Dietary Patterns in a Global Crisis:
Using COVID-19 as a Lens to Identify Optimal Nutrition for Health

CSCI 114 Final Project - Kailin Liu

I. INTRODUCTION

Diets play a key role in preventing chronic diseases, supporting immune function, and promoting overall well-being. Yet, despite its importance, determining the "best" diet remains challenging. Nutritional science offers various dietary recommendations—ranging from plant-based diets and ketogenic regimens to Mediterranean-style eating and high-protein plans. This diversity—stemming from variations in individual needs, cultural preferences, and study methods—makes it difficult to identify a universally optimal diet. These challenges are further compounded by global inequalities in food access and nutrient availability, which shape dietary patterns and health outcomes across populations.

The COVID-19 pandemic provides a unique test case to explore this gray area. While numerous factors are at play specific differences in COVID outcomes across nations and populations may offer insights into the role of diet and nutrient supply in influencing resilience and recovery. For instance, countries with diets rich in anti-inflammatory and immune-supporting nutrients—such as vitamins A, C, and D, zinc, and omega-3 fatty acids—may have experienced milder disease impacts compared to those where diets were characterized by processed foods and low nutrient density[1][2]. Conversely, dietary patterns characterized by high consumption of processed foods are linked to increased levels of pro-inflammatory biomarkers, such as C-reactive protein and interleukin-6 (IL-6). Furthermore, studies using the Dietary Inflammatory Index (DII) have shown that individuals with higher DII scores, indicating diets with greater inflammatory potential, experienced worse COVID-19 outcomes, including prolonged hospitalizations and severe symptoms [3]. Thus, analyzing how the pandemic correlates with national nutrient supply data can offer insights into identifying which nutritional habits are the most effective at supporting immune function and reducing disease severity.

This project utilizes bioinformatics data to analyze global nutrient supply during the pandemic, exploring how variations in diet and food systems may have impacted COVID-19 outcomes. By examining this unprecedented global crisis, I aim to explore the broader question of which dietary patterns may offer the greatest health resiliency. By analyzing these connections, we may infer how distribution of nutrients and eating habits can strengthen immunity and mitigate disease severity. Ultimately, this knowledge not only informs dietary recommendations to improve individual health, but can also offer guidance for public health strategies in promoting population-level resilience against future health challenges.

II. HYPOTHESIS

There is a significant relationship between the diversity of a country's food supply and its public health resilience to a global pandemic, such as COVID-19. Specifically, countries with more diverse food supplies are likely to exhibit better public health outcomes, including lower COVID-19 mortality rates and higher overall health indicators, compared to countries with less diverse food supplies. Furthermore, variation in distribution of energy intake across countries may correlate with certain recovery outcomes.

III. METHODS

1. Data Collection

The dataset used in this project was sourced from the COVID-19 Healthy Diet Dataset from the Kaggle database [4]. Specifically, the Food_Supply_kcal_Data.csv was used. The file includes kilocalorie based nutritional information (e.g., supply of food products) and health-related statistics (e.g., obesity rates, COVID-19 metrics) for multiple countries. The CSV file contains columns for different food categories and health metrics, and rows for individual countries.

2. Software and Code

A. Language: C++ (C++17 standard)

B. Editor: Visual Studio Code (Version 1.95.3) on Mac OS Sequoia 15.0.1

C. Libraries:

   a. Standard Libraries: iostream, fstream, sstream, map, vector, string, algorithm, random.

   b. Visualization: matplot for generating plots and charts [5]

D. Compilation

   a. The code was compiled in terminal using:

   clang++ -g final_proj.cpp -o finalproj -std=c++17 -I/opt/homebrew/include -ldlib -lmatplot -lblas -L/opt/homebrew/lib

E. Code availability: https://github.com/usc-csci114/assignments-liu-kailin/tree/main/project

3. Data Preparation

The data was parsed using the parseCSV function. This function reads the CSV file and populates a map structure where each key is a country name, and the value is a vector of doubles corresponding to the data for that country. The steps involved are:

1. Reading Column Names: The column names were hard-coded to account for CSV headers and special cases with commas in country names.
2. Identifying Indices: The function identified indices of relevant columns (e.g., "Animal Products", "Meat", "Vegetables") for subsequent calculations.
3. Handling Special Cases: Specific country names with commas (e.g., "Korea, North") were handled to avoid parsing issues.
4. Calculations: Additional columns were dynamically calculated, such as:
   - "Meat Products": Sum of "Animal Products" and "Meat".
   - "Vegetables": Sum of "Vegetables" and "Vegetal Products".

   NOTE: Missing or invalid values were defaulted to -1.0 to signify data issues.

