statistical comparisons, looking across Wyoming counties, identifying how limited access to nutritional foods, chronic conditions like obesity, correlate with the outcomes of COVID-19 infections. This work will help support future inquiry about the interfaces between viral infection, other health conditions and nutritional environments. Additionally, it may identify correlations to counties with at risk and rural populations such as Fremont County, containing the Wind River Indian Reservation, having a greater COVID-19 infection rate to surrounding counties at the start of the pandemic. With little research, pertaining to food systems of rural counties and their relation to COVID-19 outcomes, Wyoming data is a good place to start the analysis. Our findings may be used to inform and incentivize public health through diet.

This collective work was done by an undergraduate cohort, building community and academic relations in the process and inquiries of our research. Learning the fundamental and ethical responsibilities a researcher

has ensures our work is is readily inclusive and accessible to the people in which our work may educate and impact.

With a desire to gain deeper understanding of the relations between food systems and COVID-19 outcomes, I continued to look for further explanations of Wyoming county findings. With the additional United States COVID-19 and food environment data at the county level, the previous unseen relations became clearer with further analysis, in which the small Wyoming data set could only minimally reveal.

Project Objectives

Methods

Data Collection

Food security and food environment data was gathered from reputable open sources including Census Household Pulse Survey, Feeding America, and USDA Household Food Security data. COVID-19 data was initially gathered from the Wyoming Public Health Lab (WPHL) for Wyoming counties and county data for the United States was gathered from the John Hopkins Covid-19 Git Hub repository.

Data Wrangling

Loading the required packages for data wrangling and further analysis.

```
library(readr)
library (shiny)
library (ggplot2)
library (maps)
library (tidyr)
library (shinythemes)
library (dplyr)
library (plotly)
library (lubridate)
library(piecewiseSEM)
library(lavaan)
library(semPlot)
```

Data wrangling code

```
#Reading in the Food_data set
food <- read.csv ("Food_data.csv")

#Reading in food insecurity data
insecurity <- read.csv("The Impact of the Coronavirus on Food Insecurity v1.csv")

#####

#Reading in United States COVID-10 data from John Hopkins Git Hub repository.
#Confirmed Cases:
covid.us1 <- read.csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/confirmed_cases.csv")
#need to convert to long format
covid.us1 <- gather (covid.us1, Date, Confirmed, 12:ncol(covid.us1))

#Deaths
covid.us2 <- read.csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/deaths.csv")
#and convert to long form
covid.us2 <- gather (covid.us2, Date, Deaths, 13:ncol(covid.us2))
```

```

#Join Data Sets Together:

if (sum (covid.us2$Date!=covid.us1$Date | covid.us2$Admin2!=covid.us1$Admin2)==0){
  covid.us1$Population <- covid.us2$Population
  covid.us1$Deaths <- covid.us2$Deaths
}

covid.us3 <- covid.us1

#Fix the date column by removing the X
#get rid of Xs
covid.us3$Date <- gsub ("X", "", covid.us3$Date)

#Convert date format to month,day,year
covid.us3$Date <- mdy (covid.us3$Date)

#Adjust column names to allow for better addition of other data sets
#because github used _ instead of .
#this way you don't have to change things later on
colnames (covid.us3) <- gsub ("_", ".", colnames (covid.us3))

#the github data has additional columns not present in it
#you may not need to delete columns (who knows, that info might be useful)
#but if you refer to columns by number in you app,
#the data frames will need to match exactly
covid.us <- covid.us3[,c(12, 6, 7, 13, 15, 8, 14)]
#this should be the exact same as the "covid.us" we used before

#####

#Use same code to read in the whole food data and US Covid data
#Need to combine the whole food data and the us covid data

combined.us <- covid.us
combined.us <- combined.us [combined.us$Date==max(combined.us$Date),]
#this is where we need to combine the data so that
#population will work for calculating rates
#we want to combine the covid data with the food data but their
#lengths are not the same such they will not merge properly.
#Food data needs a column that as actually state names
food$State<-state.name[match(food$State,state.abb)]

#renaming combined.us columns to match food ie admin2
#will become County and province.state will be State
colnames(combined.us) [2] <- "County"
colnames(combined.us) [3] <- "State"

##left merging the smaller data frame into the larger
data.us <- left_join(combined.us, food)

```

```

#Using the Cases and death data to calculate per 100,000, and fatality rate
data.us$Cases_per_100000 <- (data.us$Confirmed/data.us$Population*100000)
data.us$Deaths_per_100000 <- data.us$Deaths/data.us$Population*100000
data.us$Fatality_Rate <- data.us$Deaths/data.us$Confirmed

#Subsetting the data to just Wyoming such that code wont have to be rewritten.
data.wy <- data.us[data.us$State == "Wyoming",]

point_data <- read.csv ("County_labels.csv")
#making a subregion column to match the counties.wy data set such that
#they can be joined together.
data.wy$subregion <- tolower(data.wy$County)

counties.wy <- map_data("county", region = "wyoming")

counties.covid.wy <- left_join(counties.wy, data.wy, by= "subregion")
#counties.covid.wy$Cases_per_100000

```

Analysis

Data was analyzed via the program R: a powerful programming language developed for statistical analysis. The initial analysis for Wyoming data was completed with linear models. The larger national data was initially analyzed via linear modeling, but with the larger sample of data, Statistical Equation Modeling (SEM) was a beneficial analysis, able to compare the effect of multiple variables on each other.

Nutritional Advantage: Wyoming

Summary

Through combined work of Nutritional Advantage and Ella DeWolf we were successful in our goals of creating a public website and R shiny app to convey our findings regarding food environments and COVID-19 outcomes. Here is The Nutritional Advantage website that our team made, in addition to the developed R Shiny application .

Preliminary Analysis

County Food insecurity In regards to where Wyoming sat compared to national food insecurity, I wanted to have a deeper understanding of the food insecurity experienced in each individual county. With this I graphed the histogram marking the 0.95 percentile and determined two counties that were significantly different, being Fremont and Niobrara counties.

```

hist(insecurity$X.C._2020_Food_Insecurity_. [insecurity$State == "WY"],
     xlab = "Percentage of Food Insecurity 2020", main = "",
     ylab="Counts", breaks = 9)
abline (v = mean(insecurity$X.C._2020_Food_Insecurity_.), col = "red", lwd=3)
abline (v = insecurity$X.C._2020_Food_Insecurity_.
       [insecurity$County_State == "Fremont_County,_Wyoming" ],
       col = "blue", lwd= 3)
abline (v = insecurity$X.C._2020_Food_Insecurity_.
       [insecurity$County_State == "Niobrara_County,_Wyoming" ],
       col = "hotpink", lwd= 3)
abline (v = mean(insecurity$X.C._2020_Food_Insecurity_.[insecurity$State == "WY"]),

```

```

    lwd = 3, col = "yellow")
abline (v = quantile(insecurity$X.C._2020_Food_Insecurity_[insecurity$State == "WY"],
                    probs = c(NA, 0.95)))
text(0.199, 3.5, "0.95*")
legend("topleft", c("Mean", "Wyoming", "United States", "Fremont", "Niobrara"),
      col=c(NA, "yellow", "red", "blue", "hotpink"), lwd=10)

```

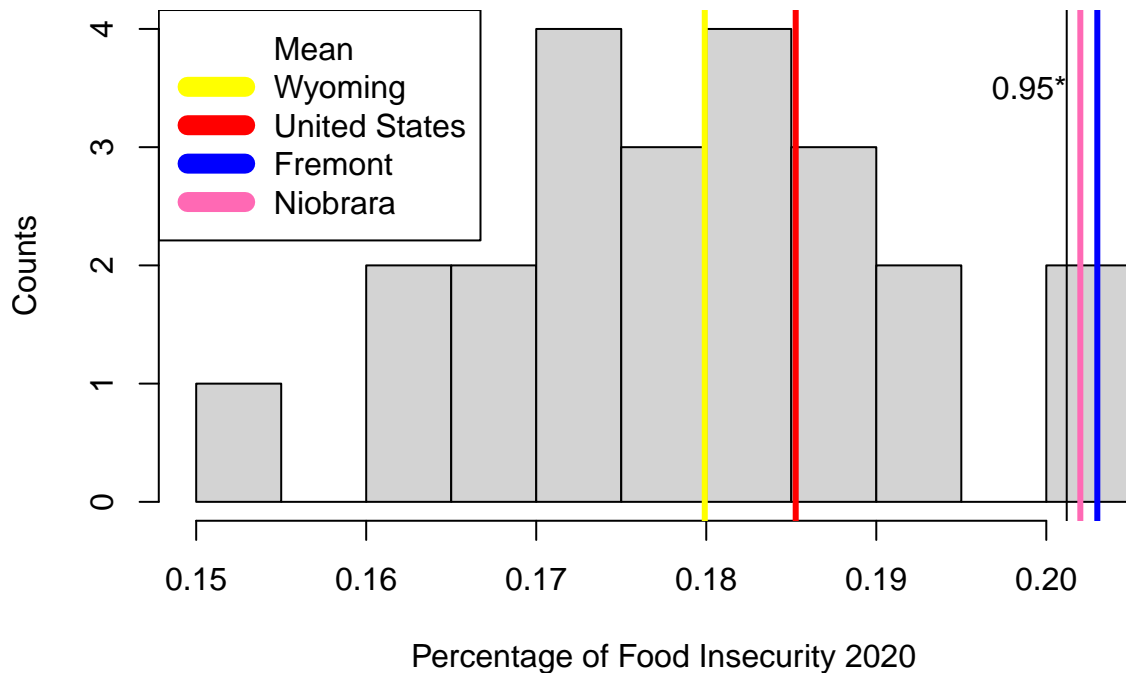


Fig 2 Histogram of 2020 Wyoming food insecurity. Identifying Niobrara and Fremont Counties as outside the norm for Wyoming food insecurity.

