Effect of Dietary Factors on COVID-19 Cases

Introduction

Suggested COVID-19 risk mitigation measures generally focus on vaccination, wearing masks, and isolation; however, we know that an individual's health can affect their susceptibility to infection and risk for severe illness or death. Since diet plays a large role in overall health, does it have any effect on COVID-19 infection and death rates?

In this project, we explore a dataset from Kaggle that incorporates data from February 2021 to compare different types of food intake, world population obesity, and global COVID-19 cases. We researched the following questions:

1. What is the relationship between COVID-19 infection and death rates and factors generally understood to be associated with health issues such as obesity and alcohol consumption?
2. Are there any strong correlations between different types of food and COVID-19 cases?
3. What are the dietary compositions of the countries with the highest and lowest COVID cases per capita?

Analysis

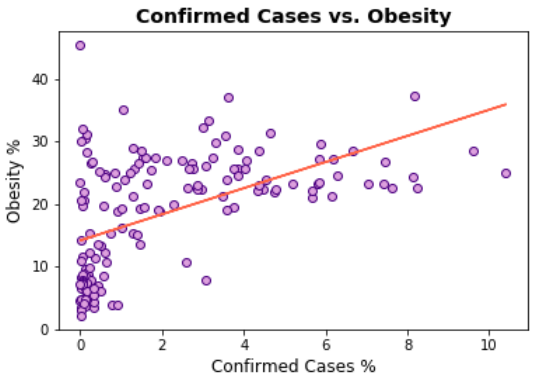
Data Cleaning

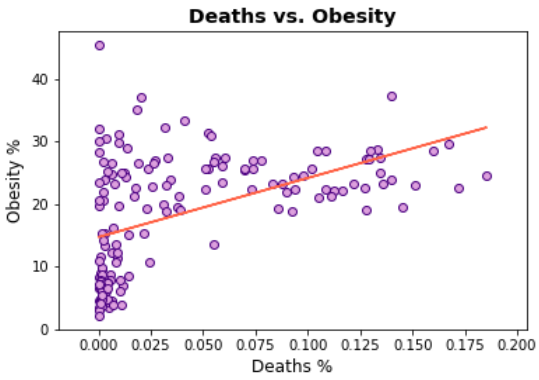
Before performing the following analysis, we removed countries with null values for confirmed cases and deaths, so this analysis is not inclusive of every country's data. We also combined some of the food categories with relatively low percentage composition across countries and expanded other food categories to be inclusive of these minimally consumed foods. This allowed us to consider diet composition on a broader scale, which was necessary for the dietary composition visualizations.

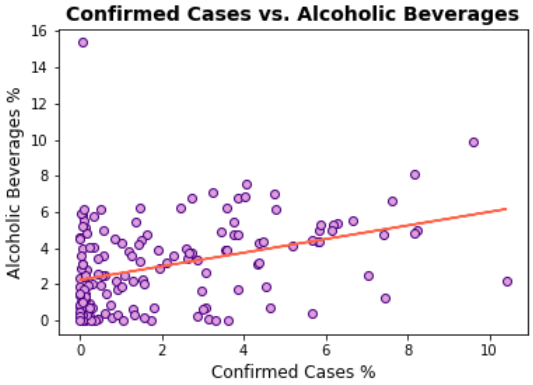
COVID-19 vs. Adverse Health Factors

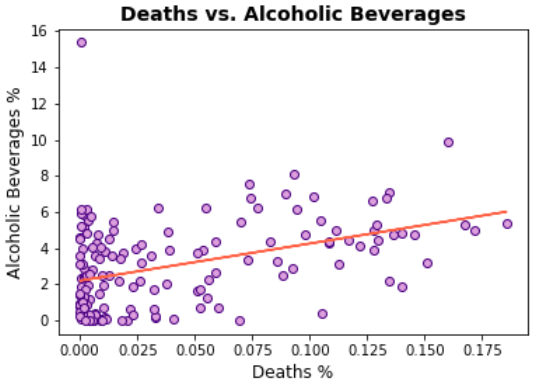
To find the relationship between COVID-19 cases and factors generally associated with health issues, we created scatter plots and performed linear regressions on obesity and alcohol consumptions vs. COVID-19 infection and death rates. As anticipated, there is a positive correlation between COVID-19 cases and each of these factors. Obesity is the most strongly correlated with COVID-19 cases, with a correlation coefficient (CC) of .53 for confirmed cases and .49 for confirmed deaths.