4. Data Analysis

A. Subsetting Countries by Obesity Rate

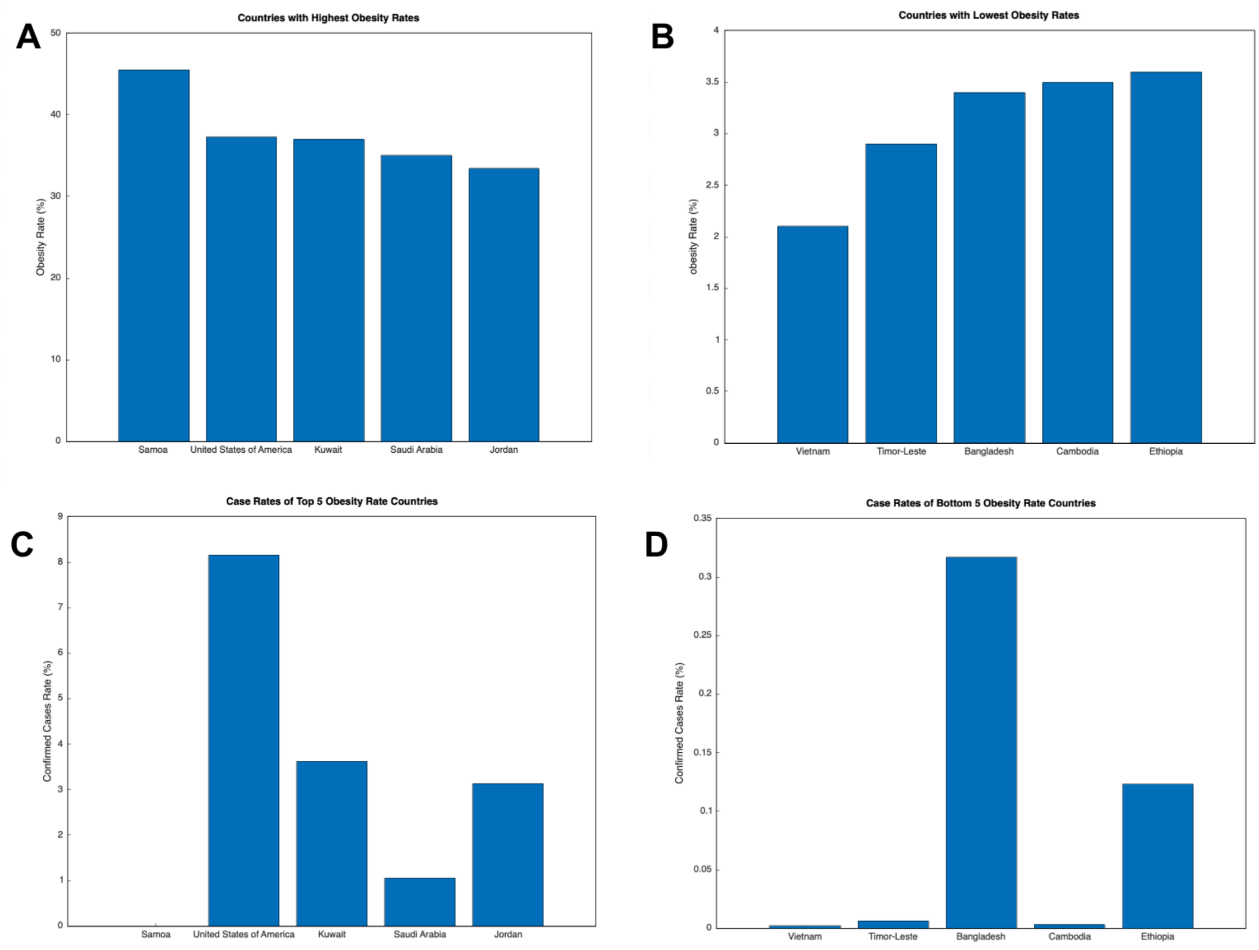
The function subsetCountries was used to identify the top 5 and bottom 5 countries based on obesity rates. The steps included:

a. Max-Heap and Min-Heap: Two heaps were employed:
    i. A max-heap to track countries with the highest obesity rates.
    ii. A min-heap to track countries with the lowest obesity rates.
b. Exclusions: Special cases like "Kiribati", "Korea, North", and "Korea, South" were excluded from the analysis due to parsing complications.

B. Visualization
    a. Three types of visualizations were generated to analyze the relationship between food supply and health metrics:
        i. Bar Plots (makeBP function): Simple bar plots showing the obesity rates of the top 5 and bottom 5 countries.
        ii. Grouped Bar Charts (makeGroupedBarCharts function):
            1. Grouped bar charts to visualize the distribution of key food categories (e.g., "Fish, Seafood", "Meat Products") for the countries with the highest and lowest obesity rates.
        iii. Scatter Plots (makeScatterPlots function):
            1. Scatter plots displaying the relationship between food supply (e.g., "Fish, Seafood") and COVID-19 metrics (e.g., "Confirmed", "Deaths").
            2. Linear Regression: For each scatter plot, a regression line was fitted using the calculateLinearRegression function to explore potential correlations.

