Exploring the relationship between dietary lifestyle and COVID-19 susceptibility

Itu Mukherjee

Abstract

This study is aimed at analyzing if a particular dietary lifestyle have any impact on COVID-19 occurrence. The dataset being used is downloaded from Kaggle and comprises of country wise consumption of different types of food along with number of confirmed cases and deaths of COVID-19. Correlation analysis revealed higher consumption of animal products to be directly associated with COVID-19 related mortality. The results also indicate a possible link between obesity and COVID-19 induced death.

Motivation

COVID-19 has turned into a global pandemic. As biomedical researchers worldwide are working around the clock to come up with a vaccine, there is also an interest in non pharmaceutical intervention strategies.

One of the major factors that can potentially play a role in determining susceptibility towards any decease is diet. With this in mind I wanted to explore links between global dietary lifestyles and COVID-19 susceptibility.

Dataset

I have used the following dataset from Kaggle.

https://www.kaggle.com/mariaren/covid19-healthy-diet-dataset

The dataset contains different categories of food supply quantity in kgs for 170 countries, calculated as the percentage of total population of a country. The dataset also has country-wise numbers of confirmed cases of COVID-19, deaths associated with COVID-19 and individuals recovered from COVID-19.

Data Preparation and Cleaning

The dataset has country-wise supply data for 23 different categories of food. In addition, the dataset also has country-wise statistics of obesity and undernourishment, confirmed cases of COVID-19, deaths associated with COVID-19 and individuals recovered from COVID-19, all showed as percentage of total population.

Quite a few columns have lot of NaN values. In particular, the undernourished column has more than one third missing values. I used pandas dropna() method to drop rows containing missing values. Since the undernourished column had a lot of missing values I decided to drop this column all together from the dataframe using pandas df.drop() method.

The final resulting dataframe had 152 countries as rows and 30 columns, out of which country label was an object while all the other parameters were floats.

Research Questions

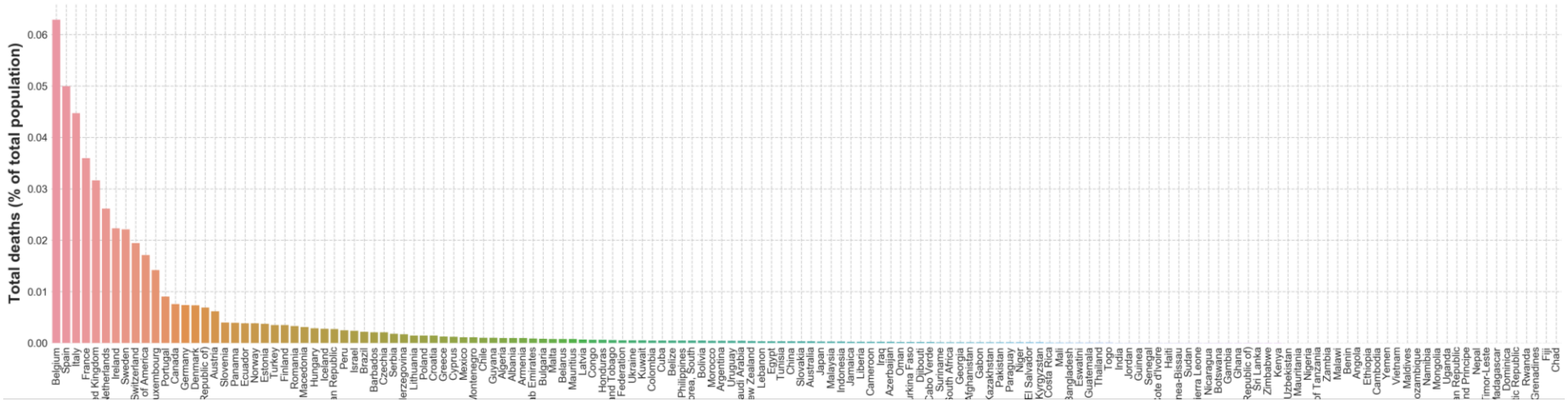
1. What are the country-wise distributions of dietary lifestyles?
2. Does dietary choice have any impact on susceptibility to COVID-19?