Additionally, there is a positive correlation between alcoholic beverages consumed and COVID-19 cases (CC of .37) and deaths (CC of .42).



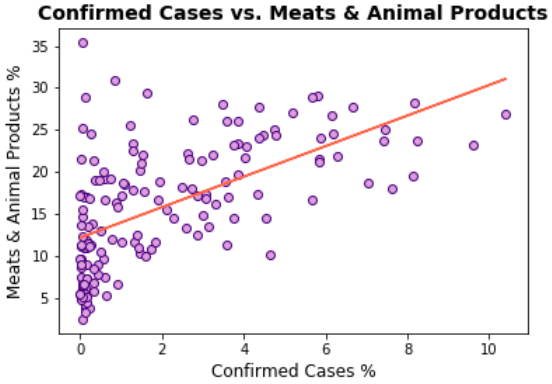


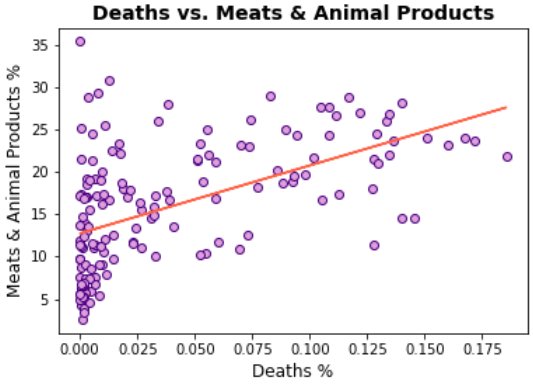


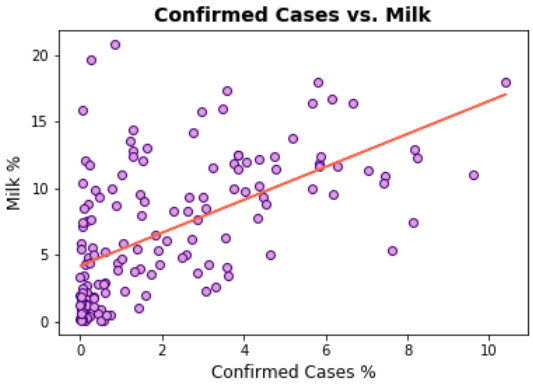


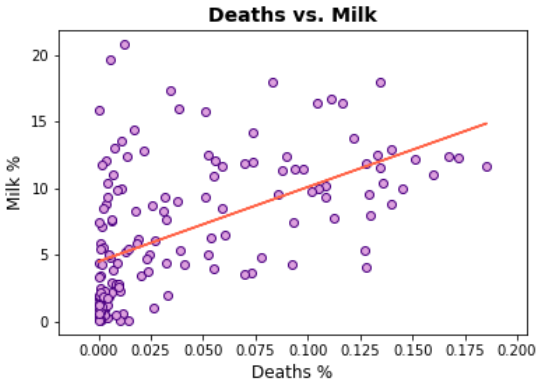
COVID vs. Diet

To find the relationship between COVID-19 cases and types of food, we created scatter plots and performed linear regressions on different categories of food vs. COVID-19 infection and death rates. The strongest correlation lies with countries that consume the most Meats and Animal Products with a CC of .59 for confirmed cases and .54 for deaths. Milk was the second most strongly correlated with COVID-19, with a coefficient of .54 for COVID-19 deaths and A .58 coefficient for positive cases.