IV. ANALYSIS

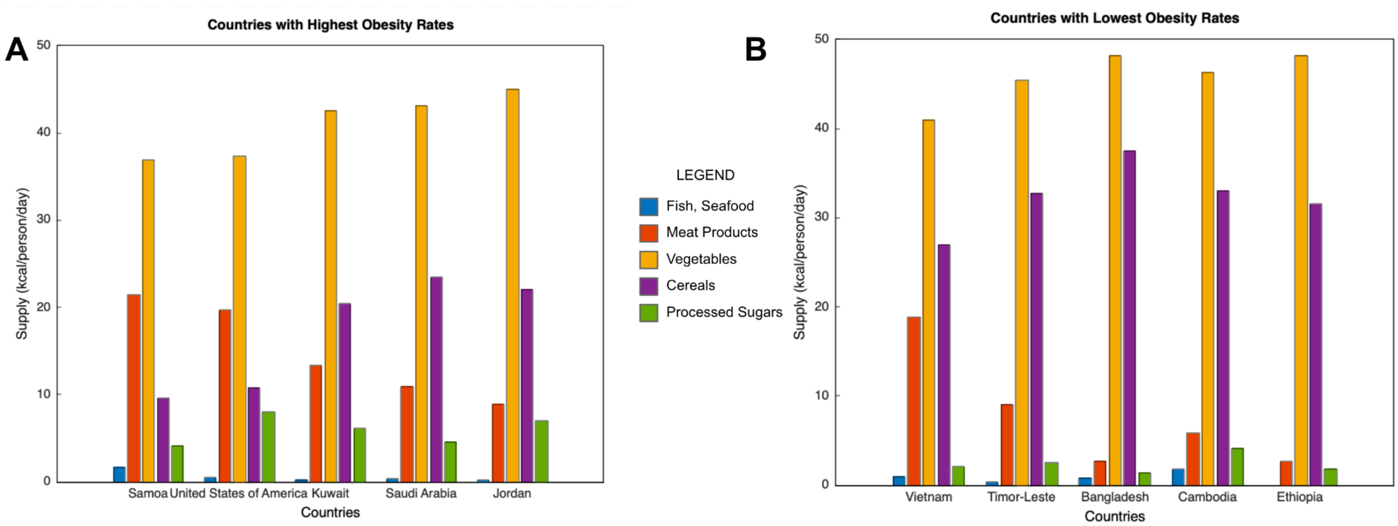
Figure 1: Simple Bar Plots by Top 5 (Highest) Obesity Rate and Bottom 5 (Lowest) Obesity Rate



[1A] Bar plot of top 5/ highest obesity rate countries: Samoa(45.5%) , United States of America (37.3%), Kuwait (37%), Saudi Arabia (35%), Jordan (33.4%). **[1B]** Bar plot of bottom 5/ lowest obesity rate countries: Vietnam (2.1%), Timor-Leste (2.9%) , Bangladesh (3.4%), Cambodia (3.5%), Ethiopia (3.6%). **[1C]** Bar plot of top 5 obesity rate countries plotted against case rate: Samoa(0.001%) , United States of America (8.16%), Kuwait (3.62%), Saudi Arabia (1.05%), Jordan (3.13%). **[1D]** Bar plot of bottom 5 obesity rate countries plotted against case rate: Vietnam (0.002%), Timor-Leste (0.006%) , Bangladesh (0.317%), Cambodia (0.003%), Ethiopia (0.123%).

Overall, the lower obesity rate countries had significantly lower case rates (less than 1%) than the higher obesity rate countries. However, there is no clear trend of case rates among top 5 / highest obesity rate countries. Samoa had a lower case rate than the lowest obesity rate country Vietnam. Without the clear correlation between obesity is not a strong predictor for and case rates. Additional metrics must be explored to see how diet influences COVID-19 resilience.

Figure 2: Grouped Bar Plot of Energy Intake by Countries with Highest and Lowest Obesity Rates