Methods

1) For plotting country-wise distributions of diet, I made use of a choropleth plot in which I overlaid country-wise dietary values on a map of the world.

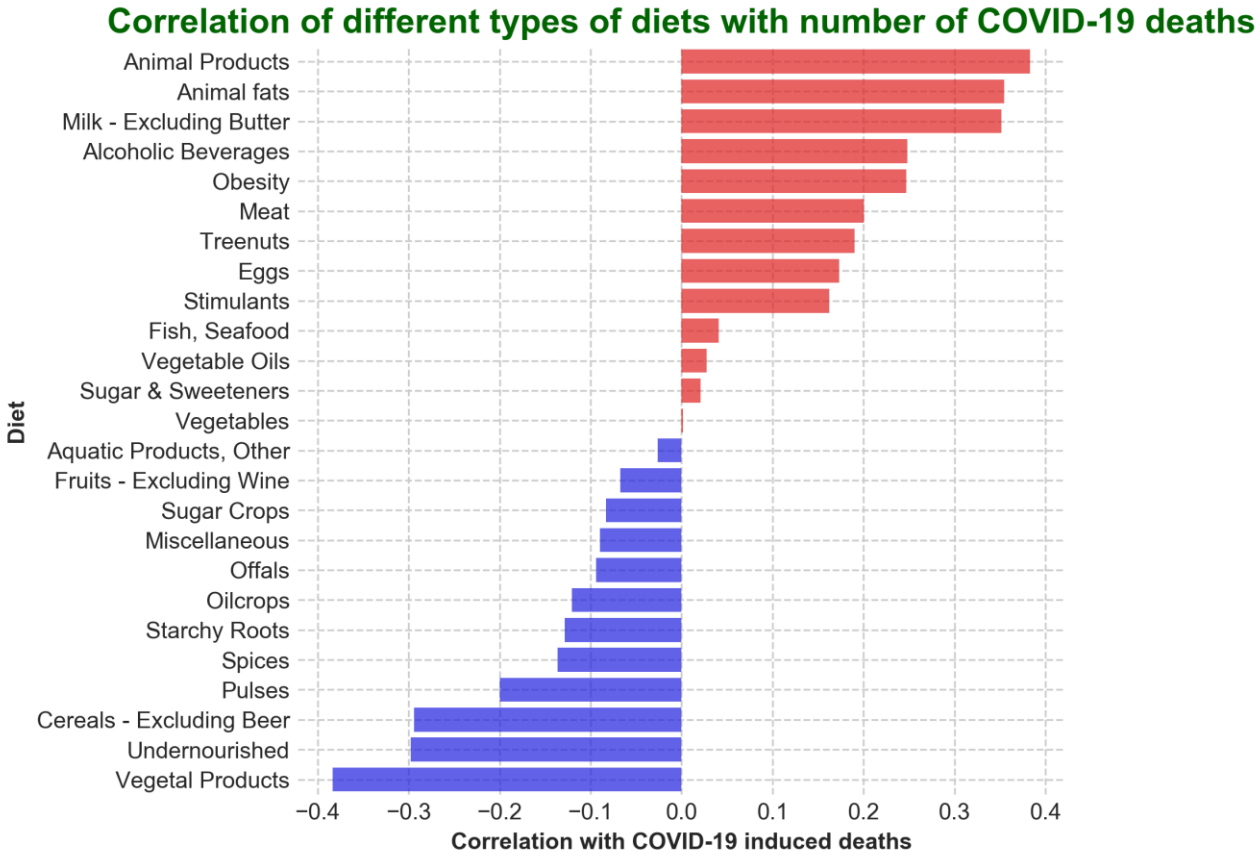
2) To explore links between different dietary choices and susceptibility to COVID-19, I calculated Pearson's correlation coefficients for the different dietary lifestyles with COVID-19 associated deaths across 152 countries.

