Final Presentation

```
library(tidyverse)
library(tidymodels)
```

```
food <- readr::read_csv("data/Food_Supply_kcal_Data.csv")
```

#Introduction and Data:

We are interested in looking at the relationship between COVID-19 outcomes and nutrition worldwide. The USDA Center for Nutrition Policy and Promotion suggests a dietary intake which consists of 30% grains, 40% vegetable, 10% fruits, and 20% proteins (dietaryguidlines.gov). Previous studies demonstrate an increased mortality in patients infected with COVID-19 which have chronic inflammatory diseases such as obesity, diabetes, and hypertension.

The prevalence of these chronic inflammatory diseases are known to be correlated with an individual's diet (Onishi 2020). Furthermore, previous studies show that maintaining a healthy diet can decrease risk of severe infection by promoting the immune system (Messina et al. 2020, Iddir et. al 2020). Adequate protein consumption is essential for antibody production and poor nutrient consumption has been shown to increase inflammation and oxidative stress (Iddir et. al 2020). We are ultimately interested in seeing if countries that tend to consume similar diets to those suggested by the USDA show decreased rates of mortality from COVID-19 and how it/if it is related to income levels.

This dataset, "COVID-19 Healthy Diet Dataset" comes from Kaggle. The dataset provides energy intake (kcal) as percentages of total diet by food group. In addition, it provides percentages of obesity and undernourished individuals. Finally it provides data for total confirmed COVID-19 cases, recovered COVID cases, COVID deaths, and active COVID cases for 170 countries. The food supply quantities in addition to the prevalence of obesity and undernourishment in the populations were obtained from the Food and Agricultural Organization of the United Nations, the population count was taken from the Population Reference Bureau, and the Johns Hopkins Center for Systems Science and Engineering was used for COVID-19 data and was last updated in Feburary of 2021.

Works Cited:

Iddir M., et al. Strengthening the Immune System and Reducing Inflammation and Oxidative Stress through Diet and Nutrition: Considerations during the COVID-19 Crisis. Nutrients. 2020;12(6):1562.

Messina G., et al. Functional Role of Dietary Intervention to Improve the Outcome of COVID-19: A Hypothesis of Work. International Journal of Molecular Sciences. 2020; 21(9):3104

Onishi J., et al. Can Dietary Fatty Acids Affect the COVID-19 Infection Outcome in Vulnerable Populations? mBio. 11(4).