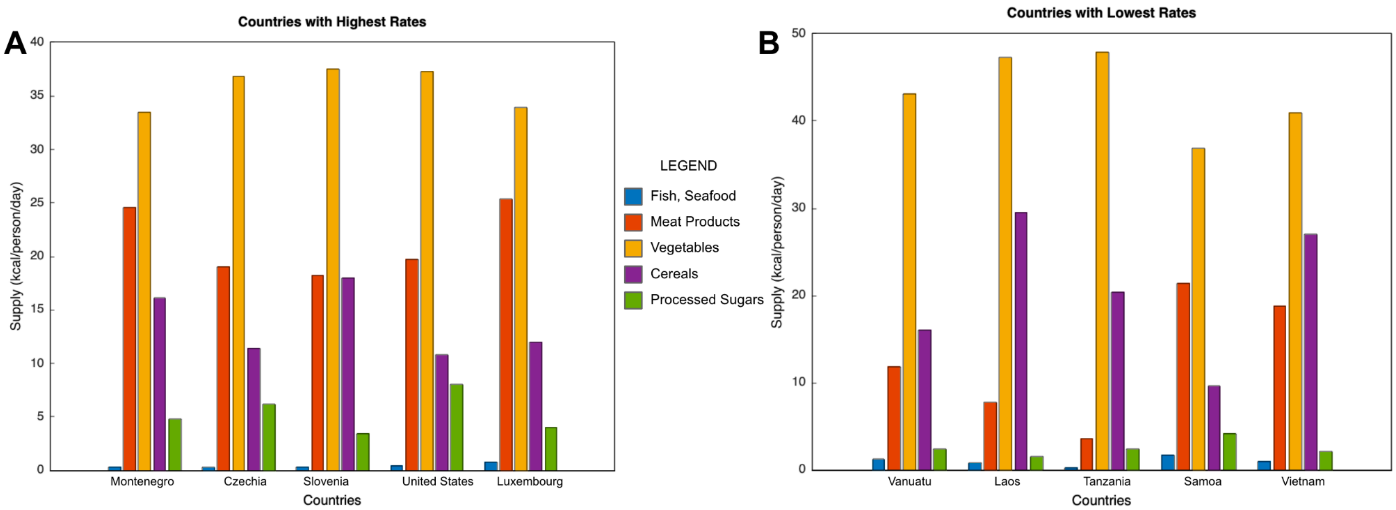


[2A] Grouped bar plot of top 5/ highest obesity rate countries. Each bar represents the % daily energy intake.
<u>SAMOA</u>: Fish, Seafood: 1.7479, Meat Products: 21.4286, Vegetables: 36.8571, Cereals: 9.6134, Processed Sugars: 4.1849
<u>UNITED STATES:</u> Fish, Seafood: 0.4779, Meat Products: 19.7266, Vegetables: 37.3025, Cereals: 10.7792, Processed Sugars: 8.0048
<u>KUWAIT:</u> Fish, Seafood: 0.3048, Meat Products: 13.3382, Vegetables: 42.5689, Cereals: 20.4209, Processed Sugars: 6.1393
<u>SAUDI ARABIA:</u> Fish, Seafood: 0.3286, Meat Products: 10.939, Vegetables: 43.1768, Cereals: 23.4585, Processed Sugars: 4.6479
<u>JORDAN:</u> Fish, Seafood: 0.2211, Meat Products: 8.972, Vegetables: 45.0626, Cereals: 22.0892, Processed Sugars: 6.9823

[2B] Grouped bar plot of bottom 5/ lowest obesity rate countries. Each bar represents the % daily energy intake.
<u>VIETNAM:</u> Fish, Seafood: 1.0385, Meat Products: 18.8117, Vegetables: 40.9942, Cereals: 26.9833, Processed Sugars: 2.094.
<u>TIMOR-LESTE:</u> Fish, Seafood: 0.374, Meat Products: 9.0463, Vegetables: 45.4184, Cereals: 32.7256, Processed Sugars: 2.5245.
<u>BANGLADESH:</u> Fish, Seafood: 0.8284, Meat Products: 2.7163, Vegetables: 48.2566, Cereals: 37.5265, Processed Sugars: 1.4255.
<u>CAMBODIA</u>: Fish, Seafood: 1.7792, Meat Products: 5.8229, Vegetables: 46.3, Cereals: 33.0166, Processed Sugars: 4.1852.
<u>ETHIOPIA:</u> Fish, Seafood: 0.0217, Meat Products: 2.6692, Vegetables: 48.2422, Cereals: 31.5321, Processed Sugars: 1.8663,

Across all 10 countries, vegetables had the highest energy intake percentage among the food group categories. All countries had relatively low fish and seafood intake. The higher obesity countries had an overall greater percentage of Meat Products intake (but the USA's intake was comparable to Vietnam's). The greatest difference is seen in the percentage of cereals. The lower obesity countries had significantly higher energy intake of cereals and somewhat lower energy intake of processed sugars.

Figure 3: Grouped Bar Plot of Energy Intake by Countries with Highest and Lowest COVID-19 Cases