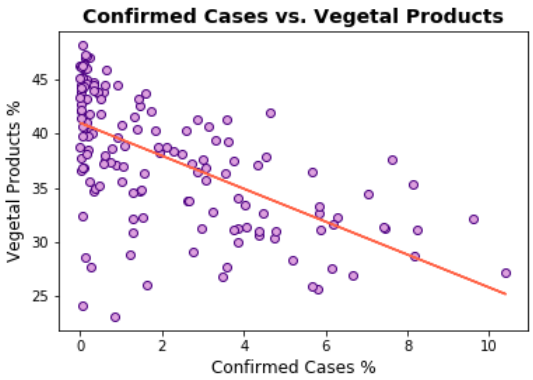


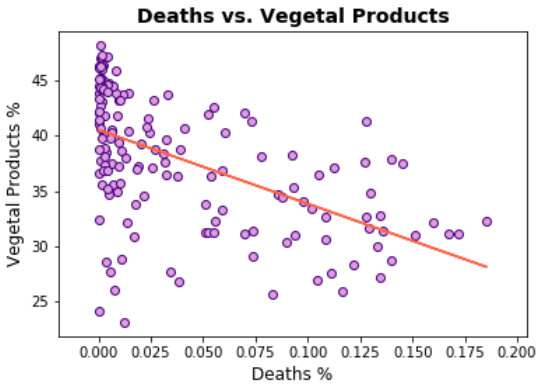


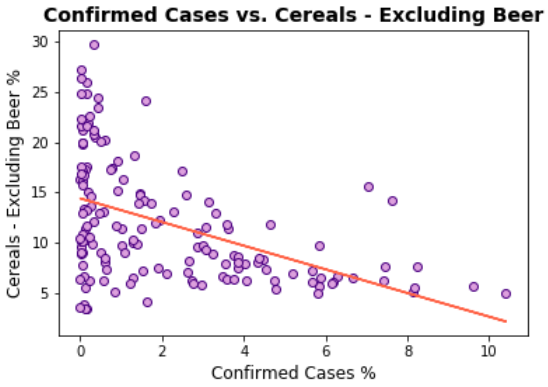


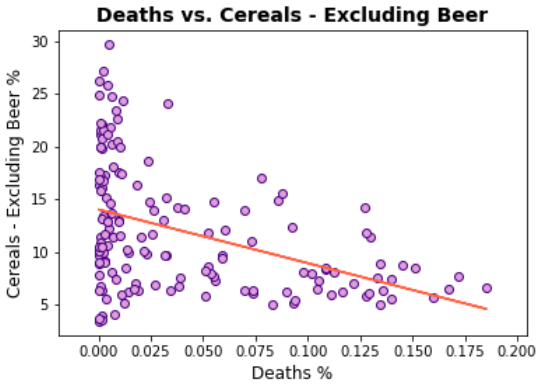


The strongest negative correlations were with vegetal products, with correlation coefficients of -.60 for confirmed cases and -.55 for deaths. The Cereals category was the second most negatively correlated, with a -.47 CC for COVID cases and -.42 for deaths.