Upper Quantile for Wyoming Food Insecurity: 0.2012

Niobrara County Percentage Food Insecure: 0.202

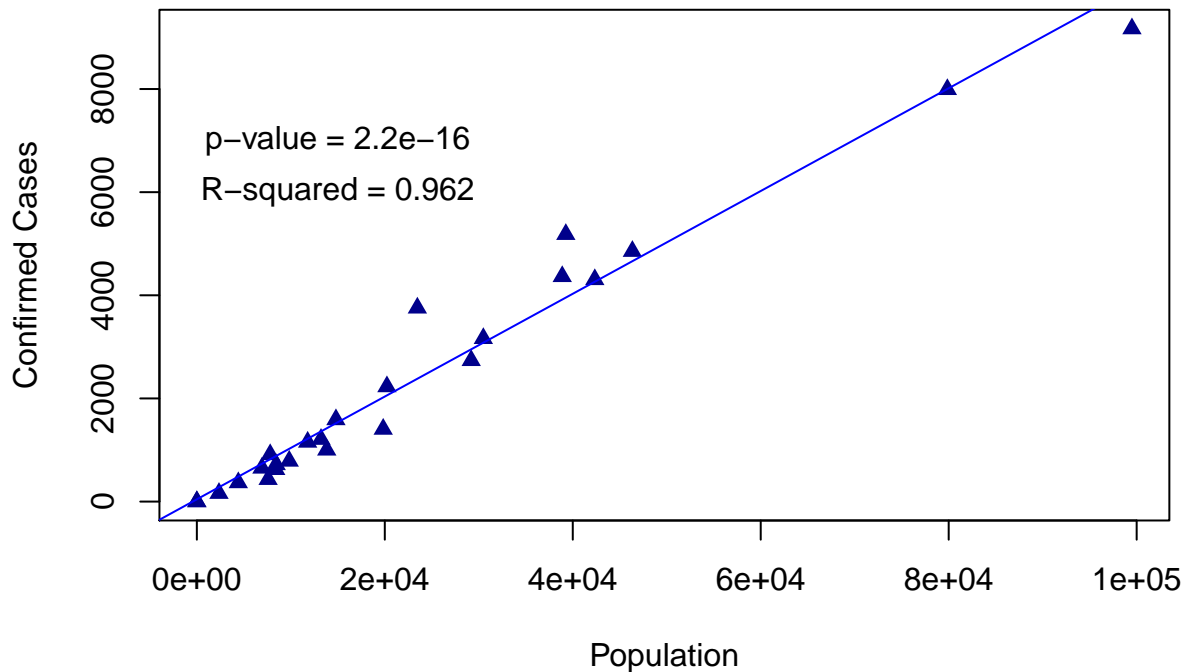
Femont County Percentage Food Insecure: 0.203

Cases & Deaths by Population Further analyses were based on simple linear regression. By plotting two data against each other a correlation can be observed if the points form a line following standard $y=mx+b$ equation. In the two graphics, Cases and Deaths can be plotted against population to indicate a linear trend. Much work cond in the R Shiny is based on this code and concept.

```

par(mfrow = c(1,1))
plot(data.wy$Population, data.wy$Confirmed, xlab = "Population",
     ylab= "Confirmed Cases", pch= 17, col = "darkblue")
abline(lm(Confirmed ~ Population, data.wy), col = "blue")
text(15000, 7000, "p-value = 2.2e-16")
text(15000, 6000, "R-squared = 0.962")

```



```
plot(data.wy$Population, data.wy$Deaths, xlab = "Population",
      ylab = "Deaths", pch=17, col= "darkblue")
abline(lm(Deaths ~ Population, data.wy), col = "blue")
text(15000, 120, "p-value = 1.32e-9")
text(15000, 110, "R-squared = 0.795")
```

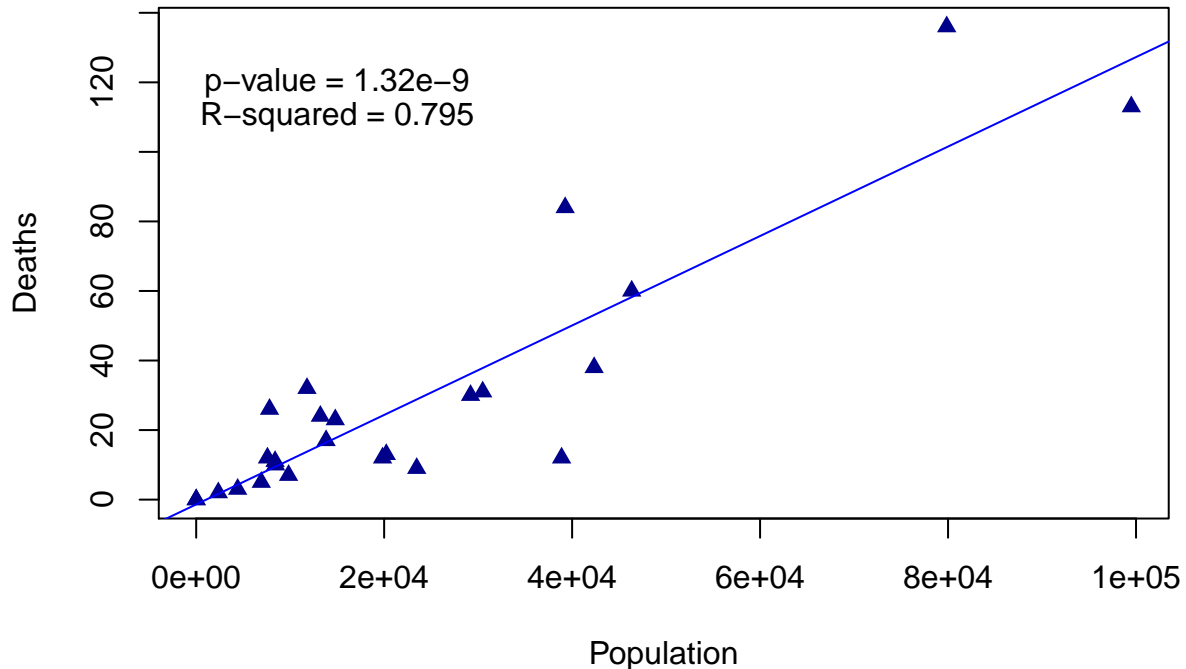


Fig 3 Plotting Wyoming county population against Confirmed Cases and Deaths with trend line.

Summary of simple linear regressions for above plots.

```
summary(lm(Population ~ Confirmed, data.wy))
summary(lm(Population ~ Deaths, data.wy))
```


ggplot Mapping We used ggplot to build a map of Wyoming, which allows us to look at multiple variables at any one time. Because the map is visual and color based, our findings can be better communicated to a broader community than traditional graphs and statistical values.

```
#plot
map <- ggplot()+
  #add counties as polygons and color by confirmed cases
  #color= black gives black borders
  geom_polygon(data = counties.covid.wy,
              aes (x=long, y= lat,group = group, fill = Deaths), color= "black") +
  #add points in each county scaled by population
  geom_point (data=point_data, aes(x=long, y= lat + 0.1, size = Population)) +
  #adding labels
  geom_text(data=point_data, aes(x=long, y= lat- 0.05, label= County), size= 3) +
  #changing the color scheme to obnoxiously bright colors
  scale_fill_gradient (low="white", high = "steelblue") +
  #getting rid of axis/labels
  theme_void() +
  #creating a title
  ggtitle("") +
  #centering the title
  theme(plot.title = element_text(hjust = 0.5))
map ## static map
```

