

## **Evaluation of Dietary Habits in Relation to Covid-19 Mortality Rate Using Machine Learning Techniques**

R. Sujatha<sup>1</sup>, Jyotir Moy Chatterjee<sup>2</sup>, S. S. Thorunitha<sup>1</sup>, Nadhiya S.<sup>1</sup>

<sup>1</sup>School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

<sup>2+</sup>Department of IT, Lord Buddha Education Foundation, Kathmandu, Nepal

jyotirchatterjee@gmail.com

**Abstract.** Coronavirus attacks have affected countless countries. The death rates between most countries are increasing day by day, and we have attempted to propose many considerations about the principal problems that cause dangerous infections across the globe. In this work, the dietary patterns of 170 countries are considered to identify correlations between diet practices and death rates, confirmed and recovered cases caused by COVID-19. We have used data from food intake by countries and data associated with the spread of COVID-19 and other health issues that help get new insights into the importance of nutrition and eating habits to combat the spreading of infectious diseases. We have built a machine learning model (regressor) such as ridge regressor, support vector regression, random forest, and XGBoost regressor to predict the mortality rate based on food intake information and Obesity. Two approaches were considered: one with all food-related features taken as parameters and a simpler one, which reduced the dimensionality by using only two features: animal products and vegetal products. Both have issues (mainly of spread and non-linearity), but we could use different models and metrics. Next, we have built a model to predict obesity rates based on eating habits in each country. The proposed model was far more effective, and the general inclination of the information was taken and anticipated. We have also used data visualization approaches to get better insights into the data considered.

**Keywords:** COVID-19, machine learning (ML), dietary patterns, mortality, ridge regressor (RR), support vector regression (SVR), random forest (RF), XGBoost regressor

## **1. Introduction**

In this work, it has been found that dietary habits are related to COVID-19 Mortality, so respect to the diet could be a method to prevent the COVID-19 death rate. Consequently, the point of this exploration is to distinguish Patterns that make it conceivable to put the focus of consideration on nations that today are in the current phases of the infection's extension but share dietary attributes with the nations that have experienced the most from this pandemic and may address a danger later on.

Our proposed framework utilizes a progressed data visualization tool to present the data in various structures to discover how a nation's eating routine relates to the COVID-19 death rate. Using ML algorithms, we can predict which food attribute prompts a country's passing rate. The dietary patterns of 170 nations were assessed to discover connections between the food habits and death rates brought about by COVID-19 utilizing ML Techniques. We built an ML model (regressor) such as RR, SVR, RF and XGBoost regressor. We evaluated the model on a few different scores like mean\_absolute\_error, mean\_squared\_error, and r2\_score both on training and testing data. Many factors are important to fight against the current COVID-19 epidemic. Maintaining good eating habits helps keep our immune system healthy and ready to combat a possible disease. In this project, we tried to explore possible patterns found in data of COVID-19 and food intake in different countries. One major goal was to find the influence of obesity rates on the disease's effect in each country. Splitting countries into HOC (High Obesity Countries) and LOC (Low Obesity Countries) groups, it was possible to create a classifier with good accuracy, predicting which group would be based on its food intake data. Having this, we created regression models to predict the Mortality of COVID-19 in countries based on their eating habits and obesity rate. Then next, we build a model to predict Obesity rates based on eating habits in each country. With different food cultures across the planet, it might be interesting to determine which food categories will best predict a country's rate of deaths. Still, there is no evidence that a country's diet impacts the spread of COVID-19.

## **2. Literature Review**

In synopsis, they need applied experiences into the pathophysiology of the unfavorable results of weight and arising proof with regards to the pathologic process in COVID-19 to recommend potential courses whereby Obesity can intensify the tissue harm related with the disease by the "severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)" infection. Thus, A developing group of proof demonstrates that weight is firmly and autonomously connected with unfavorable results of COVID-19, including demise (Lockhart et al., 2020). The daily routine is affected because of the COVID-19 situation. Grabia et al., (2020) aimed to investigate diabetes patients with Type 1 and 2. The survey was conducted to understand their periodic activities, vegetable consumption, and physical activity, and the result

showed that most people follow nutritious, healthy habits. Also, frequent use of hand sanitizer was increased. Pourhomayoun et al., (2021) collected 146 countries' laboratory-confirmed case reports. They considered more than 100 features based on their symptoms, existing conditions like diarrhea, illness, weakness, cardiac disease, etc., using ML algorithms like classification, regression, and neural networks to predict the risk of death. The model could help the clinical and pharmaceutical departments know which patients should prioritize based on their indications. Yadaw et al., (2020) broke down the patient-level Warehouse informational dataset for individuals with an affirmed finish of COVID-19 who had a wellbeing framework experience between 9th March and 6th April, 2020. The assessment was an audit that included patients with COVID-19 between 1st March and 18th April 2020. An examination of COVID-19 required a positive SARS-CoV test. In this study obesity was discovered to be a huge indicator of Mortality among inpatients with COVID-19 (Pettit et al., 2020). Klang et al., (2020) brilliantly analyzed data of patients with COVID-19 hospitalized to an enormous academic clinical center system in New York City. Data included economics, BMI, smoking, and alcohol consumption. Multivariable vital backslide models separated factors self-sufficiently associated with Mortality in patients younger and more settled than 50. Murray proposed evaluating the utilization of food and nutrients over 195 countries and evaluating the effect of their suboptimal intake, leading to Mortality. The method used by using a comparative assessment approach has examined the extent of illness explicit weight inferable from each dietary risk factor among grown-ups matured 25 years or more seasoned. Schwingshack et al., (2017) proposed a systematic survey of different kinds of food and hazards that cause Mortality. This investigation aims to meta-dissect the connection between intake of foods with hazards that cause Mortality. The utilization of low-risk food varieties brings about a 56% decrease of Mortality; consuming high-risk foods increases the risk of Mortality by 2 times. Lange (2020) intended to comprehend the dietary and way of life progressions that are significant determinants of well-being during COVID-19. The technique they have led a sectional examination through an online survey utilizing a test of 415 grown-ups. The outcome they have analyzed is that helpless dietary propensities with the undesirable way of life may prompt significant medical issues. Husain et al., (2020) proposed a dietary part for the avoidance or treatment of COVID 19. The examination recommends that less nutrients, bioactive in food or utilitarian sustenance may pass on the likelihood to amplify viral protection. Their objective is to decrease the number of diseases and relief of instances of COVID-19. Diet may accept a valuable part in keeping a sound body weight and forestalling non-transmittable conditions. Hassen et al., (2020) investigates the effect of COVID-19 customer mindfulness, mentalities, and practices related with food use. The results show clear changes in how shoppers eat, shop, and cooperate with food. In this paper the authors have revealed that passing's in Germany are decently low by contrasting with numerous European countries due to coronavirus.