Findings



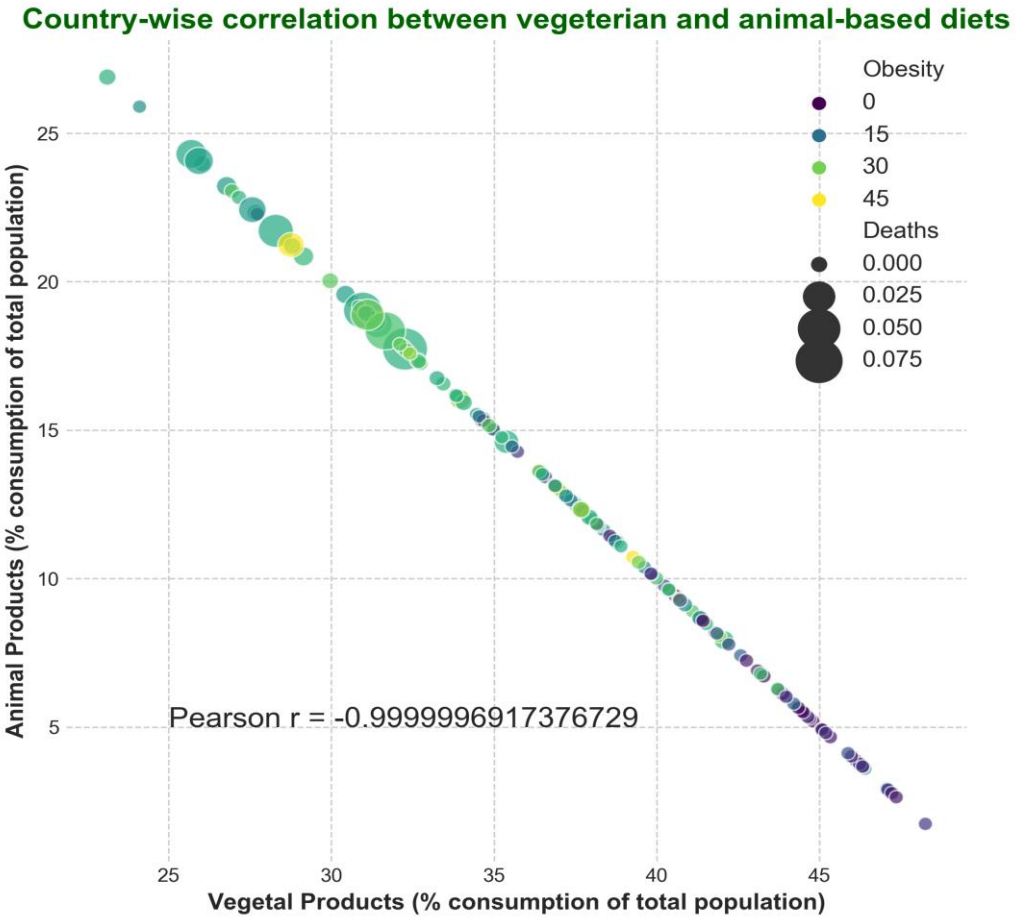
Percentage of total death rates per country

Findings



- The figure shows the correlation of different kinds of diet with COVID-19 associated deaths across 152 counties.
- High positive correlations (shown in red) indicate higher susceptibility to COVID-19 induced death associated with that attribute.
- On the other hand higher negative correlations (shown in blue) suggest a lower risk of COVID-19 induced death.