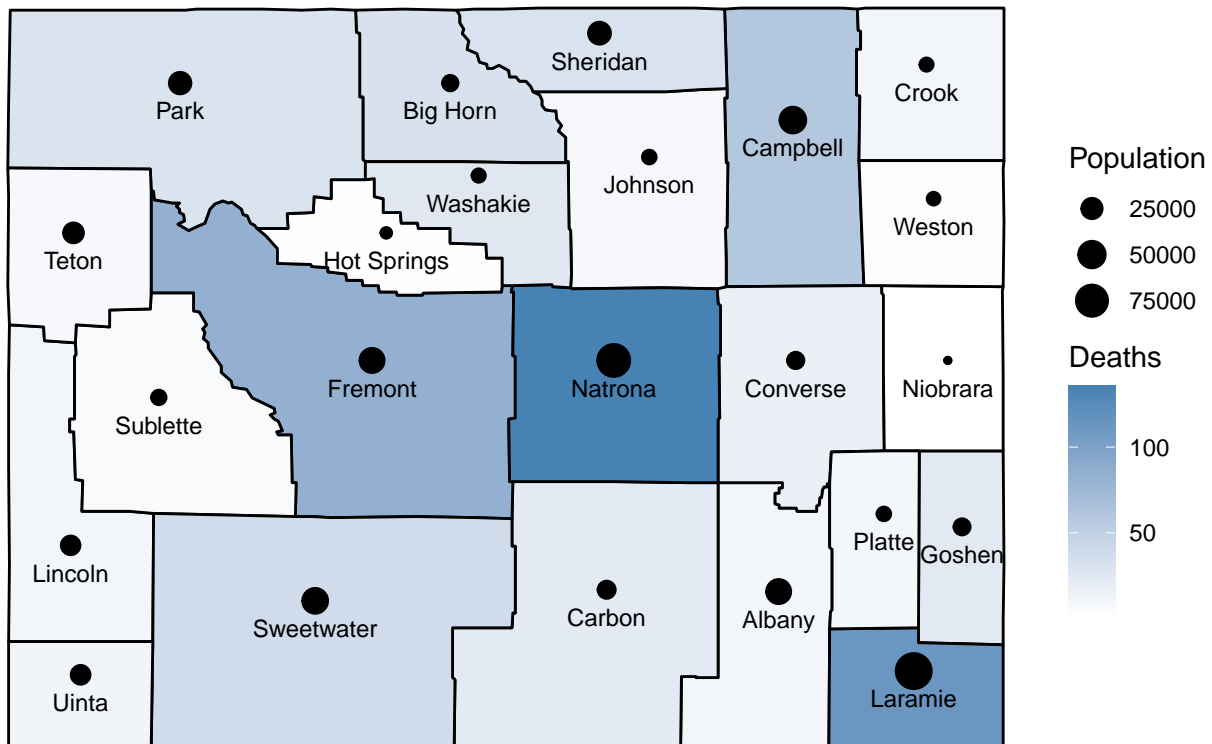


Fig 4 Map of Wyoming looking at population as the size of the point and Deaths as shade of blue.

The complete code and data for The Nutritional Advantage R Shiny can be found [Here](#)

Results

From our analysis completed above, specifically the work done in the R Shiny, our team determined a number of statistically significant findings (p-value <0.05) being:

Tab of Application	COVID Variable	Variable
Food Access & Environment	Confirmed Cases	<ul style="list-style-type: none"> • Percent Access to Stores • Access to SNAP authorized stores • Access to WIC authorized stores • Farmer's Markets
Food Access & Environment	Deaths	<ul style="list-style-type: none"> • Access to SNAP authorized stores • Access to WIC authorized stores • Students eligible for free lunch • Farmer's Markets
Food Access & Environment	Cases per 100,000	<ul style="list-style-type: none"> • Farmer's Markets
Food Access & Environment	Deaths per 100,00	<ul style="list-style-type: none"> • Students eligible for free lunch • Percent Diabetes
Food Access & Environment	Fatality Rate	<ul style="list-style-type: none"> • Students eligible for reduced priced lunch • Percent Diabetes
Obesity & Diabetes	Cases per 100,000	<ul style="list-style-type: none"> • Percent Diabetes
Obesity & Diabetes	Deaths per 100,000	<ul style="list-style-type: none"> • Percent Diabetes

Figure 1: **Table 1** A table of significant findings in the R Shiny app. All correlations are positive between Covid Variable and Food environment variable.

Discussion

Our group discussed the many possible reasons that could better explain our findings.

For decreased access to stores, we believed that individuals from rural locations would become exposed by one of two possibilities: First, the local convenience store, in which there is a potential for high transmission rate, due to small buildings, and social standards being more relaxed when it comes to mask wearing and distancing, Second, rural populations may have to shop further away from home, in location with a greater population. Due to the increased travel time, individuals would spend more time shopping in the store, thus leading to a higher chance of interacting with a positive covid-19 individual leading to transmission, infection, and dispersal of the virus back to the locality of the rural individual.

We additionally saw an increasing percentage of SNAP and WIC stores with an increase in covid cases and deaths. This observation may be related to the limited access to resources, which may indicate less privilege to socially distance, to remain at home working remotely, or not going to work.

The observation of farmers markets being significant to cases, deaths and cases per 100,000 came as a surprise. We figured that the open air markets, and healthy produce options would mitigate COVID-19. With further thought we suspected it may not be the market itself but the socializing, gathering, and additional shopping in closed environments that may have led to the increased chances of exposure and cases.

Finding significance between students eligible for free, and reduced priced lunch against Deaths per 100,000 and fatality rate. We believe this may be less associated with the students receiving the lunch programs, but the conditions found within these districts. This being low income families, living in multigenerational households, such there is greater chance of exposing an older immunocompromised family member, and not having the money or means to access health care support for the sick individual.

Our final finding indicated that diabetes and deaths per 100,000, cases per 100,00 and fatality rate were significant for two independent diabetes data sets. This understanding is grounded in the number of other immunological compromising conditions that individuals with diabetes can have making them more susceptible to COVID-19 infections.

National

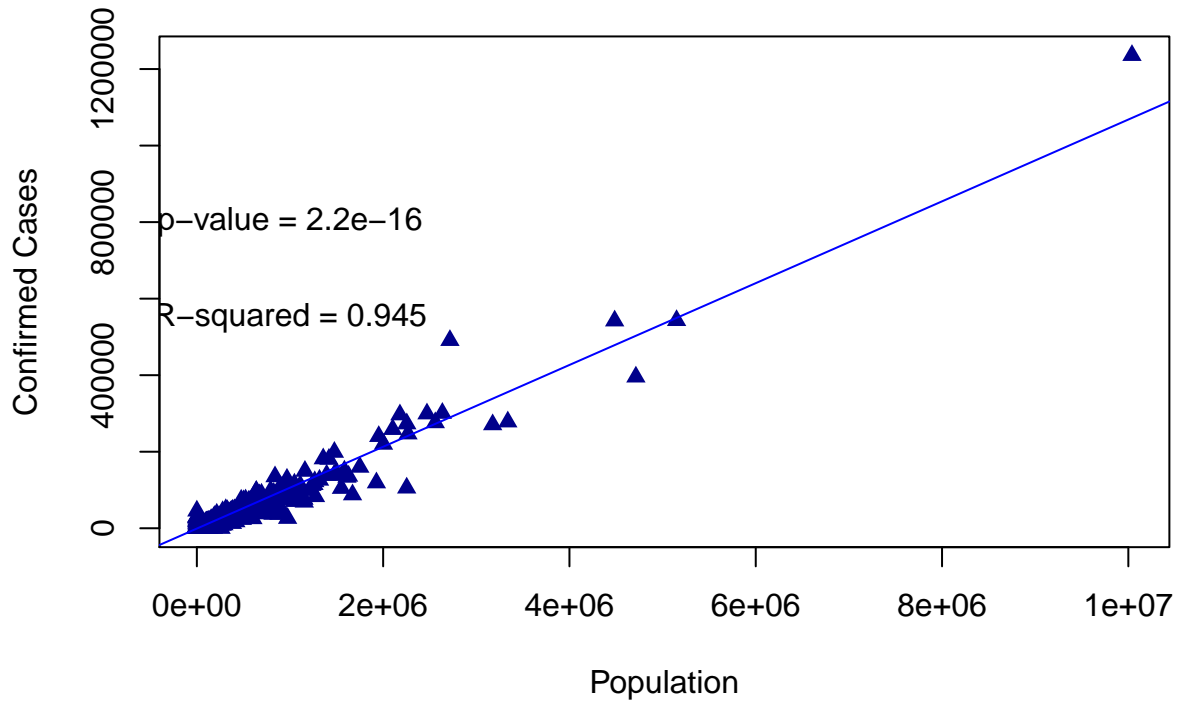
Summary

With the various spurious connections at the Wyoming county level our team had various ideas about the statistically significant findings. To ensure our findings were not just a local finding I wanted to expand our work to include all counties of the United States.