After a few clarifications proposed, early and colossal testing of the population and reasoned that it is fundamental to get diet and changing over protein level 2 in population with different COVID-19 passing rates since dietary medications may be of incredible advantage (Bousquet et al., 2020). Woolf et al., (2020) thought of and evaluated the Deaths in the long early stretches of the pandemic. These evaluations propose that the quantity of COVID-19 passing's detailed in the principal long stretches of the pandemic caught just 66% of overabundance passing's in the US. In outline, Iddir et al., (2020) covers enhancements, micronutrients, and other nutrients known to impact insusceptibility and disease risk, particularly appropriate during the COVID-19 emergency. The proof showed that an eating schedule that insistently impacts safe limit comprises acceptable proportions of protein, particularly glutamine & amino acids, high omega-3 vs lower drenched, trans fat & unsaturated fats, high fiber substance like full jots & micronutrients including supplement A, supplement E, supplement D and C, supplement B. Faour-Klingbeil et al., (2021) showed that the respondents in in some countries had limited data on SARS-CoV-2 and were stressed over getting the COVID-19 disorder from food. The results supplemented the significance for neighborhood specialists to improve the quality and level of nuances of the data and control the initiating of the viral pieces of tattle by remembering all partners for the sanitation and general wellbeing areas. Bohlouli et al., (2021) proposed a review on the relation between COVID-19 and fast-food consumption. This study concludes that intake of fast foods can decrease the availability of some necessary nutrients, increasing the vulnerability to COVID-19. Hossen et al., (2020) used various ML algorithms and data mining techniques to predict the recovery rate of the COVID-19 recovered patients in south Asian countries based on their food habits and concluded that adapting a healthy eating lifestyle can help fight the fight COVID-19. The author presented a study on the relation between Obesity and covid-19 mortality rate. Popkin et al., (2020) states that the impact of COVID-19 is significantly higher in individuals with Obesity. Rishi et al., (2020) states that the use of home cooked eating is normal, in this way improving the beneficial microflora, which may have achieved better expectation of COVID-19 patients in India interestingly in the western countries. Gasmi et al., (2021) ensure that deciding a person's present metabolic status, including full scale and micronutrients, is a fundamental factor in characterizing any people's lacks, which should be tended to desperately through a legitimate eating routine, explicit customized dietary supplementation, and way of life changes. Villadsen et al., (2021) proposed that emotional well-being is undeniably significant in well-being conduct imbalance in COVID time. Hence, the advancement of emotional well-being may be a significant segment of improving post-COVID-19 population well-being by introducing a theoretical model for understanding the quick COVID-19 related expansion in the food weakness and its effect on well-being (Leddy et al., 2020). The effect of the pandemic on food security led to amassed hazards for ongoing infection advancement,

grimness, and Mortality. (Zupo et al., 2020), presented a review article investigation the starter impacts in dietary follows throughout the COVID-19 epidemic and a public health source of inspiration to confront Obesity.

### 3. Materials and Methods

#### 3.1. Dataset description

Dataset consists of a percentage of food intake (kg)-Alcoholic Consumption, animal products, aquatic products, cereals, Animal fats, fish, sea food, fruits, meat & eggs in 170 countries worldwide. Information about Obesity, undernourished & COVID-19 cases as percentages are also included in this dataset (<http://www.fao.org/home/en/>). Data for population count for each country is taken from (<https://www.prb.org/>). Data for COVID-19 confirmed, deaths, recovered, and active cases are collected (<https://coronavirus.jhu.edu/map.html>). Some of the few attributes that are considered are presented in Table 1.

Table 1: Attribute details

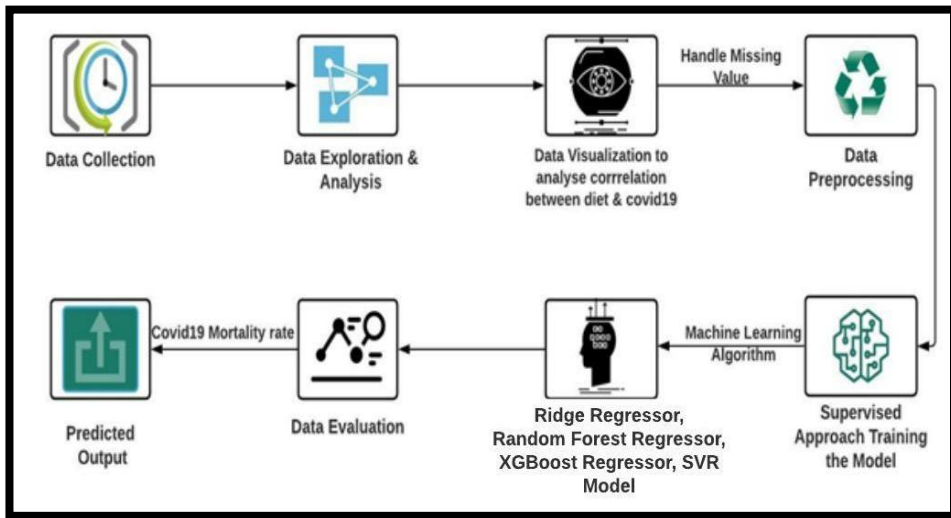
Sl. No.	Variables	Mean	Standard Deviation
1	Alcoholic Beverages	3.022971	2.382243
2	Animal fats	0.221064	0.278304
3	Animal Products	12.181871	5.852635
4	Aquatic Products, Other	0.013994	0.129382
5	Cereals - Excluding Beer	11.800347	5.824870
6	Eggs	0.470570	0.331209
7	Fish, Seafood	1.387195	1.257382
8	Fruits - Excluding Wine	5.621405	3.152849
9	Meat	3.375934	1.762911
10	Milk - Excluding Butter	6.519776	5.020379
11	Treenuts	0.117474	0.146143
12	Vegetable Oils	0.851554	0.445335
13	Vegetables	6.085912	3.560148
14	Vegetal Products	37.814834	5.852943
15	Obesity	18.707784	9.633557

This analysis aims to find out how a country's diet correlates with its COVID-19 mortality rate. With different food cultures across the world, it would be interesting to see the food categories that can best predict a country's rate of deaths.

The Fig. 1. Can be explained as: The first step carried out was Data Exploration and analysis of four different datasets are food quantity (in Kg, Kcal), Fat, and Protein. Instead of considering all the 24 attributes, if the two or more attributes are highly correlated then they are removed and a few of the attributes considered for further evaluation. The next step is data preprocessing is most important. The proposed

framework incorporates a group of approaches for preprocessing the data to get new features & dealing with missing qualities, disposing of repetitive & pointless information components & choosing the top instructive features. The data visualization shows the countries with the worst impact from COVID-19 when there is a high obesity rate and high consumption of Animal Products compared with vegetal products.