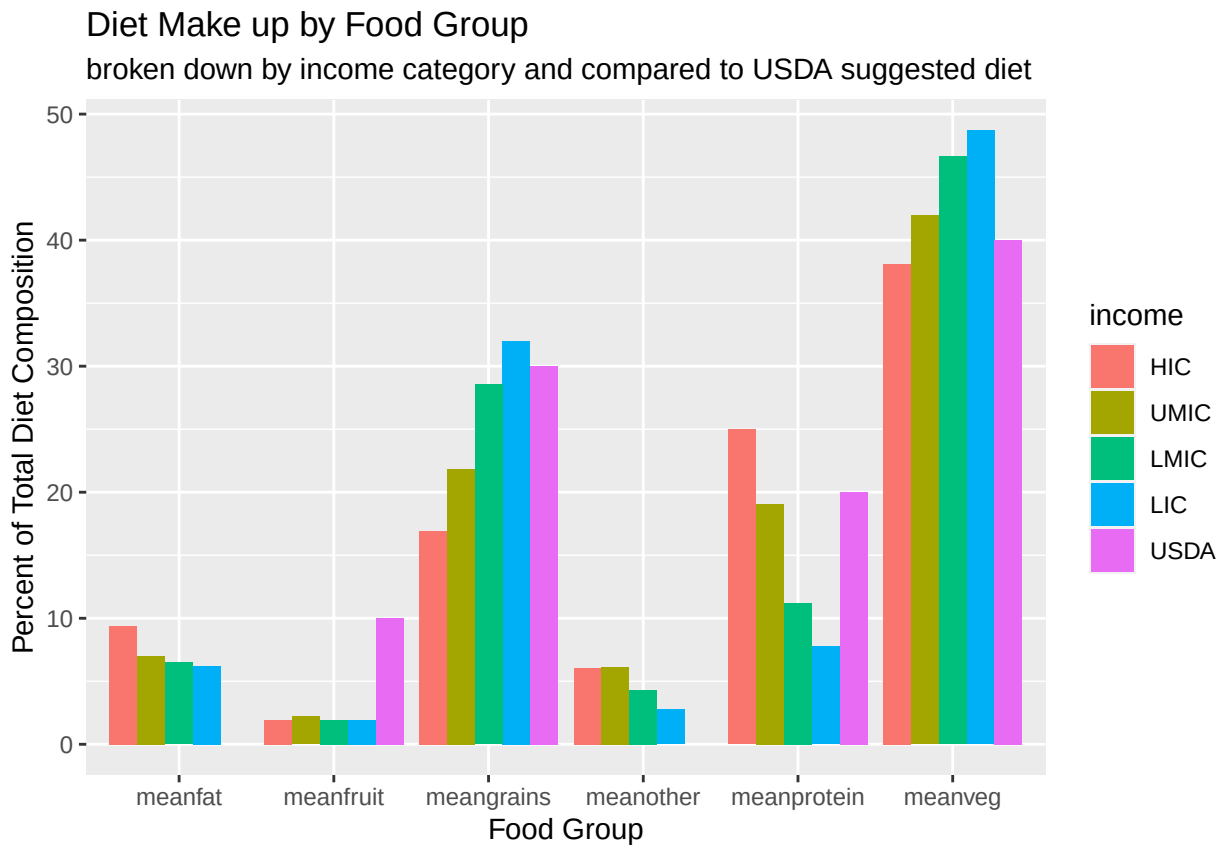
INCOME DATA FROM:

```
food <- food %>%
  mutate(income = ifelse(Country %in% c("Afghanistan","Burkina Faso", "Central African Republic", "Chad
                         ifelse(Country %in% c("Algeria","Angola","Bangladesh", "Belize", "Benin", "Boli
Pakistan", "Samoa", "Sao Tome and Principe", "Senegal", "Solomon Islands","Sri Lanka", "Suriname", "Taji
                         ifelse(Country %in% c("Albania","Argentina","Armenia","Azerbaijan", "Belarus",
Panama", "Paraguay", "Peru", "Philippines", "Republic of Moldova", "Romania", "Russian Federation", "Sa
             "UMIC", "HIC"))))
```

```r
food <- food %>%
  mutate(Grains = `Cereals - Excluding Beer` + `Starchy Roots`) %>%
  mutate(Vegetables2 = Pulses + `Vegetal Products` + Vegetables) %>%
  mutate(Fruits = `Fruits - Excluding Wine`) %>%
  mutate(Fats = `Animal fats` + Oilcrops + Treenuts + `Vegetable Oils`) %>%
  mutate(Protein = `Animal Products` + Eggs + `Fish, Seafood` + Meat + `Milk - Excluding Butter` + Offal
  mutate('Alcohol/Stimulants' = `Alcoholic Beverages` + Stimulants) %>%
  mutate(Other = `Aquatic Products, Other` + Miscellaneous + Spices + `Sugar Crops` + `Sugar & Sweetener
```

```r
meanfood <- food %>%
  summarize(meangrains = mean(Grains), meanveg = mean(Vegetables2), meanfruit = mean(Fruits), meanfat =

meanfoodincome <- food %>%
  group_by(income) %>%
  summarize(meangrains = mean(Grains), meanveg = mean(Vegetables2), meanfruit = mean(Fruits), meanfat =
  add_row(income="USDA", meangrains=30, meanveg=40, meanfruit=10, meanprotein=20) %>%
  pivot_longer(cols=meangrains:meanother, names_to = "Means", values_to = "PercentDiet")

meanfoodincome %>%
  mutate(income = fct_relevel(income,
            "HIC", "UMIC", "LMIC",
            "LIC", "USDA")) %>%
  ggplot(aes(x=Means, y=PercentDiet, fill=income))+geom_col(position="dodge") + labs(x="Food Group", y=
```

```
## Warning: Removed 2 rows containing missing values (geom_col).
```