Analysis and Results

County Population vs Cases and Deaths In the following graphic I replot the population vs Cases and Deaths graphic at the national level. This allows us to identify that the trend between population and increased cases and deaths is consistent even in larger data sets.

```
par(mfrow = c(1,1))
plot(data.us$Population, data.us$Confirmed, xlab = "Population",
      ylab= "Confirmed Cases", pch= 17, col= "darkblue")
abline(lm(Confirmed ~ Population, data.us), col = "blue")
text(1000000, 800000, "p-value = 2.2e-16")
text(1000000, 550000, "R-squared = 0.945")
```



```

plot(data.us$Population, data.us$Deaths, xlab = "Population",
      ylab= "Deaths" , pch= 17, col= "darkblue")
abline(lm(Deaths ~ Population, data.us), col = "blue")
text(1000000, 15000, "p-value = 2.2e-16")
text(1000000, 10000, "R-squared = 0.861")

```

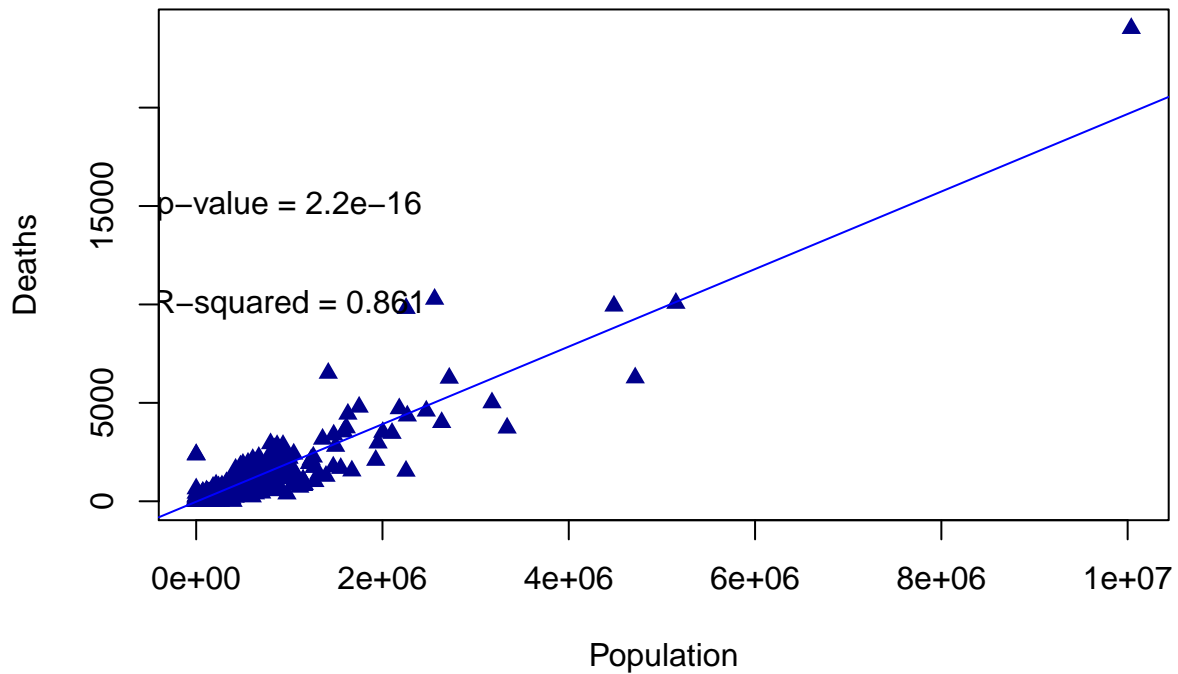


Fig 5 *Plotting National county populations against Confirmed Cases and Deaths with trend line.*
 Summary of simple linear regression lines for the national county data.

```
summary (lm(Confirmed ~ Population, data.us))
summary (lm(Deaths ~ Population, data.us))
```

Multiple Linear Models In the following code I analyze how the the food environment, demographic data and population affected the Covid-19 related deaths. This was replicated for the other COVID variables. The output has common statistic parameters like p-value and R-squared. The better way to make these models and connections would be through the structural equation modeling.

```
summary (lm(data.us$Deaths ~ data.us$PCT_NHNA10 + data.us$SNAPS17 + data.us$Population))
summary (lm(data.us$Deaths ~ data.us$PCT_NHNA10 + data.us$WICS16 + data.us$Population))
summary (lm(data.us$Deaths ~ data.us$PCT_NHNA10 + data.us$FMRKT18 + data.us$Population))
summary (lm(data.us$Deaths ~ data.us$PCT_NHNA10 + data.us$PCT_FREE_LUNCH15 + data.us$Population))
```

Structural Equation Modeling Structural equation modeling (SEM) is similar to linear model except multiple variables can be cross analyzed, while giving hierarchical affect to each variable, such population and median household income have the greatest effect on COVID-19 outcomes and food environment variable compared to the food environment variable alone. This analysis is only done at the national level due to the significantly bigger sample size compared to that of Wyoming's 23 counties.

Using sem model, outputs are similar to that of the linear models. Looking for low p-values, and estimate which identifies the rate of the variables to each other.

```
semodel <- psem(
  lm(Deaths ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population, data=data.us),
  lm(Confirmed ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population, data=data.us),
  lm(Deaths_per_100000 ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population, data=data.us),
  lm(Cases_per_100000 ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population, data=data.us),
  lm(Fatality_Rate ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population, data=data.us),
  lm(FMRKT18 ~ PCT_NHWHITE10 + MEDHHINC15 + Population,
    data=data.us),
  lm(PCT_NHWHITE10 ~ Population, data=data.us)
)
```

Levant An alternative method for SEM modeling is that of Levant, which permits for simpler integration of statistical outputs into visuals.

To make the results more clear I scaled the pertinent data.

This analysis is similar to the initial SEM, but with a little bit different formatting. The outputs for this are also slightly different, having all of the standard p-values and R Square values in addition to a relation estimate or rate between driver and variable. Note that the first model uses the percent white population as it is the inverse of the percent non-white populations that tend to be more marginalized.

```
m1 <- 'Deaths ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population
Confirmed ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population
Deaths_per_100000 ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population
Cases_per_100000 ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population
Fatality_Rate ~ PCT_NHWHITE10 + FMRKT18 + MEDHHINC15 + Population
FMRKT18 ~ PCT_NHWHITE10 + MEDHHINC15 + Population
PCT_NHWHITE10 ~ Population + MEDHHINC15'

m1.fit <- sem(m1, data=data.us.scaled)
```

Here is an example of the output for the first model.

```
summary(m1.fit, stand=TRUE, rsq=TRUE)
```

```
## lavaan 0.6-8 ended normally after 388 iterations
```

```
##
```

```
## Estimator ML
```

```
## Optimization method NLMINB
```

```
## Number of model parameters 42
```

```
##
```

```
## Used Total
```

```
## Number of observations 3108 3234
```

```
##
```

```
## Model Test User Model:
```

```
##
```

```
## Test statistic 1392.186
```

```
## Degrees of freedom 0
```

```
##
```

```
## Parameter Estimates:
```

```
##
```

```
## Standard errors Standard
```

```
## Information Expected
```

```
## Information saturated (h1) model Structured
```

```
##
```

```
## Regressions:
```

```
## Estimate Std.Err z-value P(>|z|) Std.lv
```

```
## Deaths ~
```

```
## PCT_NHWHITE10 5.766 NA 5.766
```

```
## FMRKT18 20.067 NA 20.067
```

```
## MEDHHINC15 -59.261 NA -59.261
```

```
## Population 651.854 NA 651.854
```

```
## Confirmed ~
```

```
## PCT_NHWHITE10 565.976 NA 565.976
```

```
## FMRKT18 -2458.118 NA -2458.118
```

```
## MEDHHINC15 -1585.790 NA -1585.790
```

```
## Population 37832.372 NA 37832.372
```

```
## Deaths_per_100000 ~
```

```
## PCT_NHWHITE10 -0.184 NA -0.184
```

```
## FMRKT18 -0.109 NA -0.109
```

```
## MEDHHINC15 -0.272 NA -0.272
```

```
## Population 0.051 NA 0.051
```

```
## Cases_per_100000 ~
```

```
## PCT_NHWHITE10 -0.125 NA -0.125
```

```
## FMRKT18 -0.209 NA -0.209
```

```
## MEDHHINC15 -0.081 NA -0.081
```

```
## Population 0.144 NA 0.144
```

```
## Fatality_Rate ~
```

```
## PCT_NHWHITE10 -0.120 NA -0.120
```

```
## FMRKT18 -0.016 NA -0.016
```

```
## MEDHHINC15 -0.273 NA -0.273
```

```
## Population -0.008 NA -0.008
```

```
## FMRKT18 ~
```

```
## PCT_NHWHITE10 0.025 NA 0.025
```

```
## MEDHHINC15 0.111 NA 0.111
```

```
## Population 0.773 NA 0.773
```

```
## PCT_NHWHITE10 ~
```

```

##      Population          -0.287      NA          -0.287
##      MEDHHINC15          0.219      NA          0.219
## Std.all
##
##      0.008
##      0.028
##     -0.083
##      0.932
##
##      0.015
##     -0.067
##     -0.043
##      1.042
##
##     -0.187
##     -0.111
##     -0.276
##      0.053
##
##     -0.137
##     -0.230
##     -0.089
##      0.162
##
##     -0.120
##     -0.016
##     -0.273
##     -0.008
##
##      0.025
##      0.111
##      0.784
##
##     -0.291
##      0.219
##
## Covariances:
##      Estimate      Std.Err  z-value  P(>|z|)  Std.lv
## .Deaths ~~
##   .Confirmed          188.032      NA          188.032
##   .Dths_pr_100000      64.662      NA          64.662
##   .Cass_pr_100000       1.714      NA          1.714
##   .Fatality_Rate       67.915      NA          67.915
## .Confirmed ~~
##   .Dths_pr_100000     -671.851      NA         -671.851
##   .Cass_pr_100000     1730.778      NA         1730.778
##   .Fatality_Rate    -1569.062      NA        -1569.062
## .Deaths_per_100000 ~~
##   .Cass_pr_100000       0.328      NA          0.328
##   .Fatality_Rate       0.720      NA          0.720
## .Cases_per_100000 ~~
##   .Fatality_Rate      -0.051      NA         -0.051
## Std.all
##