Fig. 1: The Architecture diagram explains the overview of overall process take place in predicting COVID-19 mortality



### 3.2. System working

The supervised learning approach used is RR, SVR, RF and XGBoost regressor. Mortality is calculated as death divided by confirmed case. Then we evaluated the model on a few different scores like mean\_absolute\_error, mean\_squared\_error & r2\_score. After metric evaluation, building the best performing model training and testing. We could show the use of different models and metrics. The final output is making predictions and visualization the plot that shows actual data and predicted data to predict the Mortality of COVID-19 and predict Obesity in countries based on their eating habits.

### 3.3. Data visualization

The data visualization showed that obese people strongly correlate with the COVID-19 mortality rate. The data visualization showed that obese people strongly correlate with the COVID-19 mortality rate. We first try to visualize the data to understand the correlation between various attributes on the collection of datasets, as in Fig. 2. First, the missing values are analyzed and filled with default parameters in the data. Then we will be using various supervised learning algorithms, and the results will be

compared. Data scientists utilize data exploration & analysis to break down & explore data sets & sum up their fundamental attributes, regularly utilizing information visualization techniques. It can likewise help decide whether the statistical procedures for data investigation are proper. A heatmap is a two-dimensional graphical portrayal of information where the individual qualities contained in a grid are addressed as colors. Heatmap comments are significant segments of a heatmap that show extra data that partners with rows and columns in the heatmap. The attributes that have a roughly normal distribution are Animal Products, Obesity and Vegetal Products. Alcoholic Beverages, Animal Fats, Milk - Excluding Butter, on the other hand, present a right skewed distribution.

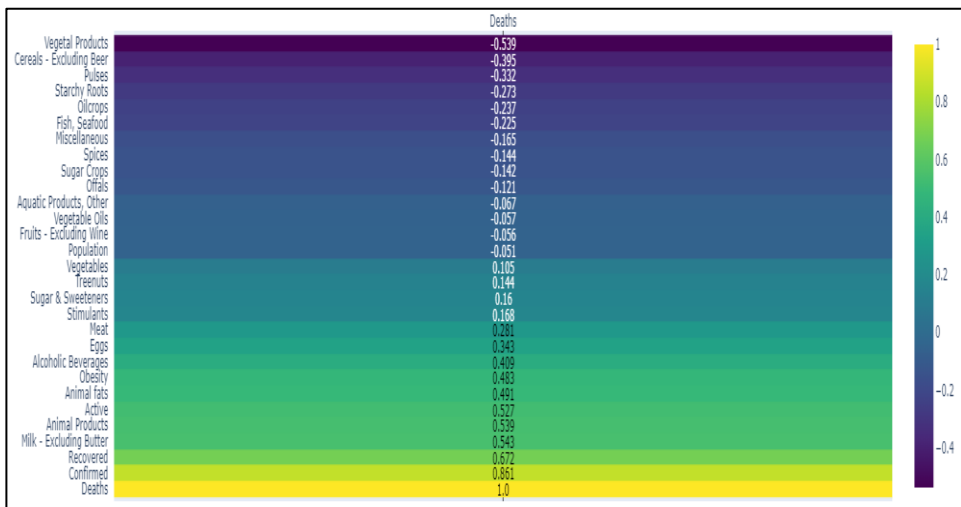


Fig. 2: Data visualization

Here we used two features: - 1) z\_score\_scaled 2) log\_scaledfeature\_names

- z\_score\_scaled\_feature\_names = ['Animal Products', 'Obesity', 'Vegetal Products']
- log\_scaled\_feature\_names = ['Animal fats', 'Milk - Excluding Butter']

### 3.3.1. Obesity between countries: High or low obesity?

There is a relation between high consumption of animal products (compared with Vegetal Products) and high obesity rates as per Fig. 3. Since the pandemic started, many examinations have detailed that a large number of the most broken down COVID-19 patients have been individuals with Obesity. As of late, that connection has come into more keen concentration as enormous new populace considers have

solidified the affiliation & shown that even individuals who are just overweight are at higher risk as per Fig 4.

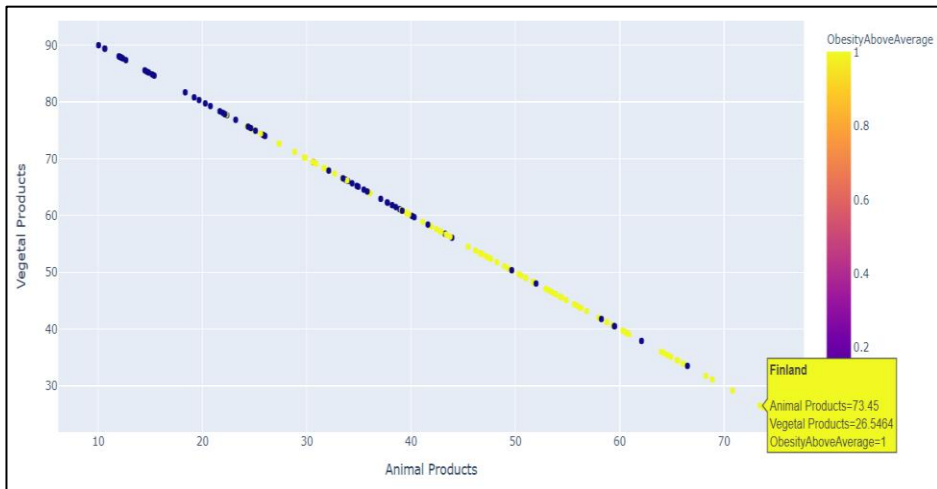


Fig. 3: Highest animal product intakes (Finland)

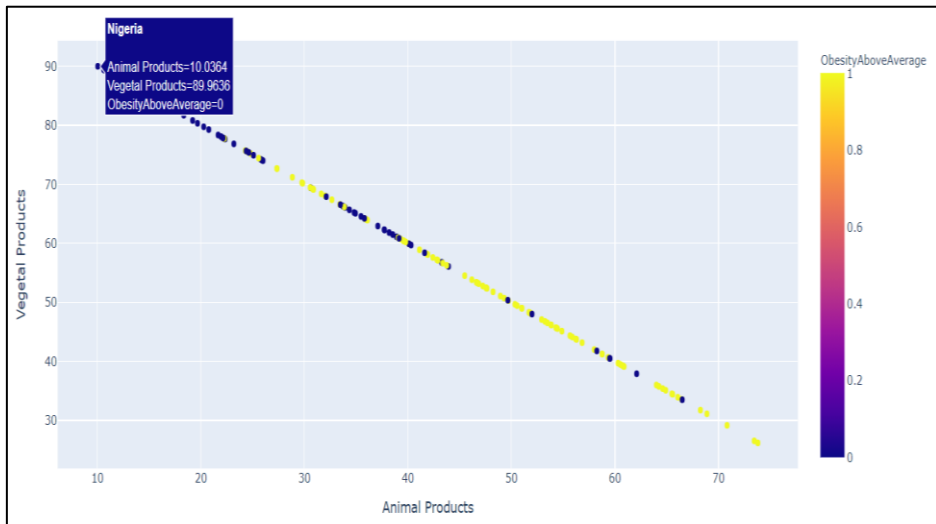


Fig. 4: Highest vegetal product intakes (Nigeria)

In Fig 5., we can see clearly that the "high obesity rate" countries have the worst impact from COVID-19. We have created a column ObesityAboveAvg that has a value 1 if the country has obesity rate above the mean of all other countries and 0 otherwise. The countries with obesity rates above the mean of all countries have a higher consumption of Animal Products and lower consumption of Vegetal Products.