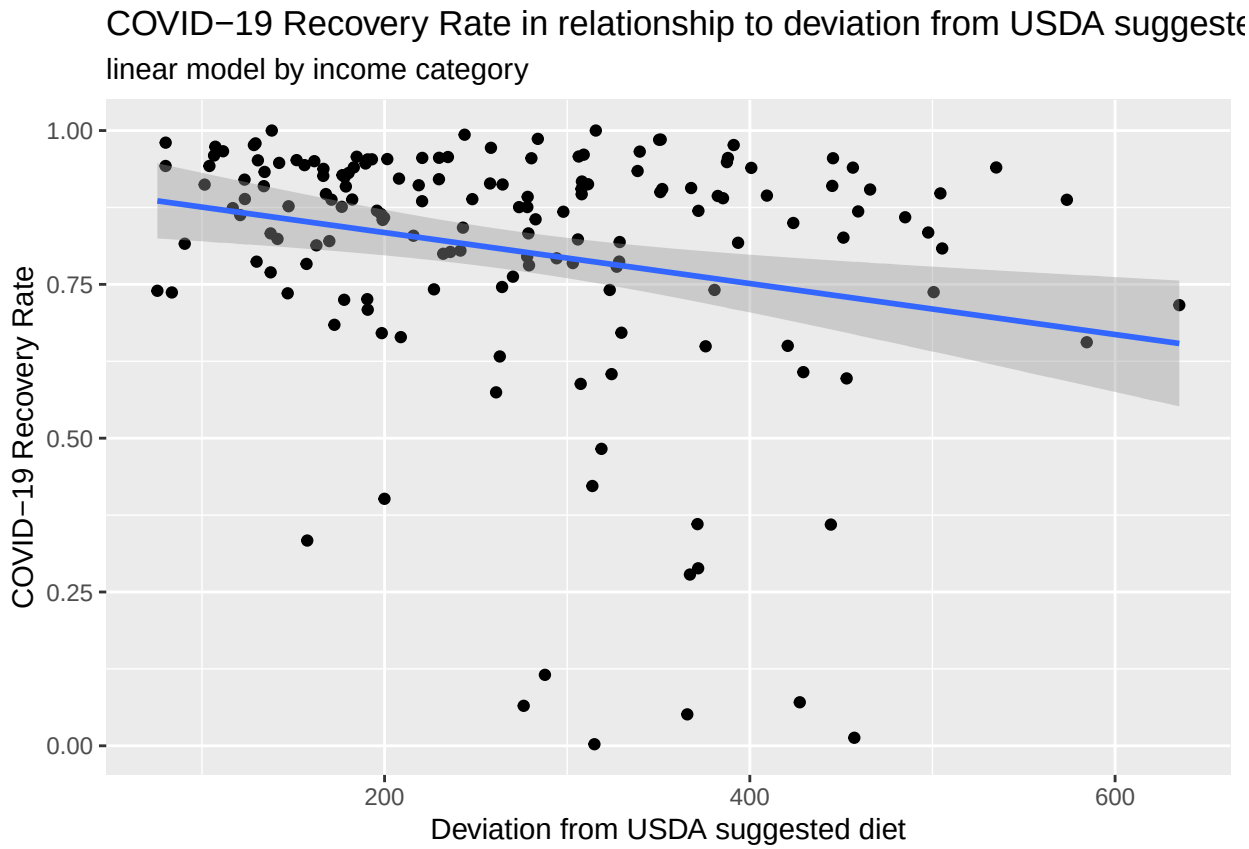


```
food <- food %>%
   mutate(rec_rate=Recovered/(Confirmed)) %>%
  mutate(adherance_grains = (30 - Grains)^2) %>%
  mutate(adherance_veg = (40 - Vegetables2)^2) %>%
  mutate(adherance_fruit = (10 - Fruits)^2) %>%
  mutate(adherance_protein = (20 - Protein)^2) %>%
  mutate(adherance_tot = adherance_grains + adherance_protein + adherance_fruit + adherance_veg) %>%
  filter(rec_rate>0)

tfood<-food %>%
  group_by(income) %>%
  summarize(meangrains = mean(Grains), meanveg = mean(Vegetables2), meanfruit = mean(Fruits), meanfat =
  add_row(income="USDA", meangrains=30, meanveg=40, meanfruit=10, meanprotein=20)

summary(aov(rec_rate~income,data=food))
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## income        3  0.324 0.10785   2.546  0.058 .
## Residuals   156  6.607 0.04236
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

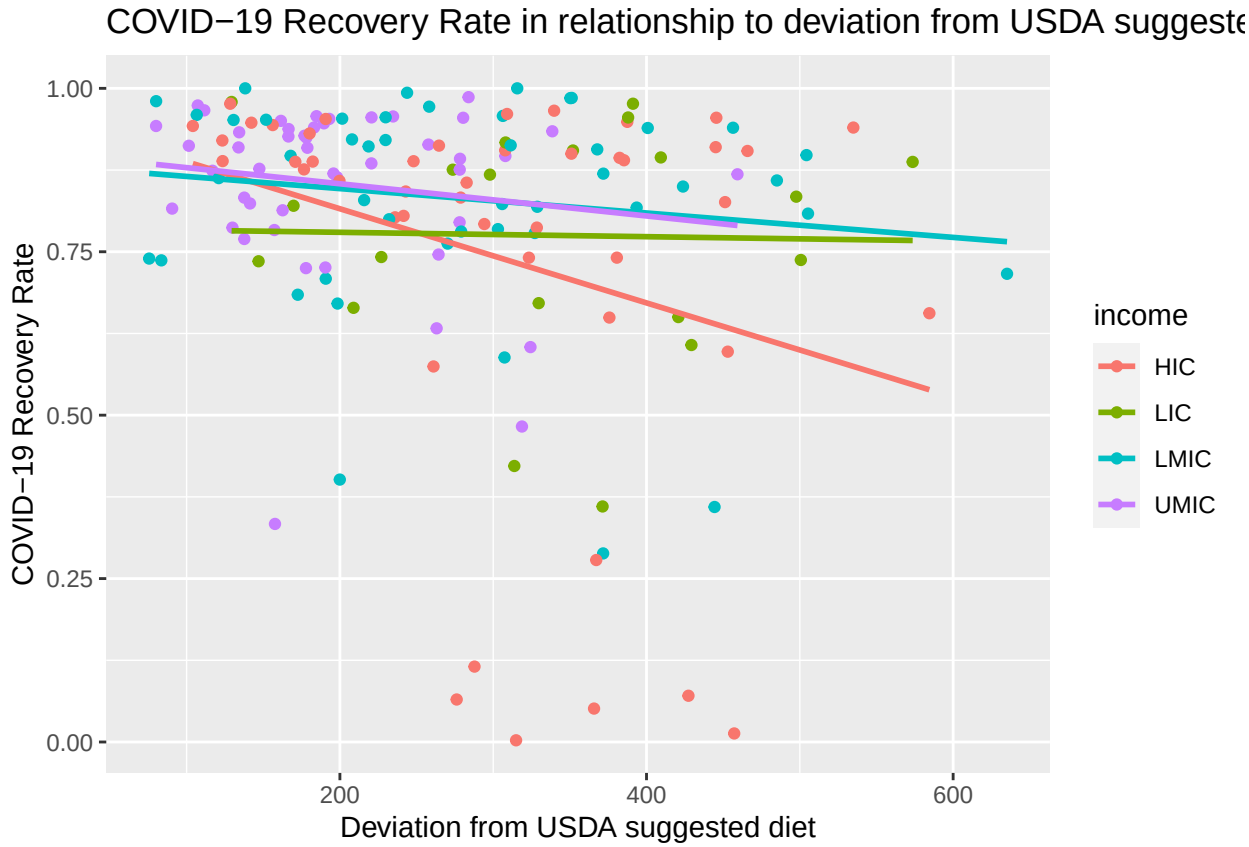
```
food <- food %>%
   mutate(rec_rate=Recovered/(Confirmed)) %>%
  mutate(adherance_grains = (30 - Grains)^2) %>%
  mutate(adherance_veg = (40 - Vegetables2)^2) %>%
  mutate(adherance_fruit = (10 - Fruits)^2) %>%
  mutate(adherance_protein = (20 - Protein)^2) %>%
  mutate(adherance_tot = adherance_grains + adherance_protein + adherance_fruit + adherance_veg) %>%