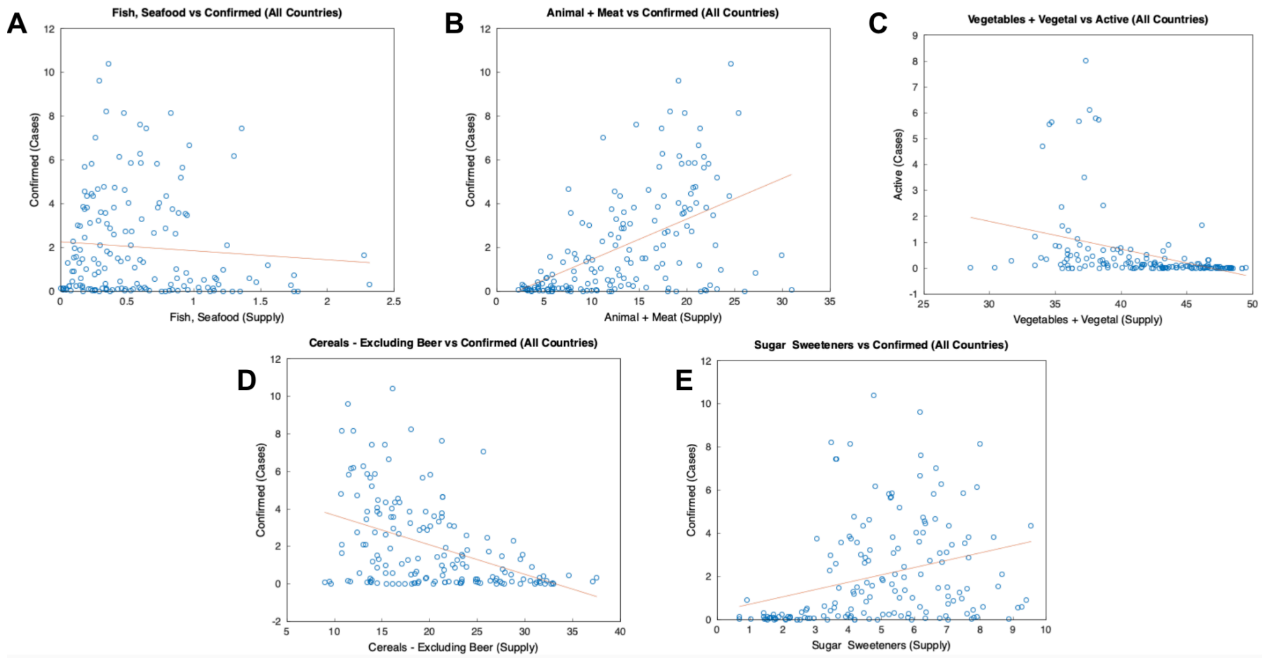


[3A] Grouped bar plot of top 5/ highest confirmed COVID-19 case countries. Each bar represents the % daily energy intake.
MONTENEGRO: Fish, Seafood: 0.3591, Meat Products: 24.5942, Vegetables: 33.4722, Cereals: 16.1184, Processed Sugars: 4.7694,
CZECHIA: Fish, Seafood: 0.29, Meat Products: 19.0781, Vegetables: 36.8742, Cereals: 11.4011, Processed Sugars: 6.1661,
SLOVENIA: Fish, Seafood: 0.3442, Meat Products: 18.2103, Vegetables: 37.5, Cereals: 18.0069, Processed Sugars: 3.4731,
UNITED STATES: Fish, Seafood: 0.4779, Meat Products: 19.7266, Vegetables: 37.3025, Cereals: 10.7792, Processed Sugars: 8.0048,
LUXEMBOURG: Fish, Seafood: 0.8246, Meat Products: 25.3673, Vegetables: 33.898, Cereals: 11.979, Processed Sugars: 4.048,

[3B] Grouped bar plot of bottom 5/ lowest confirmed COVID-19 case countries. Each bar represents the % daily energy intake.
VANUATU: Fish, Seafood: 1.2197, Meat Products: 11.8354, Vegetables: 43.1485, Cereals: 16.1044, Processed Sugars: 2.4204,
LAOS: Fish, Seafood: 0.7877, Meat Products: 7.7781, Vegetables: 47.2596, Cereals: 29.488, Processed Sugars: 1.5425,
TANZANIA: Fish, Seafood: 0.2922, Meat Products: 3.6519, Vegetables: 47.9132, Cereals: 20.409, Processed Sugars: 2.4207,
SAMOA: Fish, Seafood: 1.7479, Meat Products: 21.4286, Vegetables: 36.8571, Cereals: 9.6134, Processed Sugars: 4.1849,
VIETNAM: Fish, Seafood: 1.0385, Meat Products: 18.8117, Vegetables: 40.9942, Cereals: 26.9833, Processed Sugars: 2.094,