Findings

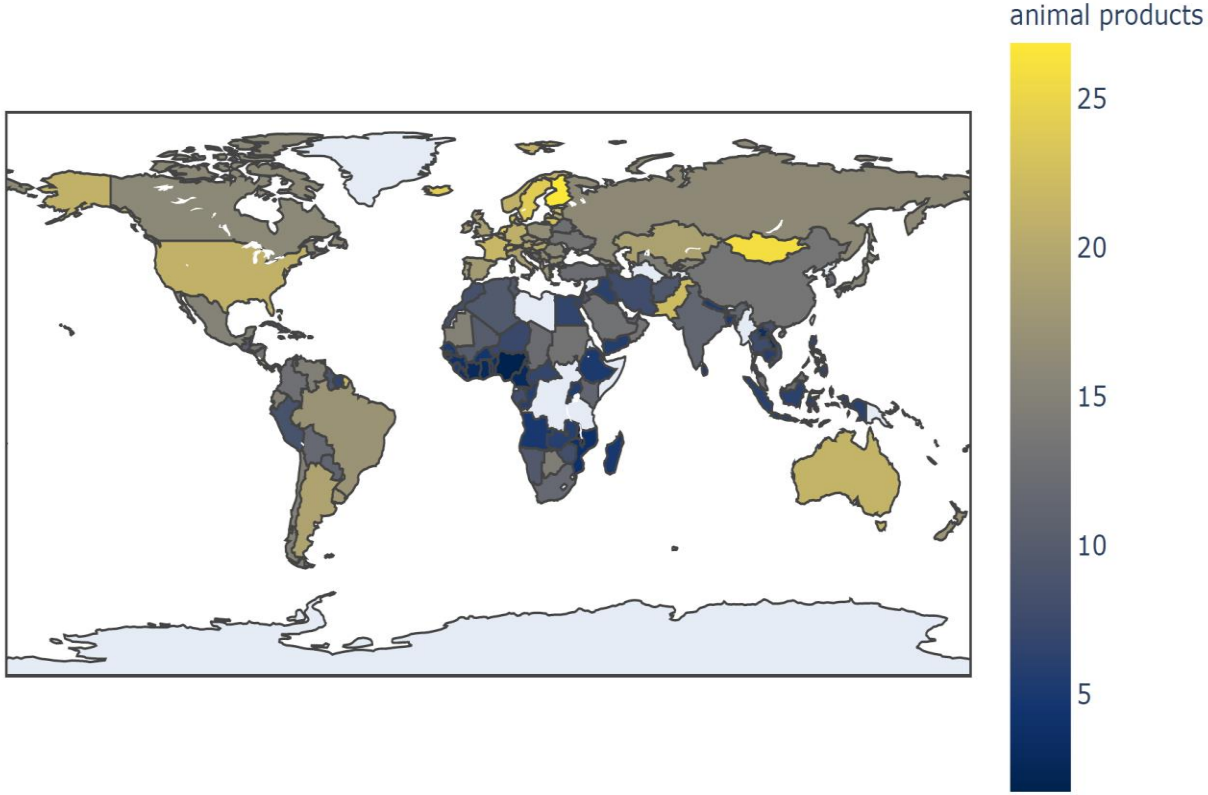


- The figure shows the country-wise correlation between vegetarian and animal-based diets of COVID-19 deaths across 152 counties
- We see a strong negative correlation between vegetal products and animal products across countries of the world.
- The plot also shows that countries with higher percentage of the population on animal products also have higher obesity.
- Therefore, these results indicate a direct association between obesity and COVID-19 induced death.

Findings

- The adjoining figure shows the distribution of animal product consumption in kgs as a percentage of the total population