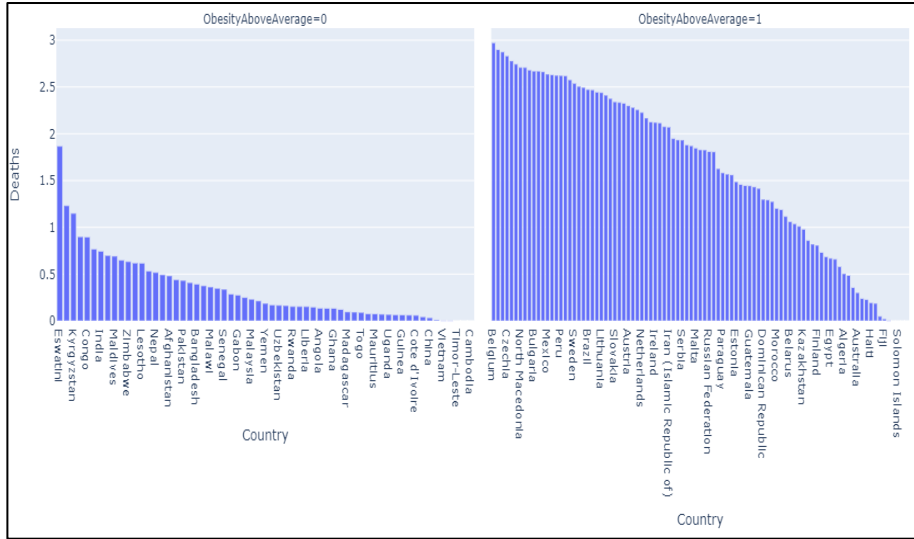


Fig. 5: HOC (Highest Obesity Countries) and LOC (Low Obesity Countries)

### 3.3.2. Health diet vs COVID-19

Obesity includes a stronger correlation with covid deaths than recovery, and malnourished patients include a stronger correlation with covid recovery than deaths as per Fig. 6. This might mean that, on average, obese patients are highly likely to die from covid whereas malnourished patients can survive. This is often why Obesity worsens outcomes from covid. Such results are to be understood as several alternative factors are to be taken into consideration - for instance, malnourished patients are likely to be in emerging countries, wherever the population is extremely young and possibly to survive.

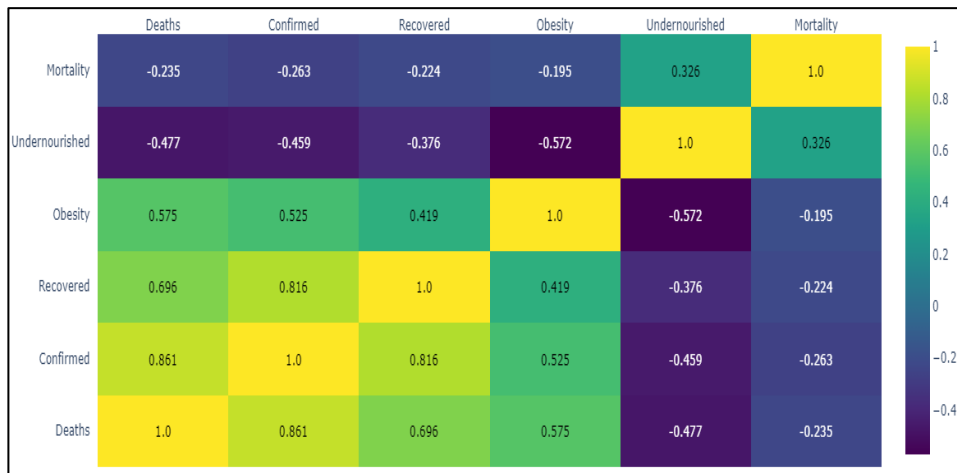


Fig. 6: Obesity correlation with Covid19 deaths

Obesity has a stronger correlation with covid deaths than recovery and undernourished patients have a stronger correlation with covid recovery more than deaths. This might mean that on average, obesity patients are more likely to die from covid whereas undernourished are to survive. This is why Obesity worsens outcomes from covid. Such results are to be taken rigorously as several different factors are to be taken under consideration - as an example, undernourished patients are to be in emerging countries, wherever the population is incredibly young and most likely to survive.

### 3.4. Proposed approach

Here we are using 4 approaches: RR, SVR, RF & XGBoost approach to check the correlations between diet practices and death rates, confirmed and recovered cases by COVID-19. We distributed the data into a training set and testing set with an 80:20 split. 80 % of the dataset is used to model training, and 20 % of the dataset is utilized for testing the model. The model building includes predicting Mortality based on all the food features and Obesity considered for predicting Mortality. A simpler model for predicting Mortality further the dimensionality reduced to two features Animal and vegetal, predicting Obesity using highest correlation food feature.

#### 3.4.1. Performance evaluators

The mean squared error (MSE), mean absolute error (MAE) &  $R^2$  to evaluate the model. MSE is an essential measure used to determine an estimator's performance. MSE is calculated as the mean of squares of the errors.

$$MSE = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z}_i)^2 \quad (1)$$

where n is the no of datapoint,  $z_i$  is the observed value and  $\bar{z}_i$  is the predicted value.

MAE is calculated as the average of all absolute errors. The formula for calculating MAE is:

$$MAE = \frac{1}{n} \sum_{i=1}^n |z_i - a_i| \quad (2)$$

Where n is the total no of datapoint,  $z_i$  is the predicted value &  $a_i$  is the true value.  $R^2$  score indicates the percentage of variance of error.  $R^2$  score is calculated as:

$$R^2 = 1 - \frac{RSS}{TSS} \quad (3)$$

where RSS is sum of sq. of residual & TSS is total sum of squares.

#### 3.4.2. RR

We will build our regressor with a standardization step in our pipeline (scale the data). Besides RR, three other models namely SVR, RF and XGBoost Regressor. RR shrivels the coefficients & it assists with diminishing the model intricacy & multi-collinearity. The obtained results for the RR are presented in Table 2.

