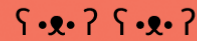


# Nutritional Profile of Countries Around the World



Group: Sahil Bansal, Vanessa Chang, Christine Han, Liza Kolev



## BACKGROUND

Nutritional profiles of countries around the world differ from each other and many factors may influence a country's nutritional profile. We thought it would be interesting to compare different countries' diets with each other to discover countries that have a similar nutritional profile and general trends across countries.

## DATA

We obtained our data from the World Health Organization, Global Nutrition Report, Food and Agriculture Organization of the UN, and ourworldindata.org. We joined together the datasets to create a database that we cleaned to conduct analysis.

## HYPOTHESES

- There is a significant difference in the percentage of fat between countries that have a higher GDP per capita purchasing power parity 2017 (henceforth GDP) versus countries that have a lower GDP.
- There is a negative correlation between alcohol consumption and life expectancy.
- There is a positive correlation between GDP and the amount of meat consumed in a country.

## METHODS

We used the following tests for each hypothesis.

1. Two Sample T-Test, with p-value 0.05
2. Spearman and One Sample T-Test
3. Spearman and One Sample T-test

### ML Models

1. Linear Regression: GDP vs Fat Percentage
2. KMeans Clustering of Nutritional Profiles



Figure 1: Bar Chart of Data Analyzed For Hypothesis One

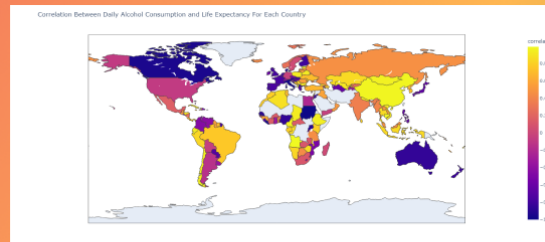


Figure 2: Choropleth Map of Hypothesis Two's Correlation Results

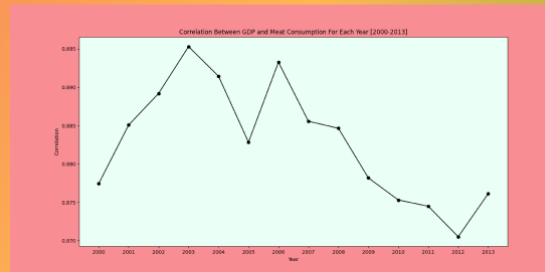


Figure 3: Line Graph of Hypothesis Three's Correlation Results

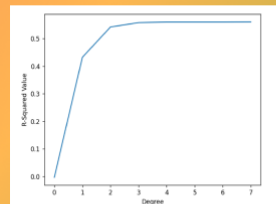


Figure 4: Graph of Average R-Squared Result Across Different Degrees of Machine Learning Model One

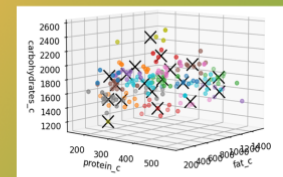


Figure 5: 3D Plot of Clusters of Nutritional Profile of Countries in 2013 from Machine Learning Model Two

## ML MODELS

- For ML Model 1, our best regression was for degree 2. It had an r-squared value of 0.54 meaning 54% of the variation in the dependent variable, percent of fat, is explained by the independent variable, GDP, in the regression model.
- For ML Model 2, we clustered the nutritional profile of different countries and used them in the interactive component of our project.

## RESULTS

1. The Two Sample T-tests returned p-values close to 0. Thus, we rejected the null hypothesis to conclude there is a **statistically significant difference in mean percentage of fat between countries that have a higher GDP v.s. lower GDP**
2. The Spearman Rank-Order Correlations and One Sample T-Tests returned p-values of 1.0, therefore we could not reject the null. Therefore, there is a **positive correlation between alcohol consumption and life expectancy**. However, it is worth noting that some countries' life expectancies had a negative correlation with their daily alcohol consumption.
3. The Spearman Correlation across countries and years returned a value of 0.88, which indicates positive correlation, and the One Sample T-Test returned a p-value well under 0.05. Therefore, we reject the null and conclude there is a **positive correlation between GDP and amount of meat consumed in a country**.

While our findings returned these results, it is important to keep in mind that our results are purely based off the data we aggregated and does not take into different cultural and societal nuances that could affect the nutritional diet of populations across the world. We also acknowledge there are some gaps in our data and therefore recommend using our research in conjunction with other findings rather than as a standalone to inform any decisions.

## CONCLUSION

Overall, in this project, there were many trends discovered among different countries and their nutritional profiles. The results show that there were indeed some correlations between some different attributes such as alcohol and life expectancy. This project could be extended to look at different attributes and nutritional aspects than the ones analyzed in this project.