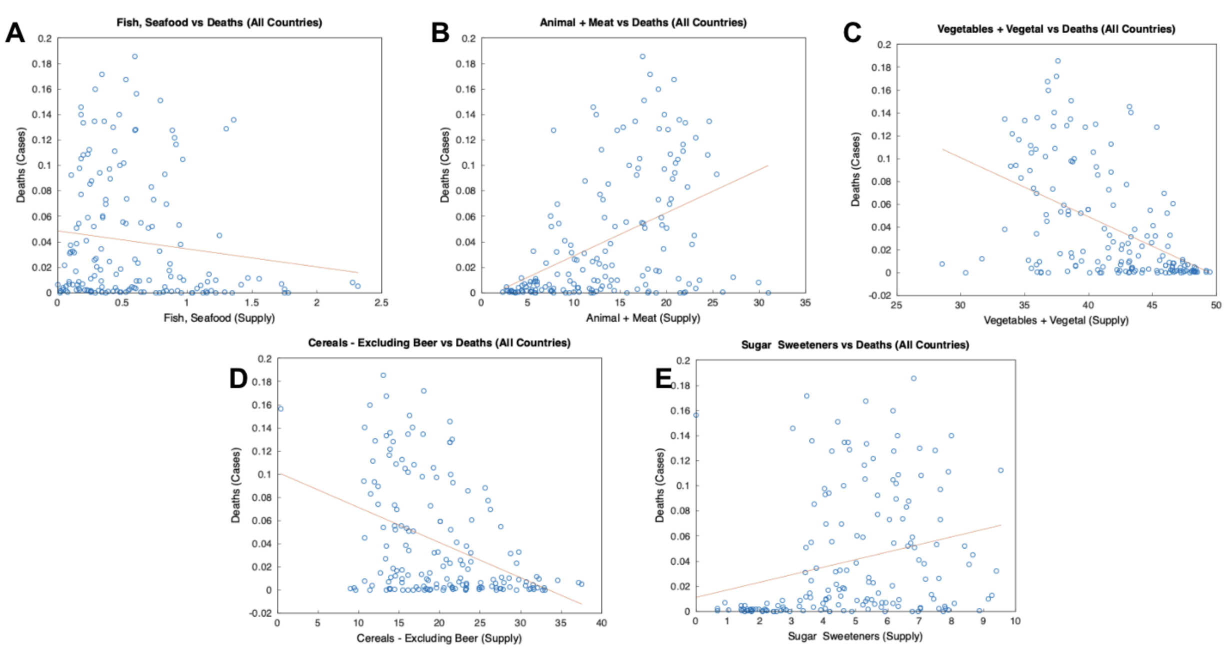
Again, across all 10 countries, vegetables had the highest energy intake percentage among the food group categories while fish and seafood were the lowest. The higher COVID rate countries had an overall greater percentage of Meat Products and Processed Sugars intake. Comparing Figures 2 and 3 there is no clear correlation between obesity and confirmed COVID rate. Samoa and Vietnam were the highest and lowest obesity rate countries respectively, but both were among the lowest COVID rates.

Figure 4: Scatter Plots of Confirmed Cases by Food Group (all countries)



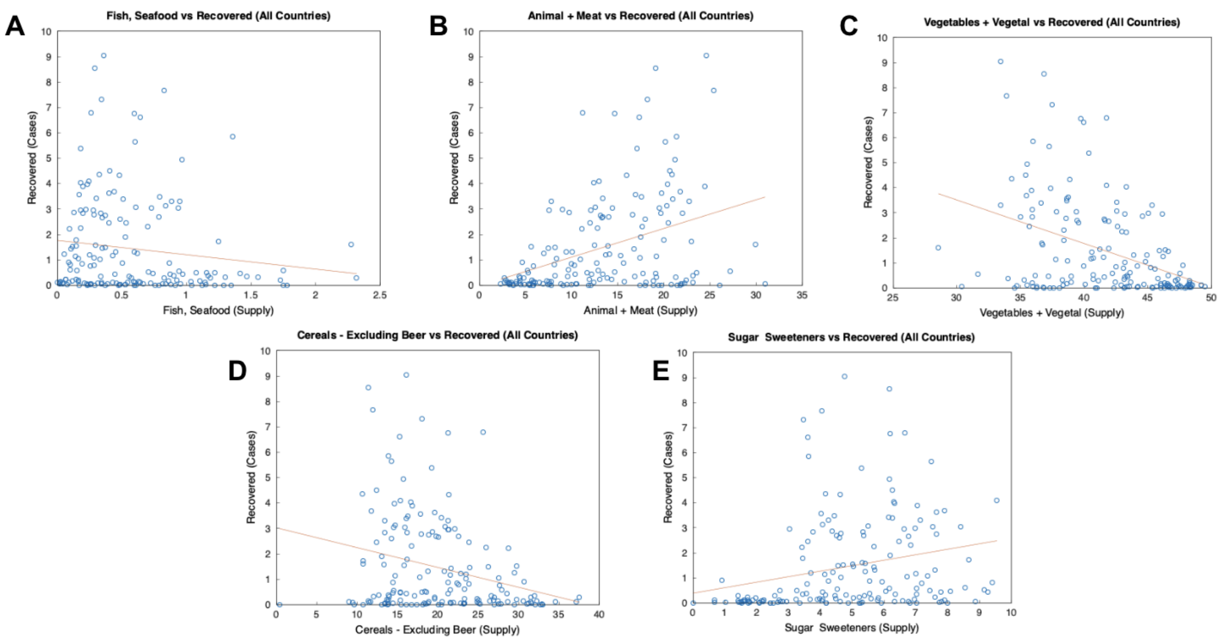
[4A] Scatter plot of percent energy consumption of Fish and Seafood and confirmed COVID-19 case rates. There is a slight downwards trend where greater percent intake is correlated with lower confirmed cases. **[4B]** Scatter plot based on percent energy consumption of Meat Products. There as a clear positive trend with greater meat consumption and higher confirmed cases. However there is also a greater spread as meat consumption increases. **[4C]** Scatter plot based on percent energy consumption of Vegetables. There is a slight downwards trend where greater percent intake is correlated with lower confirmed cases. **[4D]** Scatter plot based on percent energy consumption of Cereals. There is a downward trend where greater percent intake is correlated with lower confirmed cases and the distribution is much more limited at higher intake. **[4E]** Scatter plot based on percent energy consumption of Sugars. There is a marked upwards trend as processed sugar consumption increases.