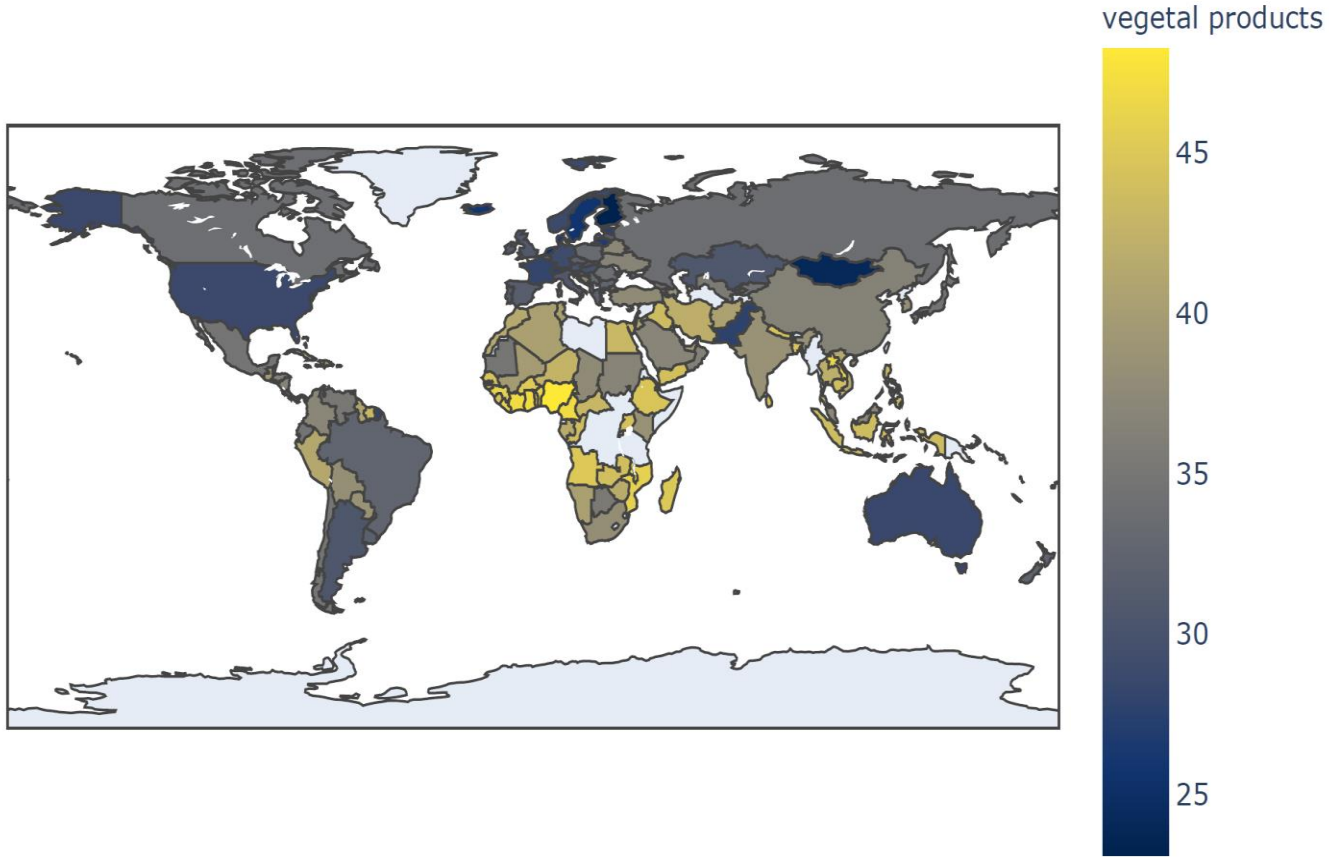
- From this figure we can see that in Western European Countries, Australia and USA the rate of animal product consumption is higher than that of in the African counties and other continents and has a direct correlation with the number of COVID-19 induced deaths