```

```

##      0.000
##      0.282
##      0.008
##      0.284
##
##     -0.091
##      0.242
##     -0.204
##
##      0.408
##      0.833
##
##     -0.060
##
## Variances:
##           Estimate      Std.Err  z-value  P(>|z|)  Std.lv
## .Deaths          63294.385         NA         NA      63294.385
## .Confirmed       65455107.360         NA         NA     65455107.360
## .Dths_pr_100000      0.828         NA         NA         0.828
## .Cass_pr_100000     0.782         NA         NA         0.782
## .Fatality_Rate      0.903         NA         NA         0.903
## .FMRKT18           0.340         NA         NA         0.340
## .PCT_NHWHITE10     0.902         NA         NA         0.902
## Std.all
##      0.125
##      0.048
##      0.857
##      0.944
##      0.899
##      0.337
##      0.899
##
## R-Square:
##           Estimate
## Deaths          0.875
## Confirmed        0.952
## Dths_pr_100000   0.143
## Cass_pr_100000  0.056
## Fatality_Rate    0.101
## FMRKT18          0.663
## PCT_NHWHITE10    0.101

```

```

m2 <- 'Deaths ~ PCT_NHWHITE10 + PCT_FREE_LUNCH15 + MEDHHINC15 + Population
Confirmed ~ PCT_NHWHITE10 + PCT_FREE_LUNCH15 + MEDHHINC15 + Population
Deaths_per_100000 ~ PCT_NHWHITE10 + PCT_FREE_LUNCH15 + MEDHHINC15 + Population
Cases_per_100000 ~ PCT_NHWHITE10 + PCT_FREE_LUNCH15 + MEDHHINC15 + Population
Fatality_Rate ~ PCT_NHWHITE10 + PCT_FREE_LUNCH15 + MEDHHINC15 + Population
PCT_FREE_LUNCH15 ~ PCT_NHWHITE10 + MEDHHINC15 + Population
PCT_NHWHITE10 ~ Population'

```

```

m2.fit <- sem(m2, data=data.us.scaled)
#summary(m2.fit, stand=TRUE, rsq=TRUE)

```

```

m3 <- 'Deaths ~ PCT_NHWHITE10 + WICS16 + MEDHHINC15 + Population
Confirmed ~ PCT_NHWHITE10 + WICS16 + MEDHHINC15 + Population

```



```
Deaths_per_100000 ~ PCT_NHWHITE10 + WICS16 + MEDHHINC15 + Population
Cases_per_100000 ~ PCT_NHWHITE10 + WICS16 + MEDHHINC15 + Population
Fatality_Rate ~ PCT_NHWHITE10 + WICS16 + MEDHHINC15 + Population
WICS16 ~ PCT_NHWHITE10 + MEDHHINC15 + Population
PCT_NHWHITE10 ~ Population'
```

```
m3.fit <- sem(m3, data=data.us.scaled)
#summary(m3.fit, stand=TRUE, rsq=TRUE)
```

```
m4 <- 'Deaths ~ PCT_NHWHITE10 + SNAPS17 + MEDHHINC15 + Population
Confirmed ~ PCT_NHWHITE10 + SNAPS17 + MEDHHINC15 + Population
Deaths_per_100000 ~ PCT_NHWHITE10 + SNAPS17 + MEDHHINC15 + Population
Cases_per_100000 ~ PCT_NHWHITE10 + SNAPS17 + MEDHHINC15 + Population
Fatality_Rate ~ PCT_NHWHITE10 + SNAPS17 + MEDHHINC15 + Population
SNAPS17 ~ PCT_NHWHITE10 + MEDHHINC15 + Population
PCT_NHWHITE10 ~ Population'
```

```
m4.fit <- sem(m4, data=data.us.scaled)
#summary(m4.fit, stand=TRUE, rsq=TRUE)
```

```
par (mfrow = c(2,2))
semPaths(m1.fit, what="std", whatLabels="std", residuals=FALSE)
semPaths(m2.fit, what="std", whatLabels="std", residuals=FALSE)
semPaths(m3.fit, what="std", whatLabels="std", residuals=FALSE)
semPaths(m4.fit, what="std", whatLabels="std", residuals=FALSE)
```

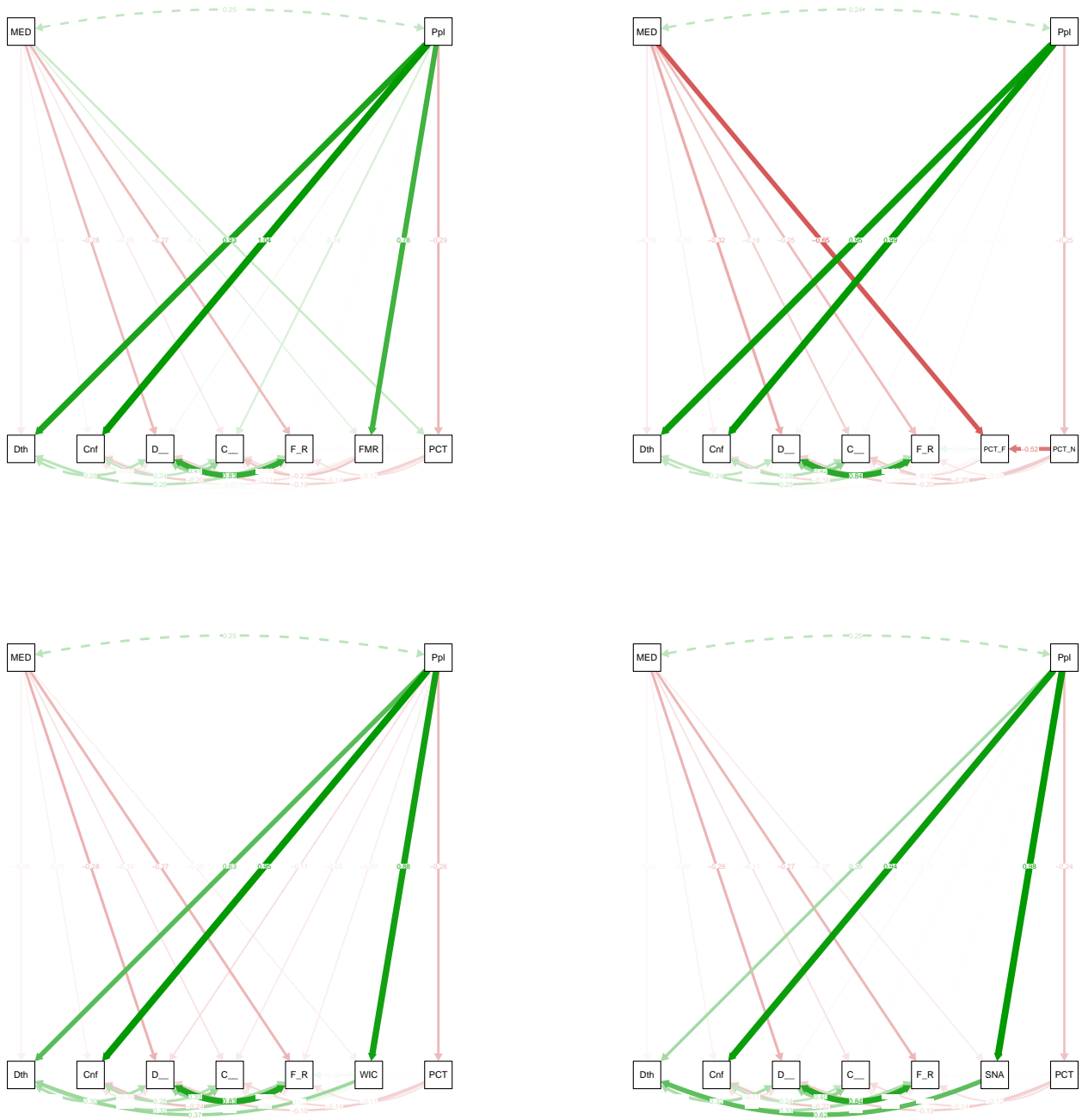


Fig 6 *SEM graphic outputs indicating positive correlation in green, negative correlation in red and rate by shade of the line. The coordinating variables to shortend code are,

- MED** : Median Household Income
- Ppl** : Population
- Dth** : Number of Deaths
- Cnf** : Number of Confirmed Cases
- D_** : Deaths per 100,000
- C_** : Cases per 100,000
- F_R** : Fatality Rate
- FMR** : Access to Farmers Markets
- PCT_F** : Percent Free School Lunches
- WIC** : Access to WIC Stores

SNA : Access to SNAP Stores
PCT : Population Percentage White *

Discussion

The biggest findings in these SEM are that population has a strong positive correlation with deaths, cases, access to SNAP stores, access to WIC stores, and farmers markets. These findings for the covid-19 outcomes is understandable as COVID-19 is a readily communicable virus easily transmitted in areas of high population density. The same areas that need greater food access to both marginalized and non-marginalized such greater number of stores, and SNAP and WIC stores, in addition to farmers markets. Opposed to the Wyoming specific analysis we can better understand that the access to farmers markets and increased covid outcomes is driven more by the population.