  filter(rec_rate>0)

food %>%
  filter(adherance_tot<750) %>%
ggplot(aes(x = adherance_tot, y = rec_rate)) +
  geom_point()+geom_smooth(method="lm") + labs(
    title = "COVID-19 Recovery Rate in relationship to deviation from USDA suggested Dietary Intake",
    subtitle="linear model by income category",
    x = "Deviation from USDA suggested diet",
    y = "COVID-19 Recovery Rate")
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
food %>%
  filter(adherance_tot<750) %>%
ggplot(aes(x = adherance_tot, y = rec_rate, color=income)) +
  geom_point()+geom_smooth(method="lm", se = FALSE) + labs(
    title = "COVID-19 Recovery Rate in relationship to deviation from USDA suggested Dietary Intake",
    x = "Deviation from USDA suggested diet",
    y = "COVID-19 Recovery Rate")
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
res <- linear_reg() %>%
  set_engine("lm") %>%
  fit(rec_rate ~ adherance_tot, data = food) %>%
  tidy() %>%
  print()
```

```
## # A tibble: 2 x 5
##   term           estimate std.error statistic  p.value
##   <chr>             <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)    0.886     0.0377       23.5  1.83e-53
## 2 adherance_tot -0.000292  0.000123     -2.36 1.94e- 2
```