Figure 5: Scatter Plots of Deaths by Food Group (all countries)



[5A] Scatter plot of percent energy consumption of Fish and Seafood and COVID-19 death rates. There is a downwards trend where greater percent intake is correlated with lower confirmed cases but overall the distribution is varied **[5B]** Scatter plot based on percent energy consumption of Meat Products. There is a clear positive trend with greater meat consumption and higher confirmed cases. However there is also a greater spread as meat consumption increases. **[5C]** Scatter plot based on percent energy consumption of Vegetables. There is a stronger downwards trend where greater percent intake is correlated with fewer deaths. **[5D]** Scatter plot based on percent energy consumption of Cereals. There is a downward trend where greater percent intake is correlated with lower deaths. However there is also a wide spread. **[5E]** Scatter plot based on percent energy consumption of Sugars. There is a upwards trend as but variation greatly increases around 3% energy intake.

Figure 6: Scatter Plots of Recovery by Food Group (all countries)



[6A] Scatter plot of percent energy consumption of Fish and Seafood and COVID-19 death rates. There is a slight downwards trend where greater percent intake is correlated with lower confirmed cases. However, there are also much fewer countries with higher seafood consumption **[6B]** Scatter plot based on percent energy consumption of Meat Products. There positive trend with greater meat consumption and higher confirmed cases. However there is also a greater spread and fewer points as meat consumption increases. **[6C]** Scatter plot based on percent energy consumption of Vegetables. There is a downwards trend where greater percent intake is correlated with lower confirmed cases. **[6D]** Scatter plot based on percent energy consumption of Cereals. There is a downward trend where greater percent intake is correlated with lower recovery cases. However, there is a wide spread overall. **[6E]** Scatter plot based on percent energy consumption of Sugars. There is an upwards trend as processed sugar consumption increases but there is also a wide spread overall.

Unexpectedly, the trends remained consistent for each food group across different metrics. For example, Fish and Seafood displayed a downward trend in both Deaths and Recovery. Utilizing the information from previous figures, this could be correlated with countries with higher obesity and consumption is correlated with higher confirmed cases.

Despite the consistency, there are some specific trends and variations worth noting. Based on Figure 2, Vegetables and Cereals were two of the highest food groups, with Cereals having the greatest difference between higher obesity countries and lower obesity countries. Looking first at vegetables, comparing to Figure 3C and Figure 3D, there is a stronger correlation between greater vegetable energy consumption and lower death rates. The distribution of vegetable energy intake is less widespread/more concentrated. However, the trend for cereals is less prominent. In Figure 3D, there is a clear correlation between higher cereal intake and lower confirmed cases, but the wide distribution in figures 4D and 5D make prognosis based on cereal intake difficult to conclude.

V. CONCLUSION

Discussion

Obesity is a well-documented risk factor for a wide range of health complications, including cardiovascular diseases, diabetes, and certain cancers. Its role as a driver of metabolic dysfunction is particularly critical, as excess body fat leads to chronic inflammation, insulin resistance, and hormonal imbalances that increase disease susceptibility [7]. Thus, my first path of analysis was to verify the link between obesity and COVID-19 response, given that obesity is linked to compromised immune function and respiratory challenges, and subsequently diminished health outcomes. In Figure 1 demonstrates that countries with lower obesity rates had an overall lower COVID-19 case rate (Figure 1D), with case rates in low-obesity countries like Vietnam (0.002%) and Timor-Leste (0.006%) being far lower than those in high-obesity countries like the United States (8.16%) and Kuwait (3.62%) (Figure 1C). However, looking within high-obesity countries (Figure 1A and 1C), there is no clear trend in case rates. For example, Samoa had both the highest obesity rate (45.5%) and one of the lowest case rates (0.001%). This suggests that obesity alone is not a definitive predictor of COVID-19 cases.

Subsequently, if obesity alone is not a clear factor of COVID-19 cases, but overall obesity does play a role in modulating response, I further explored dietary distribution to find potential determinants of COVID-19 susceptability. In Figures 2 and 3, I explored the diets of countries subsetted for obesity and COVID-19 case rate respectively.