A second finding is the relation of median household income to deaths per 100,000, cases per 100,000, fatality rate, percent free lunch. For all four it was a negative associate between median household income and the variable. To better understand this relation as median household income increased the deaths per 100,000 decreased, thus the opposite is true when median household income decreases the COVID-19 outcomes increase. This indicates counties with lower median household incomes are more likely to experience disproportionate impacts of pandemics like COVID-19. The one food variable that was positively corelated with median household income was access to farmers markets. This relation can be explained by the higher cost of goods found at farmers markets that can only be afforded by the middle and upper economic classes.

The third findings is associated with demographic. Consistent between all models was with increasing population a decrease in percent white can be observed. With that, median household income increased the percentage of white also increased, a possible indication of systemic racism. Further more, as white population increased there was consistent decrease of all COVID-19 outcomes. Reversed, this means populations of color experience greater number of COVID-19 outcomes.

In all findings we can broadly identify marginalized groups that have been disproportionately impacted due to COVID-19. With areas of high population, being epicenters for increased diversity and marginalized groups it would be of the utmost importance to have directed resources in their direction in addition to prioritizing these areas for vaccine distribution to help mitigate the current and potential future effects of COVID-19 and other highly transmissible pathogens.

National vs Wyoming Trends

Analysis

Population & COVID-19 Outcomes

```
par(mfrow = c(1,2))
par(oma = c(3, 3, 0, 0))
par(mar = c(4, 4, 1, 1)) # make the plots be closer together

plot(data.wy$Population, data.wy$Confirmed, xlab = "", ylab= "Confirmed Cases",
      col=ifelse( data.wy$Confirmed> quantile(data.wy$Confirmed, na.rm= TRUE,
      probs = c(0.95)) & data.wy$Population> quantile(data.wy$Population,
      na.rm= TRUE, probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(Confirmed ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$Confirmed, labels=data.wy$County, cex= 0.7, pos=1)
text(28000, 8000, "p-value = 2.2e-16")
```

```

text(28000, 7500, "R-squared = 0.96")

plot(data.wy$Population, data.wy$Deaths, xlab = "", ylab= "Deaths",
      col=ifelse( data.wy$Deaths> quantile(data.wy$Deaths, na.rm= TRUE,
      probs = c(0.95)) & data.wy$Population> quantile(data.wy$Population,
      na.rm= TRUE, probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(Deaths ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$Deaths, labels=data.wy$County, cex= 0.7, pos=1)
text(28000, 120, "p-value = 1.68e-9")
text(28000, 110, "R-squared = 0.79")

mtext('Population', side = 1, outer = TRUE, line = 1)

```

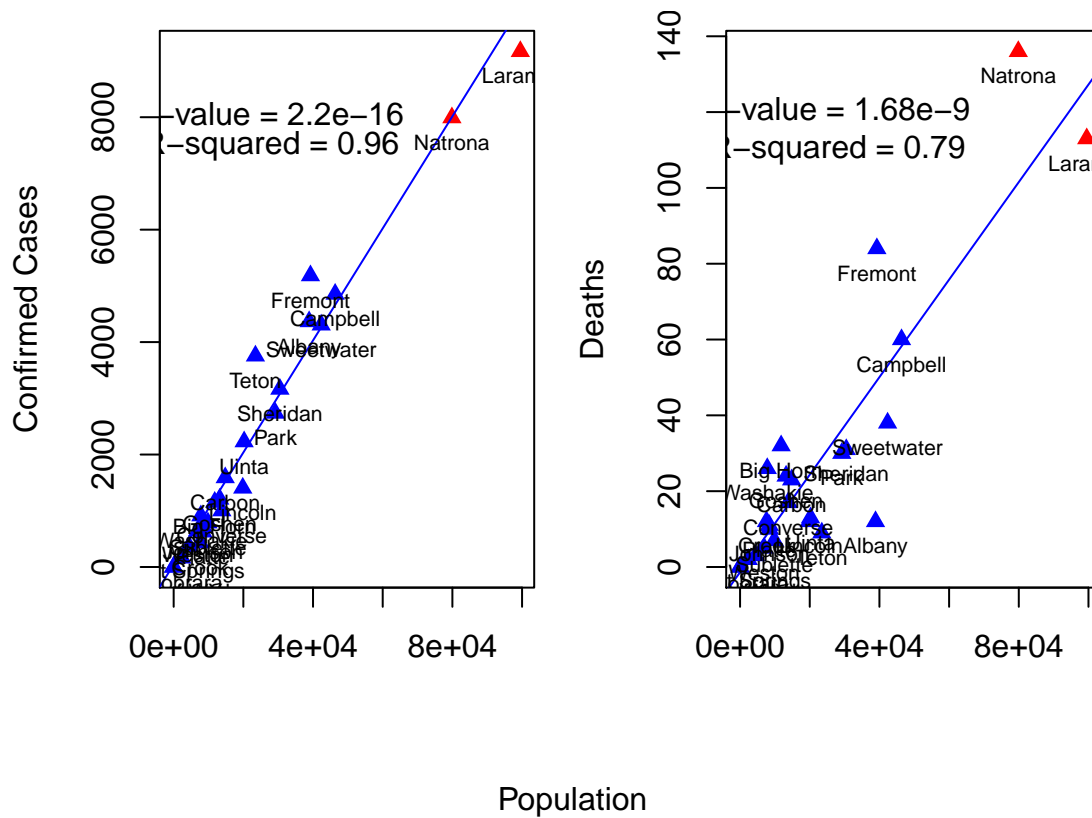


Fig 8 Comparing Population to COVID-19 outcomes. Red points are outliers for both population and COVID-19 outcomes.

Population & Food Environment

```

par(mfrow = c(1,4))
par(oma = c(3, 3, 0, 0))
par(mar = c(4, 4, 1, 1)) # make the plots be closer together

```

```

plot(data.wy$Population, data.wy$FMRKT18, xlab = "", ylab= "Farmers Market",
     col=ifelse( data.wy$FMRKT18> quantile(data.wy$FMRKT18,
     na.rm= TRUE, probs = c(0.95)) & data.wy$Population> quantile(data.wy$Population,
     na.rm= TRUE, probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(FMRKT18 ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$FMRKT18, labels=data.wy$County, cex= 0.7, pos=1)
text(39000, 5, "p-value = 0.075")
text(40000, 4.75, "R-squared = 0.10")

plot(data.wy$Population, data.wy$WICS16, xlab = "", ylab= "WICS16",
     col=ifelse( data.wy$WICS16> quantile(data.wy$WICS16, na.rm= TRUE,
     probs = c(0.95))& data.wy$Population> quantile(data.wy$Population, na.rm= TRUE,
     probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(WICS16 ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$WICS16, labels=data.wy$County, cex= 0.7, pos=1)
text(39000, 7.8, "p-value = 3.8e-6")
text(40000, 7.5, "R-squared = 0.62")

plot(data.wy$Population, data.wy$SNAPS17, xlab = "", ylab= "SNAPS17",
     col=ifelse( data.wy$SNAPS17> quantile(data.wy$SNAPS17,
     na.rm= TRUE, probs = c(0.95))& data.wy$Population> quantile(data.wy$Population,
     na.rm= TRUE, probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(SNAPS17 ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$SNAPS17, labels=data.wy$County, cex= 0.7, pos=1)
text(42000, 43, "p-value = 5.14e-14")
text(40000, 41, "R-squared = 0.93")

plot(data.wy$Population, data.wy$PCT_FREE_LUNCH15, xlab = "", ylab= "PCT_FREE_LUNCH15",
     col=ifelse( data.wy$PCT_FREE_LUNCH15> quantile(data.wy$PCT_FREE_LUNCH15,
     na.rm= TRUE, probs = c(0.95))& data.wy$Population> quantile(data.wy$Population,
     na.rm= TRUE, probs = c(0.95)), "red", "blue" ), pch = 17)
abline(lm(PCT_FREE_LUNCH15 ~ Population, data.wy), col = "blue")
text(data.wy$Population,data.wy$PCT_FREE_LUNCH15, labels=data.wy$County, cex= 0.7, pos=1)
text(39000, 42, "p-value = 0.146")
text(42000, 41, "R-squared = 0.057")

mtext('Population', side = 1, outer = TRUE, line = 1)