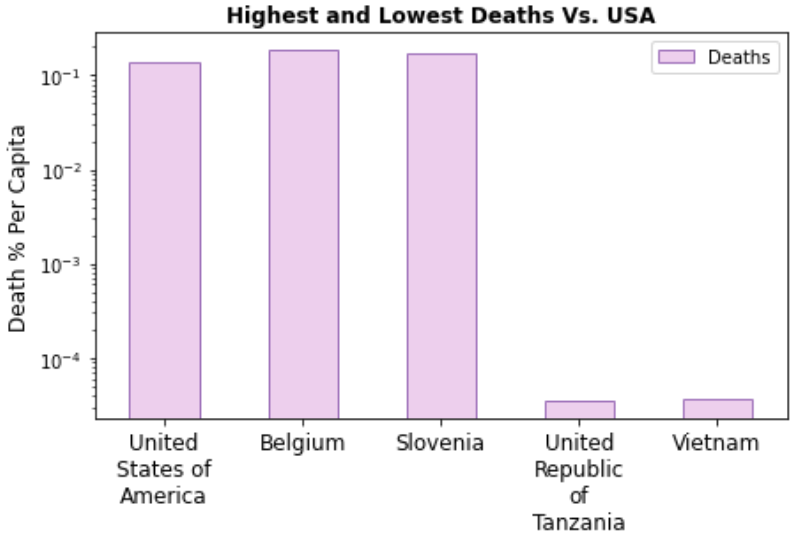




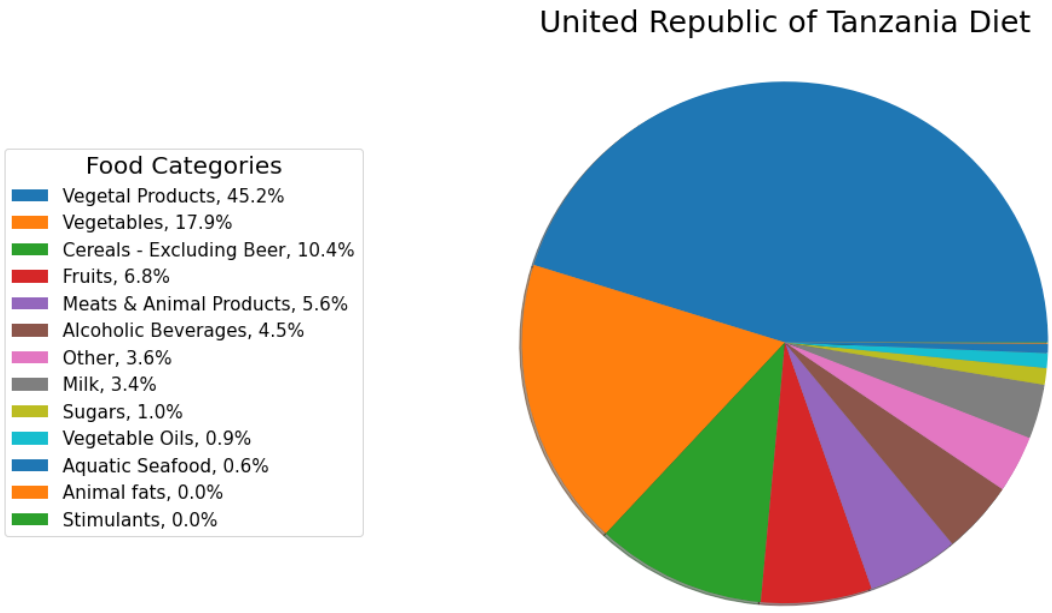


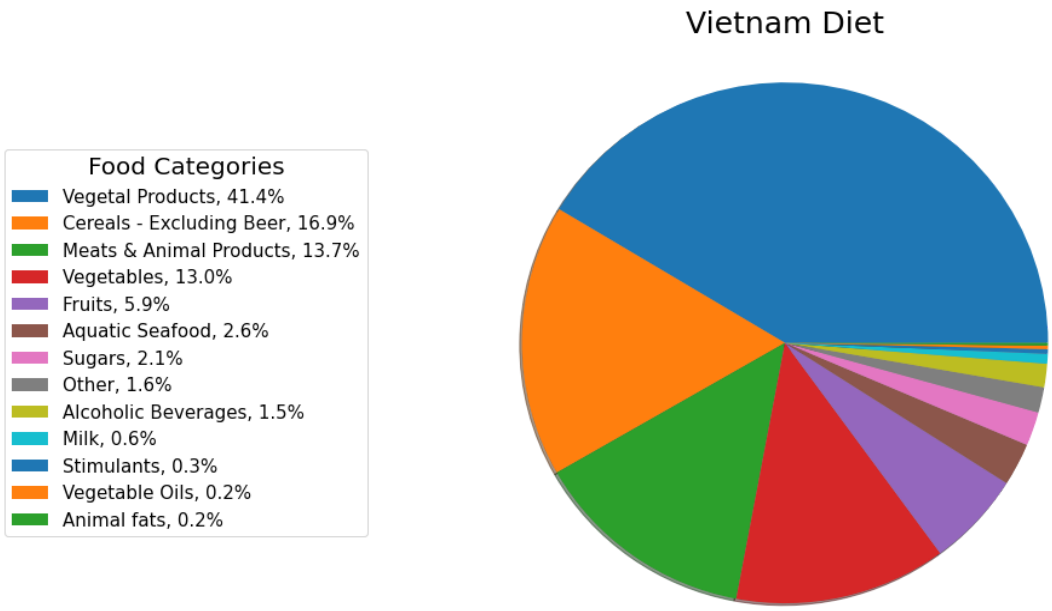
Dietary Composition of Significant Countries

COVID deaths rates across countries varied widely, with the lowest being .0000351 per capita and the highest being .185. For reference with a familiar place, the USA death rate per capita was .140.