Table 2: Result obtained for RR

Sl. No.	Parameter	Test	Train
1	MSE	0.0003	0.0002
2	MAE	0.013	0.011
3	R <sup>2</sup>	-0.0196	0.197

### 3.4.3. SVR

SVR can in like manner, be used as a regression technique, keeping up every last one of the principles includes that depict the computation. The SVR uses comparable standards as the SVM for grouping, several minor differentiations. Regardless of anything else, because yield is a genuine number, it ends up being difficult to anticipate the current information, which has endless potential outcomes. Because of relapse, an edge of resilience (epsilon) is set in assessment to the SVM, which would have adequately referenced from the issue. However, there is likewise a more complicated explanation other than this reality. The algorithm is more complicated accordingly to be taken in thought. Nonetheless, the principal thought is consistently something similar: to limit error, individualizing the hyperplane that expands the margin, and remembering that the error is endured. The obtained results for the SVR are presented in Table 3.

Table 3: Result obtained for SVR

Sl. No.	Parameter	Test	Train
1	MSE	0.00079	0.00068
2	MAE	0.024	0.024
3	R <sup>2</sup>	-1.262	-1.3855

### 3.4.4. RF

Tree model don't require scaling the data as we have done for RR. It's a meta-assessor that fits various characterizing decision trees on different sub-examples of the dataset & used averaging to improvise the predictive precision & control over-fitting. The sub-example size is controlled with the max\_samples boundary. The obtained results for the RF are presented in Table 4.

Table 4: Result obtained for RF

Sl. No.	Parameter	Test	Train
1	MSE	0.00038	4.3245
2	MAE	0.0142	0.0049
3	R <sup>2</sup>	-0.1154	0.850

### 3.4.5. XGBoost

XGBoost (Extreme Gradient Boosting) has a group of boosting algorithms and utilizes the gradient boosting (GBM) system at its center. It is an optimized distributed gradient boosting library. XGBoost model doesn't require scaling the data. The obtained results for the XGBoost Regressor are presented in Table 5.

Table 5: Result obtained for XGBoost

Sl. No.	Parameter	Test	Train
1	MSE	0.00039	3.045
2	MAE	0.0145	0.0041
3	R <sup>2</sup>	-0.1408	0.9989

## 4. Results nad Discussions

### 4.1. Predicting mortality

We built a model (regressor) to predict the mortality rate based on food intake information and Obesity. With the correlation matrix we can sort the values of the Mortality. So, the highest correlation obtained with Mortality is Milk - Excluding Butter, and this highest intake from Highest Obesity Country and Low Obesity Country. The first thing will predict the Mortality that is calculated as Deaths by Confirmed using Obesity and food features Animal products and Vegetal Products such as Animal fats, Aquatic Products, Eggs, Fish, Seafood, Meat, Milk - Excluding Butter, Offal, Alcoholic Beverages, Cereals - Excluding Beer, Fruits - Excluding Wine, Pulses, Spices, Starchy Roots, Stimulants, Sugar & Sweeteners, Sugar Crops, Treenuts, Vegetable Oils, Vegetables. The number of confirmed cases should be the sum of deaths, recovered and active. To investigate the impact of deaths by COVID-19, we will create a column Mortality which will be calculated as Deaths by Confirmed. Table 6 presents the detailed model comparison.

Table 6: Model comparison for mortality

Approach	Parameter	Test	Train
----------	-----------	------	-------

<b>RR</b>	<b>MSE</b>	<b>0.0003</b>	<b>0.0002</b>
	<b>MAE</b>	<b>0.013</b>	<b>0.011</b>
	<b>R<sup>2</sup></b>	<b>-0.0196</b>	<b>0.197</b>
<b>SVR</b>	MSE	0.0008	0.0007
	MAE	0.024	0.024
	R <sup>2</sup>	-1.262	-1.3855
<b>RF</b>	MSE	0.0004	4.3245
	MAE	0.0142	0.0049
	R <sup>2</sup>	-0.1154	0.85
<b>XGBoost</b>	MSE	0.0004	3.045
	MAE	0.0145	0.0041
	R <sup>2</sup>	-0.1408	0.9989

The best model among all Ridge Regressor. Training MSE and Testing MSE are almost similar. The model neither underfit nor overfit. In case of RF, SVR, XGBoost the MSE was high for training compared to RR. The Fig. 7-10 below shows the mortality dependency on food features and Obesity that has the most positive correlation with the target.

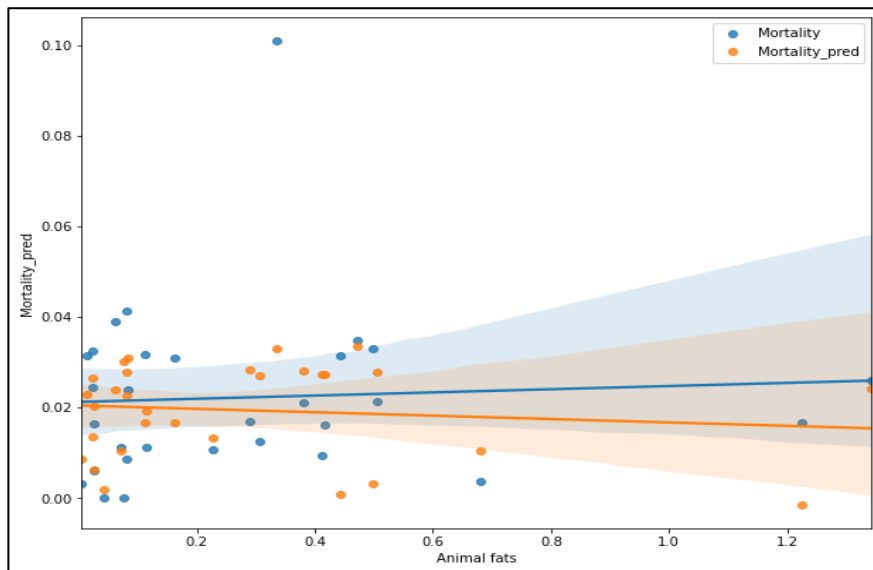


Fig. 7: Animal fat (mortality vs predicted mortality)

The above Fig. 7 shows simple visualization of our model's predictions using the food feature (Animal fats).

The next Fig. 8 shows simple visualization of our model's predictions using the food feature (Meat).

And the next Fig. 9 shows simple visualization of our model's predictions using Obesity.