```

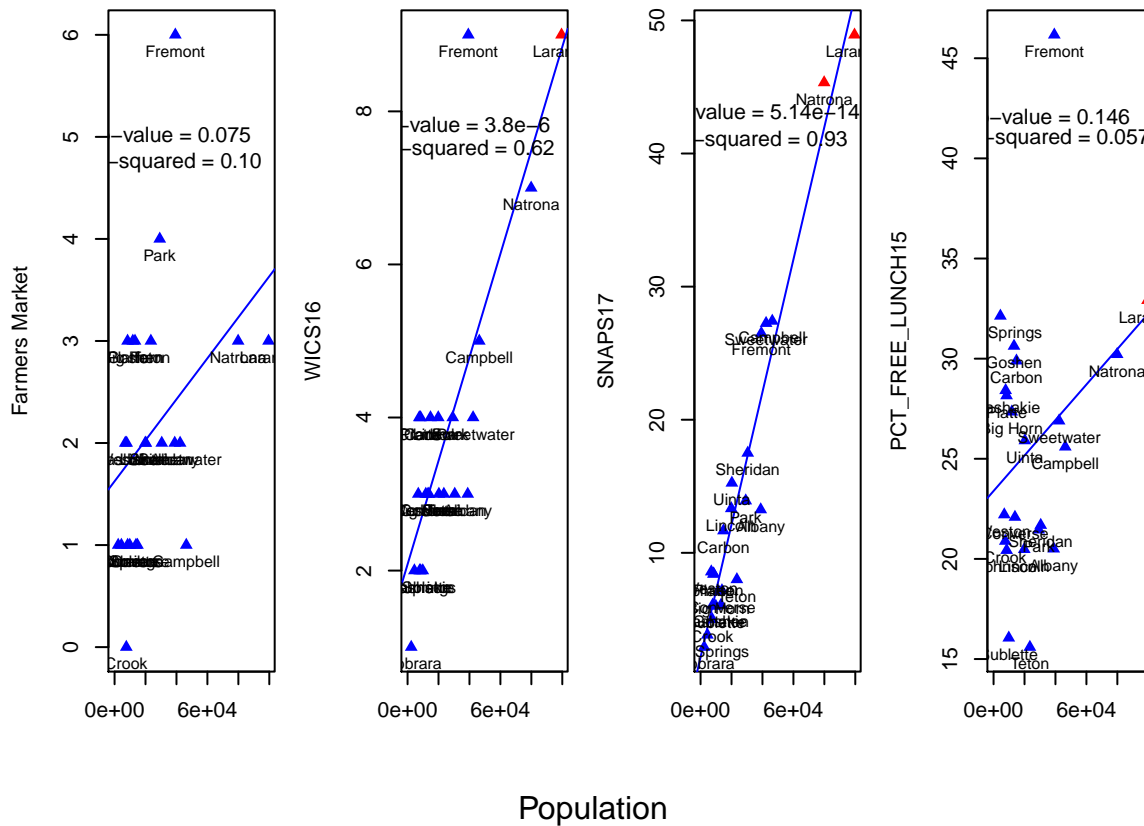


Fig 9 Comparing Population to Food Environment Variables. Red points are outliers for both population and Food Environment Variables.

Median Household Income & COVID-19 Outcomes

```

par(mfrow = c(1,3))
par(oma = c(3, 3, 0, 0))
par(mar = c(4, 4, 1, 1)) # make the plots be closer together

plot(data.wy$MEDHHINC15, data.wy$Cases_per_100000, xlab = "", ylab= "Cases per 100,000",
      col=ifelse( data.wy$Cases_per_100000 > quantile(data.wy$Cases_per_100000,
      na.rm= TRUE, probs = c(0.95)) , "red", "darkblue" ), pch = 17)
abline(lm(Cases_per_100000 ~ MEDHHINC15, data.wy), col = "blue")
#abline (v = quantile(data.wy$MEDHHINC15, na.rm= TRUE, probs = c(0.05, 0.95)))
text(data.wy$MEDHHINC15,data.wy$Cases_per_100000, labels=data.wy$County, cex= 0.7, pos=1)
text(60000, 14900, "p-value = 0.3104")
text(60000, 14500, "R-squared = 0.003")

plot(data.wy$MEDHHINC15, data.wy$Deaths_per_100000, xlab = "", ylab= "Deaths per 100,000",
      col=ifelse( data.wy$Deaths_per_100000 > quantile(data.wy$Deaths_per_100000,
      na.rm= TRUE, probs = c(0.95)), "red", "darkblue" ), pch = 17)
abline(lm(Deaths_per_100000 ~ MEDHHINC15, data.wy), col = "blue")

```

```

#abline (v = quantile(data.wy$MEDHHINC15, na.rm= TRUE, probs = c(0.05, 0.95)))
text(data.wy$MEDHHINC15,data.wy$Deaths_per_100000, labels=data.wy$County, cex= 0.7, pos=1)
text(60000, 300, "p-value = 0.265")
text(60000, 288, "R-squared = 0.013")

plot(data.wy$MEDHHINC15, data.wy$Fatality_Rate, xlab = "", ylab= "Fatality_Rate",
      col=ifelse( data.wy$Fatality_Rate> quantile(data.wy$Fatality_Rate,
          na.rm= TRUE, probs = c(0.95)) , "red", "darkblue" ), pch = 17)
abline(lm(Fatality_Rate ~ MEDHHINC15, data.wy), col = "blue")
#abline (v = quantile(data.wy$MEDHHINC15, na.rm= TRUE, probs = c(0.05, 0.95)))
text(data.wy$MEDHHINC15,data.wy$Fatality_Rate, labels=data.wy$County, cex= 0.7, pos=1)
text(60000, 0.025, "p-value = 0.2809")
text(60000, 0.024, "R-squared = 0.010")

mtext('Median Household Income', side = 1, outer = TRUE, line = 1)

```

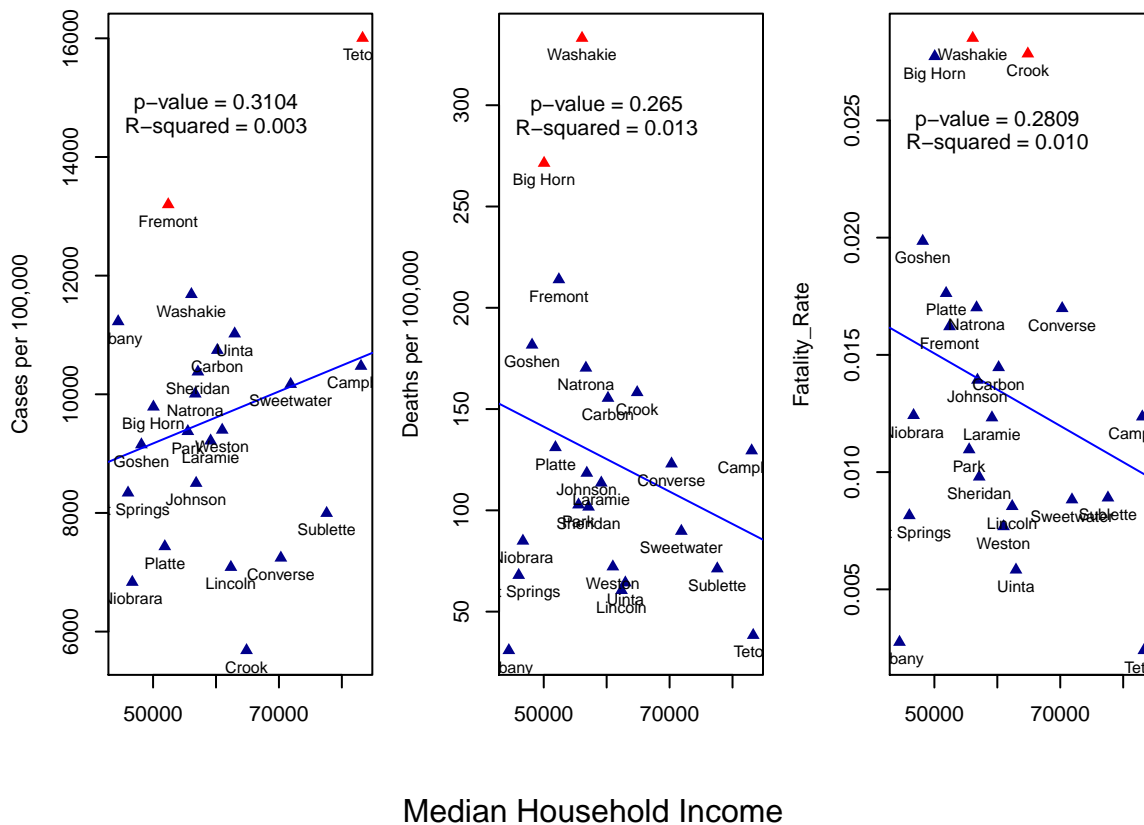


Fig 10 Comparing Median Household Income to COVID-19 outcomes. Red points are outliers for only COVID-19 outcomes.