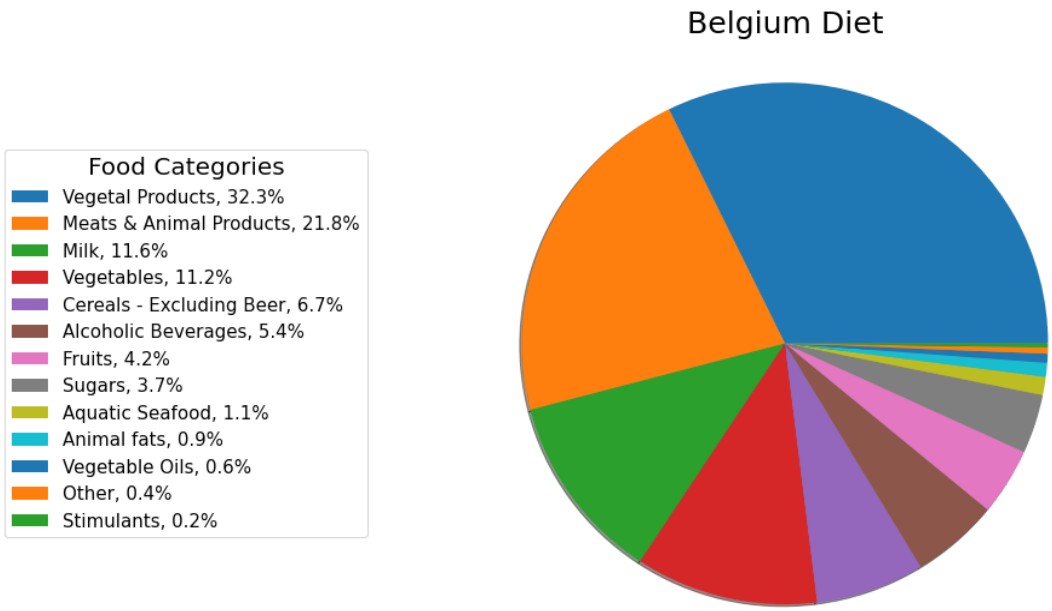


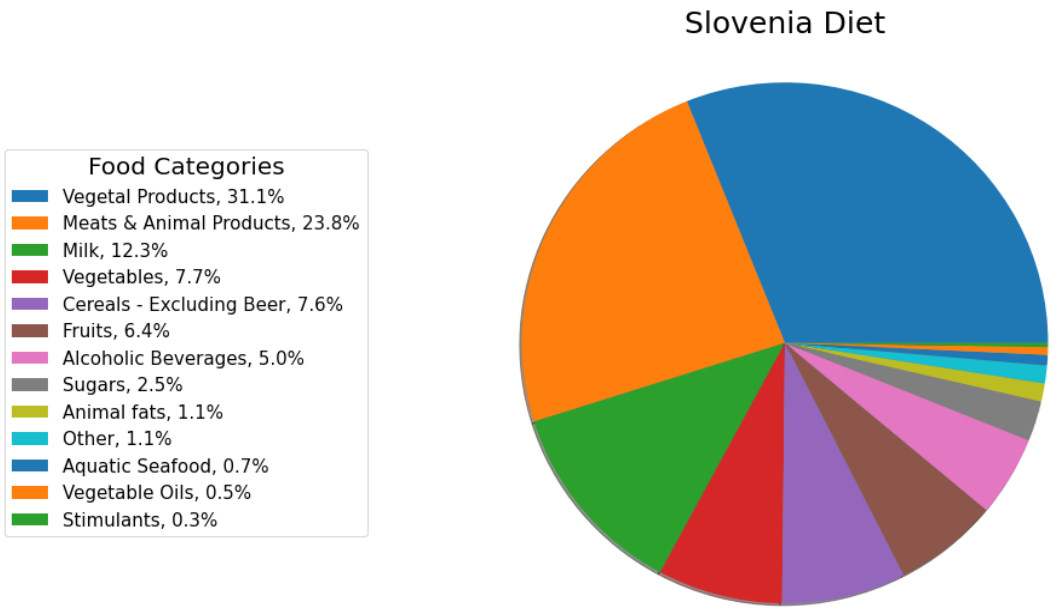
The countries with the two lowest death rates are The United Republic of Tanzania (URT) (.0000351) and Vietnam (.0000363). Both of these countries have diets high in vegetal products, the food category most negatively correlated with COVID deaths, and relatively low amounts of meat and animal products,, which are most positively correlated with COVID deaths.





The countries with the highest death rates are Belgium (.185) and Slovenia (.172). These countries have lower percentages of the foods negatively correlated with COVID than Vietnam and URT (Vegetal Products, Cereals). They also have higher percentages of foods positively correlated with COVID, including Meats/Animal Products and Milk.





Conclusion

While certain foods and health factors were more strongly correlated with COVID than others, none of these correlation coefficients were particularly high, since a high correlation coefficient is often considered to be .75 or even .90.

When considering this data, it's also important to look at the food category descriptions, as they require some scrutiny to determine how accurately this data was compiled. For example, "Vegetal Products" is extremely broad and contains over 30 different food items, many also existing in other categories such as various fruits, vegetables, and alcoholic products. Since the parameters of this category are unclear, the results of this analysis may be misleading.

All in all, this data provided some interesting insights into possible relationships between diet and COVID cases; however, additional analysis and/or studies are needed to determine if these relationships are a direct effect of diet, or if other characteristics of a country are a result of these trends. For example, it may be that countries with higher milk consumption have poorer health care or more lenient COVID precautions.