In Figure 2, regardless of obesity levels, vegetables consistently represented the highest percentage of energy intake while fish and seafood represented the lowest percentage. Comparing across obesity levels, higher-obesity countries consume a greater percentage of meat products and processed sugars, while lower-obesity countries consume more cereals and less processed sugar (Figure 2A and 2B). A notable comparison is that the United States (second highest obesity) consumed 19.7% of daily energy from meat products and 8% from processed sugars. In contrast, Ethiopia (second lowest obesity) consumed only 2.7% from meat products and 1.9% from processed sugars.

Following the trend from Figure 1, where lower obesity was correlated with lower COVID-19 case rates, I also subsetted for highest and lowest COVID-19 confirmed cases. If this hypothesis was supported, higher obesity countries should be represented in the higher COVID-19 case subset and vice versa for the lower COVID-19 case subset. However, as seen in Figure 3B, Samoa and Vietnam were both present among the lowest COVID rates, but were the highest and lowest obesity rate countries respectively. While this pattern contradicts the general trend established from Figure 1, it presents interesting conclusions when analyzing the distribution of energy intake by food group. In Figure 3, vegetables remained the most consumed food group across both high- and low-case countries, with low-case countries (ex. Laos: 47%, Tanzania: 47.9%) consuming a slightly higher percentage than high-case countries (ex. Montenegro: 33.5%, Luxembourg: 33.9%). Meanwhile, high COVID-19 case countries consumed greater percentages of meat products (ex. Luxembourg: 25.4%, Montenegro: 24.6%) and processed sugars (ex. Czechia: 6.2%, United States: 8%) compared to low-case countries (ex. Laos: 7.8% meat products, 1.5% processed sugars). Thus, both figures suggest that a diet higher in meat products and processed sugars correlates with greater health risks, including higher obesity prevalence and higher COVID-19 case rates. Conversely, a diet richer in cereals and vegetables is associated with better health outcomes.

To further explore how the distribution of energy intake influences not just COVID-19 susceptibility but also prognosis, Figures 5 and 6 explore the relationship between food group and COVID-19 death rate or recovery rate.

Here, the trends follow the patterns established in Figure 3. Both figures 5 and 6 reinforce that diets high in vegetables and cereals are associated with better COVID-19 outcomes, including lower death rates and higher recovery rates. Countries consuming more vegetables (ex. over 40% energy) showed consistently fewer COVID-19 deaths, with a narrower distribution of outcomes compared to other food groups (Figure 5C and 6C). Similarly, countries with higher cereal intake had better recovery rates (Figure 5D and 6D), though the wide spread of outcomes suggests other confounding factors may influence recovery. Conversely, diets rich in meat products and processed sugars are correlated with poorer outcomes, highlighting their potential role in exacerbating disease severity and impeding recovery. Countries with greater processed sugar intake (above 3% energy) also showed higher death rates, with significant variation in outcomes as intake increased (Figure 5E and 6E).

Overall Conclusions and Future Directions

In conclusion, these findings support most components of the initial hypothesis about relationship between food diversity and public health resilience to COVID-19, but they do not fully resolve it due to limitations of the dataset and methods taken. The figures implemented primarily focus on dietary composition rather than the diversity of the food supply. While a diverse food supply could be inferred from the variety of food groups (seafood, meats, vegetables, cereals, and processed sugars), the analysis does not directly assess or compare the variety of available foods within each country. However, the results do suggest that specific dietary patterns— greater intake of vegetables and cereals and lower consumption of processed sugars and meats—are associated with better health outcomes (ex, lower death rates and higher recovery rates from COVID-19). These findings align with the idea that countries with healthier, more balanced diets (which could reflect a more diverse food supply in some cases) exhibit lower COVID-19 mortality and better recovery outcomes. Thus, while food diversity isn't directly measured, the analysis nonetheless supports the broader assumption that diverse and balanced diets can contribute to better public health resilience.

Looking ahead, further research is necessary to clarify the mechanisms linking dietary patterns, obesity, and infectious disease outcomes. Longitudinal studies could investigate causality, while controlled interventions might assess the direct impact of dietary patterns on disease resilience. Furthermore, diet is just a snapshot of larger societal influences. Exploring the role of other factors such as physical activity, socioeconomic disparities, healthcare infrastructure, and genetic predispositions would provide a more thorough understanding of prognoses. Ultimately, synthesizing these results and expanding to other health crises could help generalize the findings and inform broader public health strategies.

VI. References

[1] https://nutrition.bmj.com/content/early/2023/10/04/bmjnph-2023-000688

[2] https://nutrition.bmj.com/content/4/1/257

[3] https://www.frontiersin.org/journals/nutrition/articles/10.3389/fnut.2023.1075061/full

[4] https://www.kaggle.com/datasets/mariaren/covid19-healthy-diet-dataset

[5] https://alandefreitas.github.io/matplotplusplus/

[6] CSCI 114 Class Slides

[7] https://www.sciencedirect.com/science/article/pii/S2211383523000126