Discussion

In **Figure 8** it can be identified that Wyoming follows the national trends for population, cases and deaths. To additionally identify that both nationally and locally, the trends are positively correlated with a high significance and a large R-squared value. Between either conditions it can be noted that population correlates stronger with cases opposed to deaths, indicating deaths by COVID-19 are driven by other variables like comorbidities.

Figure 9 points out two strong relationships between population and SNAPs and WIC store access indicated

by the low p-value and the high r-squared. This relationship is consistent at the national and local levels opposed to the loss of significance of both farmers markets and percent free lunches noted by their high p-value and low r-squared.

Figure 10 identifies that the national and local trends do not always match, as noted in the first graphic comparing cases per 100,000 and median household income. This relationship may be due to Teton County, in which another microcapstone group looked into the recreation and tourism and its effects on local counties, also identifying Teton County as an outlier. In the graphics looking at deaths per 100,000 and fatality rate, it is noted that both follow the national trend but additionally identify three additional counties of interest for their outlying data. Washakie, Big Horn and Crook Counties are all outliers for at least one of the two COVID-19 outcomes and with further analysis of current data it was unclear to what variable could have driven these counties outside of the normal.

Summary

The relationships found between food environment, COVID-19 Outcomes, and population, help to connect the importance of strong food environments and success. As noted above the relationship of SNAPs and WIC access to population was strong additional data such as SNAP and WIC usage may better connect the SNAP and WIC variable to COVID-19 outcomes and demographic data like median household income. As much of the data used in this project was from pre 2017 newer data may soon be available to add to our collection and analysis.

Of the weak relationships between food environment and median household income at the local level it brings to question, if there is a better demographic variable, like homelessness or poverty, that may better encompass the co-conditions experienced with food insecurity, and such better connect COVID-19 to food environment.

The IRB Process

Summary

Nutritional Advantage, in conjunction to the quantitative analysis of the COVID-19 outcomes and food environments, wanted to have a better connection and understanding of how people viewed COVID-19 in relation to their food environment. With this we began the process to study human subjects. This process starts with completion of the CITI Human Research Subjects training course, requires the writing up of the proposed research, and approval by the University of Wyoming IRB (institutional review board) before any research can begin.

Discussion

Our interest in collecting these personal stories and experiences came from a conversation early in the fall semester with several community members of the WRIR. With much of the discussion being about the importance of how, and who tells the story it became clear that personal testimonies from the pandemic could give better insights to the inner workings, hard falls and experiences that were had in relation to the local food networks that quantitative data could not.

In the completion of the CITI Training, a number of modules looked at historical accounts where human rights were violated during research. With time many of the violations developed into change via laws and policies that would better protect the subjects of the research. Additional modules covered the importance of protecting at risk populations, such as children, pregnant women, and inmates in addition to the concerns that arise when using computer technology to run research. The second of the concerns being where we as a team needed to delve deeper as we hoped our stories could be collected via email or google form with minimal risk to our research subjects.

The process of writing the IRB was a struggle at times, as the forms questions are often vague and can be understood in a number of ways. With the assistance of previously approved IRBs, from various academic resources, the requirements of the form became more clear. As a team we wrote and revised the IRB form,

and the supplemental poster, informed consent form, and response page. Our submission took the whole 15 business days before getting back to us in a meeting over zoom.

In this meeting several holes to our project proposal were revealed. High on this list was: how do we ensure informed consent, can we ensure participants will remain anonymous and or confidential, and general clarification of what we plan on doing with the data? Our work began to refine our proposal to better describe these ideas. For informed consent it was recommended by the IRB to have the consent process be over zoom. This would allow for a face to face interaction, the reading of the research consent form, the ability for the researcher present to answer questions and ensure the participant was of age, or with legal guardian present. In regards to participants remaining anonymous versus confidential, we decided it would be best to only present the works that could be kept confidential, meaning any identifiers are removed from the work. We began expanding our research plan, such that a savvy subject would be informed to the fullest extent about our project, such we were planning to analyse works by common themes, and county.

Further discussion around how the works were to remain anonymous and confidential led our team to decide against the IRB process, thus preventing us from analyzing the works. Our concern was rooted in the presentation of art works without being able to give proper credit to the artist. The alternative route we have decided to pursue is by maintaining our current community call for art based upon individuals experiences with COVID-19 and their food environment. Again, although we will not be able to formally analyze these works, it will allow individuals to take credit for their contributions in a meaningful way.

Conclusion

In this work we identify population being a major driver towards the food environment variables and COVID-19 outcomes. With the understanding that population correlates with more severe COVID-19 outcomes it becomes a concern that these localities also need to be the first in stopping the spread of the virus. In building experience, education and future plans out of the COVID-19 pandemic, future novel viruses may be better controlled, preventing countless deaths associated with the lack of public education on diseases and many governments reluctance to act.

The connection between population, increasing demographic diversity, and increasing food security programs and COVID-19 outcomes highlights another potential issue. It identifies populations of already marginalized individuals may also be the same individuals who suffer with food insecurity historically in addition to suffering from the disproportionate impacts from COVID-19. To make these communities more resilient to future adversities, more than increasing access to stores, and food programs like SNAP, WIC, and Free and reduced School lunch programs will have to be done. Not just providing more nutritious and low cost options in these communities but building sustainable food networks through local food systems and education about and action towards antiracist food systems.

References

Alcendor, Donald J. 2020. "Racial Disparities-Associated COVID-19 Mortality among Minority Populations in the US." *Journal of Clinical Medicine*, 9(8), 2442. doi: 10.3390/jcm9082442.

Centers for Disease Control and Prevention (CDC). 2020. "Coronavirus Disease 2019 (COVID-19) Data Visualization." U.S. Department of Health & Human Services. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/data-visualization.html>

Centers for Disease Control and Prevention (CDC). 2020. "Healthy Weight, Nutrition, and Physical Activity: About Adult BMI". U.S. Department of Health & Human Services. https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html

Coleman-Jensen, Alisha, Rabbitt, Matthew P., Gregory, Christian A., and Anita Singh. "Household Food Security in the United States in 2019." Economic Research Report no. 275, U.S. Department of Agriculture, Economic Research Service, 2020.

- Feeding America. n.d. "Hunger in America." Accessed September 30, 2020. <https://www.feedingamerica.org/hunger-in-america>
- Kassir, Radwan. 2020. "Risk of COVID-19 for patients with obesity." *Obesity reviews: an official Journal of the International Association for the Study of Obesity*, 21(6), e13034. <https://doi.org/10.1111/obr.13034>
- Law, Alma. 2020. Food Equity Initiatives in Fremont County. September 14, 4pm.
- Love, Charlotte V., Taniguchi, Tori E., Williams, Mary B., Noonan, Carolyn J., Wetherill, Marianna S., Salvatore, Alicia L., Jacob, Tvli, Cannady, Tamela K., Standridge, Joy, Spiegel, Jennifer, and Valarie Jernigan. 2018. "Diabetes and Obesity Associated with Poor Food Environments in American Indian Communities: The Tribal Health and Resilience in Vulnerable Environments (THRIVE) Study." *Current Developments in Nutrition*, 3(Suppl 2), 63–68. <https://doi.org/10.1093/cdn/nzy099>
- Melvin, Sandra C., Wiggins, Corey, Burse, Nakeitra, Thompson, Erica, and Maude Monger. 2020. "The Role of Public Health in COVID-19 Emergency Response Efforts From a Rural Health Perspective." *Preventing Chronic Disease*, 17(E70). <https://doi.org/10.5888/pcd17.200256>
- Patrick, Stephen W., Henkhaus, Laura E., Zickafoose, Joseph S., Lovell, Kim, Halvorson, Alese, Loch, Sarah, Letterie, Mia, and Matthew M. Davis. 2020. "Well-Being of Parents and Children During the COVID-19 Pandemic: A National Survey." *Pediatrics*. <https://doi.org/10.1542/peds.2020-016824>.
- Ralston, Katherine, Treen, Katie, Coleman-Jensen, Alisha, and Joanne Guthrie. "Children's Food Security and USDA Child Nutrition Programs." *Economic Information Bulletin no. 174*, U.S. Department of Agriculture, Economic Research Service, 2017.
- Wyoming Hunger Initiative. 2019. "2019 Annual Report." Wyoming Hunger Initiative, 2019. <https://www.nohungerwyo.org/>
- Wyoming Office of Health Equity's Chronic Disease Prevention Program. 2020. "Focus on Healing." Wyoming Department of Health. <https://health.wyo.gov/wp-content/uploads/2020/08/2020-WRIR-Publ-Focus-on-Healing-f-1.pdf>