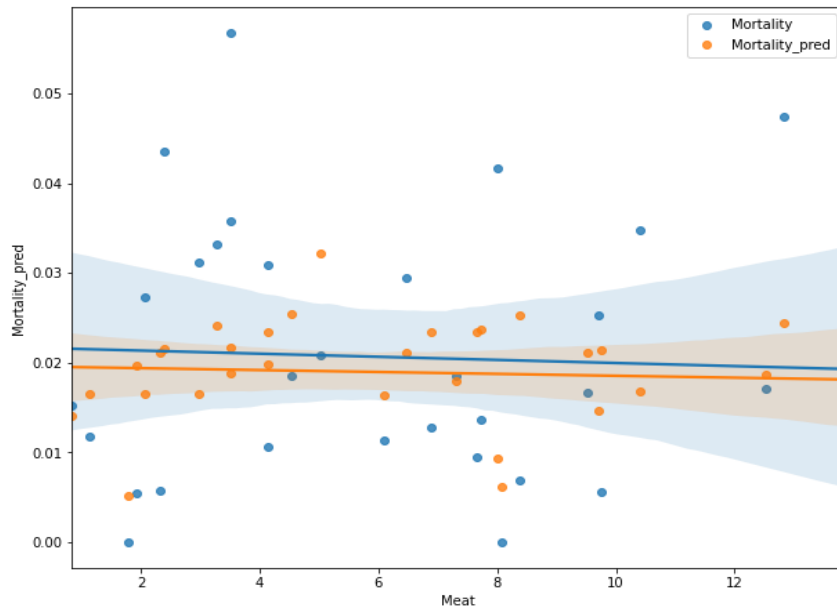


Fig. 8: Meat (mortality vs predicted mortality)

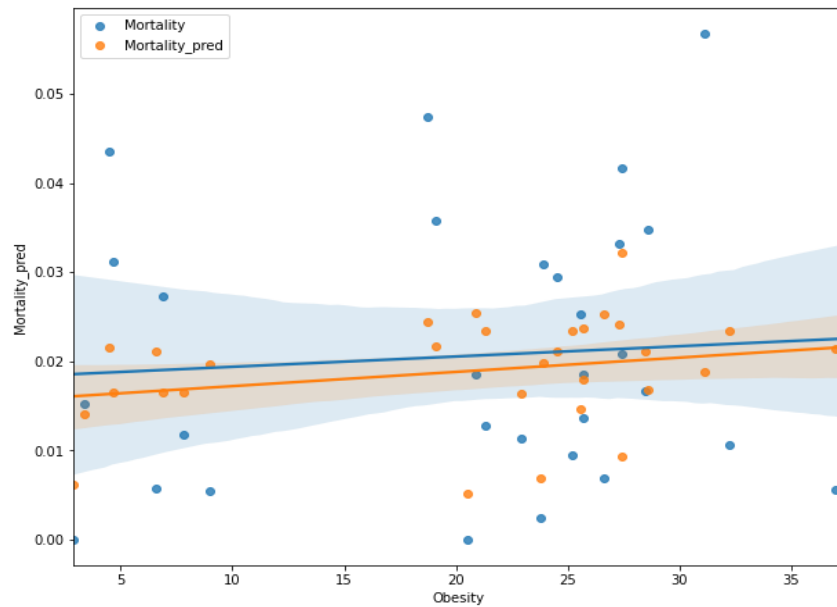


Fig. 9: Obesity (mortality vs predicted mortality)

The next Fig. 10 shows simple visualization of our model's predictions using food feature (Cereals-Excluding Beer). The model makes a prediction that somehow captures the tendency of our target.

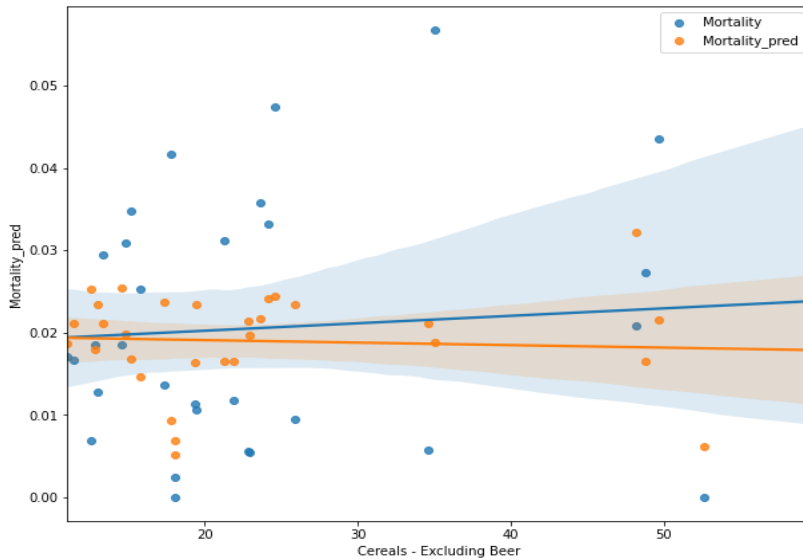


Fig. 10: Cereals-excluding beer (mortality vs predicted mortality)

#### 4.2. Predicting obesity

As Obesity has a higher correlation with all "food features" let's build our models to predict the actual obesity rate. We expect these models to have better metrics than those built to predict Mortality. Comparing various metrics for each model is presented in Table 7. First, the Obesity dependency on Cereals-Excluding Beer has the most negative correlation with the target. Second, the Obesity dependency on Meat has the most positive correlation with the target.

Table 7: Model comparison for obesity

Approach	Parameter	Test	Train
RR	MSE	0.817	18.01
	MAE	11.97	3.31
	R <sup>2</sup>	-17.529	0.783
SVR	MSE	59.54	29.96
	MAE	6.172	4.254
	R <sup>2</sup>	0.2778	0.639
RF	MSE	64.86	4.309
	MAE	5.703	1.497

	<b>R<sup>2</sup></b>	0.2132	0.948
<b>XGBoost</b>	<b>MSE</b>	<b>6.41</b>	<b>4.087</b>
	<b>MAE</b>	<b>5.997</b>	<b>0.0004</b>
	<b>R<sup>2</sup></b>	<b>0.1944</b>	<b>0.999</b>

The best performing model when predicting Obesity is XGBoost Regressor. The model shows MAE and MSE errors less both training and testing when compared to another model where RF works well in training the error is less, but in testing, the overfitting problem arise. The Fig. 11-14 below shows the obesity dependency on food features with the most positive correlation with the target (Obesity). Fig. 11 presents the relationship between meat and obesity.