Findings

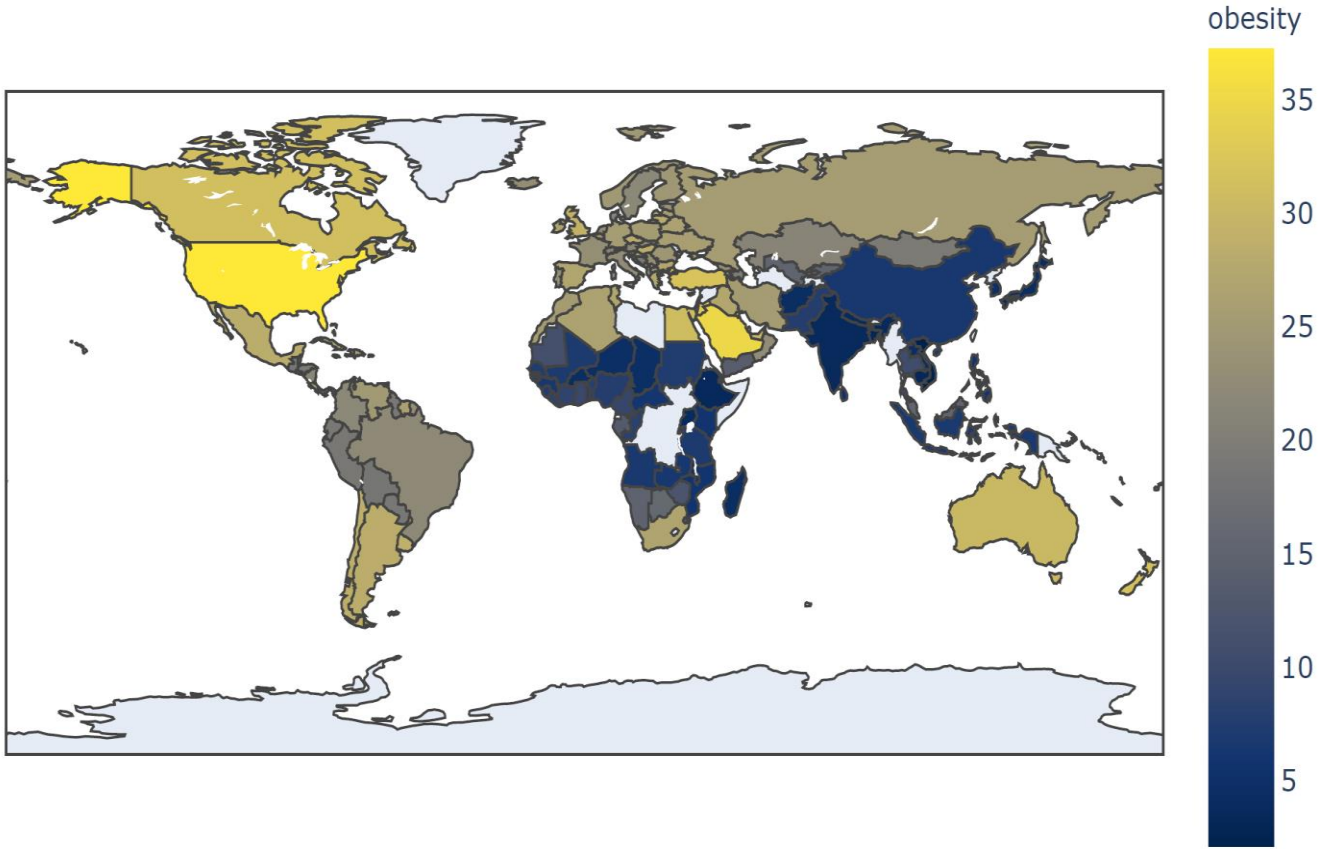
- The adjoining figure shows the distribution of vegetal product consumption in kgs as a percentage of the total population

- From this figure we can see that in African Countries and few other parts of Asia the rate of vegetal product consumption is higher than that of in the other countries and having comparatively less death rates.



Findings

- The adjoining figure shows the distribution of obesity as a percentage of the total population
- From this figure we can see that high income countries USA, Canada, Australia and few middle east countries are quite high in obesity which may adversely effect prognosis during COVID-19.



Limitations

- One potential limitation of the present study is that the presence of a lot of missing values led to the truncation of the original dataset which had 170 countries down to 152 countries. Although the loss of information of 18 countries out of a list of 170 does not seem to be a lot, but having the full dataset could possibly add to the strength of the associations observed in this study.

- Although consumption of higher amounts of animal products and obesity are positively associated with susceptibility to COVID-19, it is to be noted that one cannot infer causality from a purely correlational analysis.

Conclusions

- The study reveals global variations in dietary choices with higher income countries showing an increased affinity towards a diet based on animal products whereas most middle to low income countries follow a predominantly vegetal product based diet.
- Direct association of animal product consumption and indirect association of vegetal product consumption with COVID-19 related death indicates that higher consumption of animal products could make an individual susceptible to mortality due to COVID-19.
- The analysis also predicts obesity as a potential risk factor for COVID-19 related mortality

Acknowledgements

- Dataset : Kaggle
- I had insightful discussions with a friend of mine who is a biomedical researcher.

References

- Dataset : www.Kaggle.com
- www.stackoverflow.com
- www.plotly.com : for projections of data on the world map
- www.pandas.pydata.org