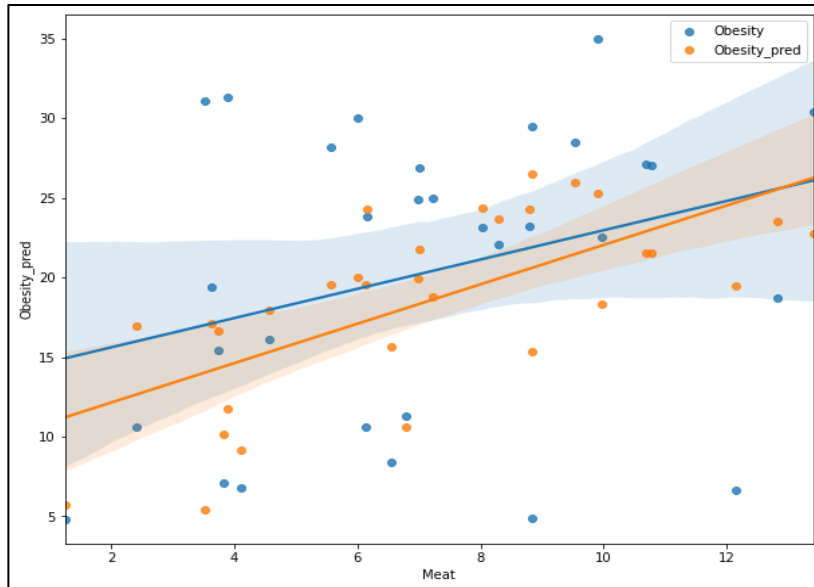


Fig. 11: Meat (actual obesity vs predicted obesity)

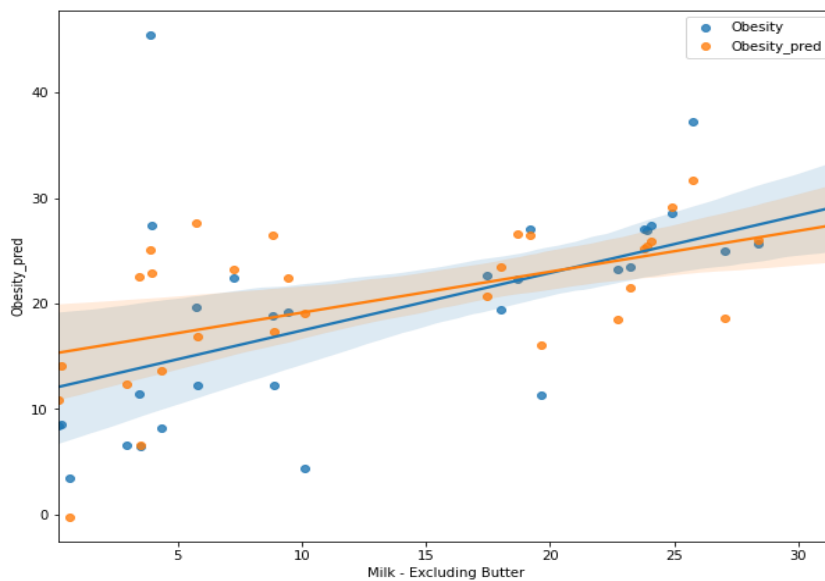


Fig. 12: Milk – excluding butter (actual obesity vs predicted obesity)

Fig. 12 presents the relationship between milk and obesity. Fig. 13. presents the relationship between sugary content & Obesity.

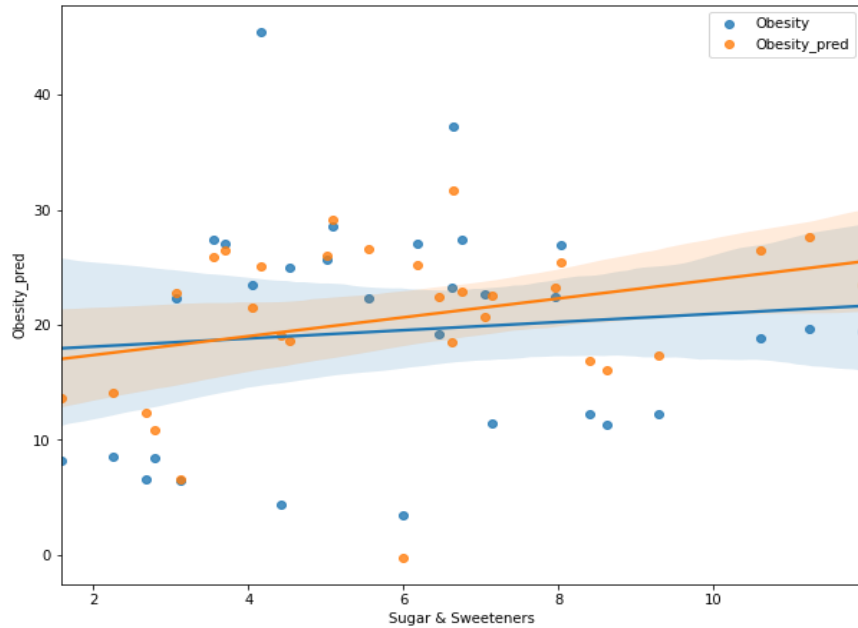


Fig. 13: Sugar & sweeteners (actual obesity vs predicted obesity)

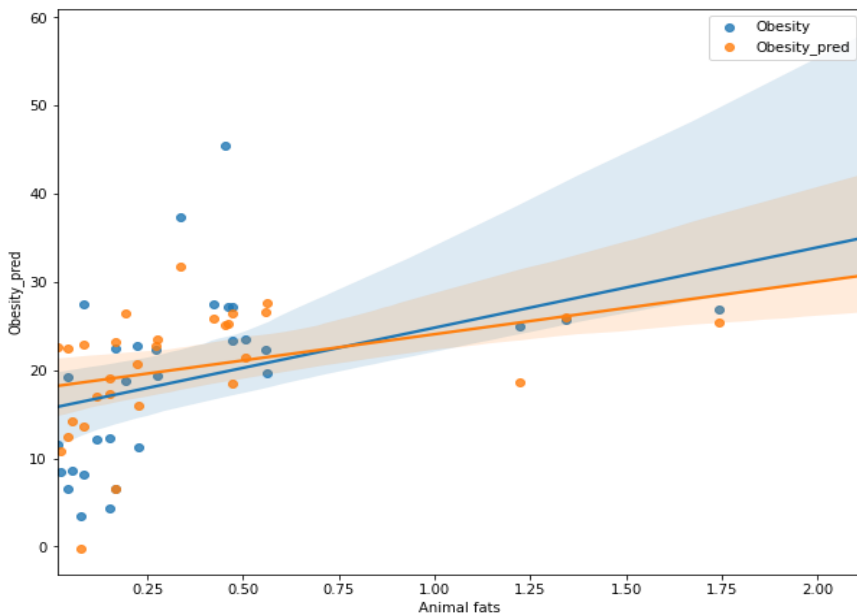


Fig. 14: Animal fats (actual obesity vs predicted obesity)

The model is doing a good job predicting this feature's influence on obesity. These models have better metrics than those built to predict Mortality as shown in Fig. 14.

## 5. Conclusion

In summary, a country's COVID-19 confirmed and active cases will somehow be explained by food classes like the calorie contents of oil crops and, therefore, protein content in baby food and miscellaneous food. As well as on the other hand, it cannot be said regarding the death and recovered cases. This might flow because these models don't satisfy the necessary model assumptions of getting equal variance and unremarkably distributed residuals. We created regression models to predict the Mortality of COVID-19 in countries based on their eating habits and obesity rate. Two approaches were taken: one with all food-related features taken as parameters and a simpler one, which reduced the dimensionality by using only two features: Animal Products and Vegetal Products. Both have issues (mainly of spread and non-linearity), but we could use different models and metrics. Next, we build a model to predict Obesity rates based on eating habits in each country. This model was far more fruitful & the general inclination of the information was caught & anticipated. However, recall that this model talks about the correlation between food classes and the rate of deaths. Also, several alternative factors inflicting the spread of COVID-19 are uncorrelated with diet, For example. How active the overall public are, the preventive measures enforced by the countries, population density, etc.

## References

Bohlouli, J., Moravejolahkami, A. R., Ganjali Dashti, M., Balouch Zehi, Z., Hojjati Kermani, M. A., Borzoo-Isfahani, M., & Bahreini-Esfahani, N. (2021). COVID-19 and fast foods consumption: A review. *International Journal of Food Properties*, 24(1), 203-209.

Bousquet, J., Anto, J. M., Iaccarino, G., Czarlewski, W., Haahtela, T., Anto, A., Akdis, C. A., Blain, H., Canonica, G. W., Cardona, V., Cruz, A. A., Illario, M., Ivancevich, J. C., Jutel, M., Klimek, L., Kuna, P., Laune, D., Larenas-Linnemann, D., Mullol, J., Papadopoulos, N. G., Pfaar, O., Samolinski, B., Valiulis, A., Yorgancioglu, A., & Zuberbier, T. (2020). The ARIA group, is diet partly responsible for differences in COVID-19 death rates between and within countries? Bousquet et al. *Clin Transl Allergy*.

Faour-Klingbeil, D., Osaili, T. M., Al-Nabulsi, A. A., Jemni, M., & Todd, E. C. (2021). The public perception of food and non-food related risks of infection and trust in the risk communication during COVID-19 crisis: A study on selected countries from the Arab region. *Food Control*, 121, 107617.

Gasmi, A., Chirumbolo, S., Peana, M., Noor, S., Menzel, A., Dadar, M., & Bjørklund, G. (2021). The role of diet and supplementation of natural products in COVID-19 prevention. *Biological Trace Element Research*, 1-4.

Grabia, M., Markiewicz-Żukowska, R., Puścion-Jakubik, A., Bielecka, J., Nowakowski, P., Gromkowska-Kępa, K., Mielcarek, K., & Socha, K. (2020). The

nutritional and health effects of the COVID-19 pandemic on patients with diabetes mellitus. *Nutrients*, 12(10), 3013.

Hassen, T. B., El Bilali, H., & Allahyari, M. S. (2020). Impact of COVID-19 on food behavior and consumption in Qatar, 12, 6973.

Hossen, M. S. & Karmoker, D. (2020). Predicting the probability of Covid-19 recovered in south asian countries based on healthy diet pattern using a machine learning approach. In *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, IEEE, 1-6.

<http://www.fao.org/home/en/>

<https://www.prb.org/>

<https://coronavirus.jhu.edu/map.html>

Husain, W. & Ashkanani, F. (2020). Does COVID-19 change dietary habits and lifestyle behaviors in Kuwait: A community-based cross-sectional study. *Environ Health Prev Med*, 25, 61.

Iddir, M., Brito, A., Dingo, G., Fernandez Del Campo, S. S., Samouda, H., La Frano, M. R., & Bohn, T. (2020). Strengthening the immune system and reducing inflammation and oxidative stress through diet and nutrition: considerations during the COVID-19 crisis. *Nutrients*, 12(6), 1562.

Klang, E., Kassim, G., Soffer, S., Freeman, R., Levin, M. A., & Reich, D. L. (2020). Severe obesity as an independent risk factor for COVID-19 mortality in hospitalized patients younger than 50. *Obesity*, 28(9), 1595-1599.

Klaus, W. L. (2020). Food science and human wellness. Department of Experimental Psychology, University of Regensburg, Regensburg, Germany.

Leddy, A. M., Weiser, S. D., Palar, K., & Seligman, H. (2020). A conceptual model for understanding the rapid COVID-19–related increase in food insecurity and its impact on health and healthcare. *Am J Clin Nutr*, 112, 1162–1169.

Lockhart, S. M. & O' Rahilly, S. (2020). When two pandemics meet: Why is obesity associated with increased COVID-19 mortality? *Med*, 1(1), 33–42.

Murray, C. J. L. (2019). Institute for Health Metrics Evaluation, University of Washington, Seattle, WA 98121, USA.

Pettit, N. N., MacKenzie, E. L., Ridgway, J. P., Pursell, K., Ash, D., Patel, B., & Pho, M. T. (2020). Obesity is associated with increased risk for Mortality among hospitalized patients with COVID-19. *Obesity*, 28(10), 1806-1810.

Popkin, B. M., Du, S., Green, W. D., Beck, M. A., Algaith, T., Herbst, C. H., & Shekar, M. (2020). Individuals with obesity and COVID-19: A global perspective on the epidemiology and biological relationships. *Obesity Reviews*, 21(11), e13128.

Pourhomayoun, M. & Shakibi, M. (2021). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health*, 20, 100178.

Rishi, P., Thakur, K., Vij, S., Rishi, L., Singh, A., Kaur, I. P., & Kalia, V. C. (2020). Diet, gut microbiota and COVID-19. *Indian Journal of Microbiology*, 1-10.

Schwingshackl, L., Schwedhelm, C., Hoffmann, G., Lampousi, A. M., Knüppel, S., Iqbal, K., Bechthold, A., Schlesinger, S., & Boeing, H. (2017). Food groups and risk of all-cause mortality: A systematic review and meta-analysis of prospective studies. Imperial College London, London, United Kingdom; and Institute for Biometry and Epidemiology, Leibniz Institute for Diabetes Research, Germany, 105,1462–73.

Villadsen, A., Patalay, P., & Bann, D. (2021). Mental health in relation to changes in sleep, exercise, alcohol and diet during the COVID-19 pandemic: Examination of five UK cohort studies. medRxiv.

Woolf, S. H., Chapman, D. A., Sabo, R. T., Weinberger, D. M., & Hill, L. (2020). Excess deaths from COVID-19 and other causes. March-April 2020. *Jama*, 324(5), 510-513.

Yadaw, A. S., Li, Y. C., Bose, S., Iyengar, R., Bunyavanich, S., & Pandey, G. (2020). Clinical features of COVID-19 mortality: Development and validation of a clinical prediction model. *The Lancet Digital Health*, 2(10), e516-e525.

Zupo, R., Castellana, F., Sardone, R., Sila, A., Giagulli, V. A., Triggiani, V., Cincione, R. I., Giannelli, G., & De Pergola, G. (2020). Preliminary trajectories in dietary behaviors during the COVID-19 pandemic: A public health call to action to face obesity. *International Journal of Environmental Research and Public Health*, 17(19), 7073.