Food Insecurity and Climate Change

Team Zambies: Michael Donoso, Sarah Liu, Daniel Fiume, Carmen Yu

1 | BACKGROUND

Climate change poses a significant global challenge, affecting people worldwide. Among its various impacts, the alteration of weather patterns, reduced water availability, and diminished crop yields are just a few consequences. Our project aims to employ data science techniques to explore and analyze the adverse effects of climate change on agricultural production by predicting crop yield and investigating undernourishment across the world.

2 | HYPOTHESIS

We hypothesize that the impact of crop yields, climate change, and other global indicators on food insecurity is significant.

3 | DATA

The datasets we used for this project are:

- FAOSTAT crops and livestock products
- FAOSTAT Food Security and Nutrition Indicators
- Global dataset on projected climate change indicators

4 | METHODOLOGY

We used the DBSCAN clustering algorithm to cluster countries by 26 food insecurity indicators present in the FAO dataset and compared these results to food insecurity levels in 2000 and 2019. We also used a Decision Tree Regressor for predicting the yield of three crops (rice, maize, and wheat) based on climate data.

5 | RESULTS

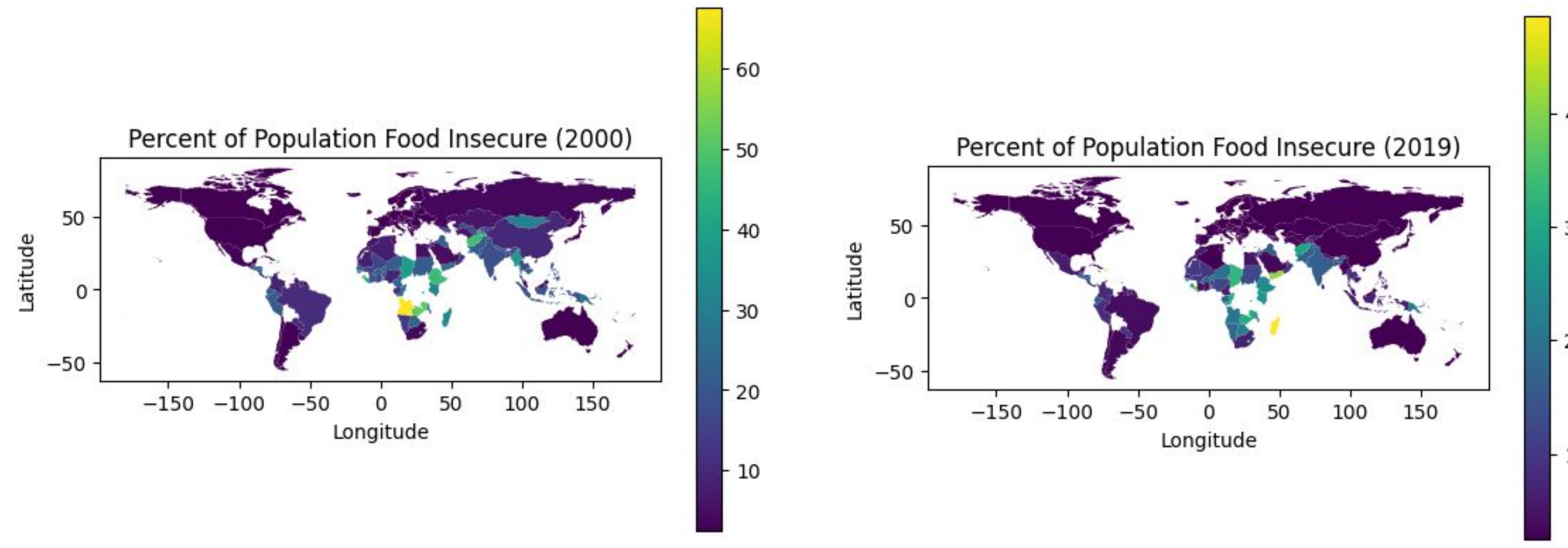

Figure 1: Percent of population by country that is food insecure in 2000 (left) and 2019 (right)

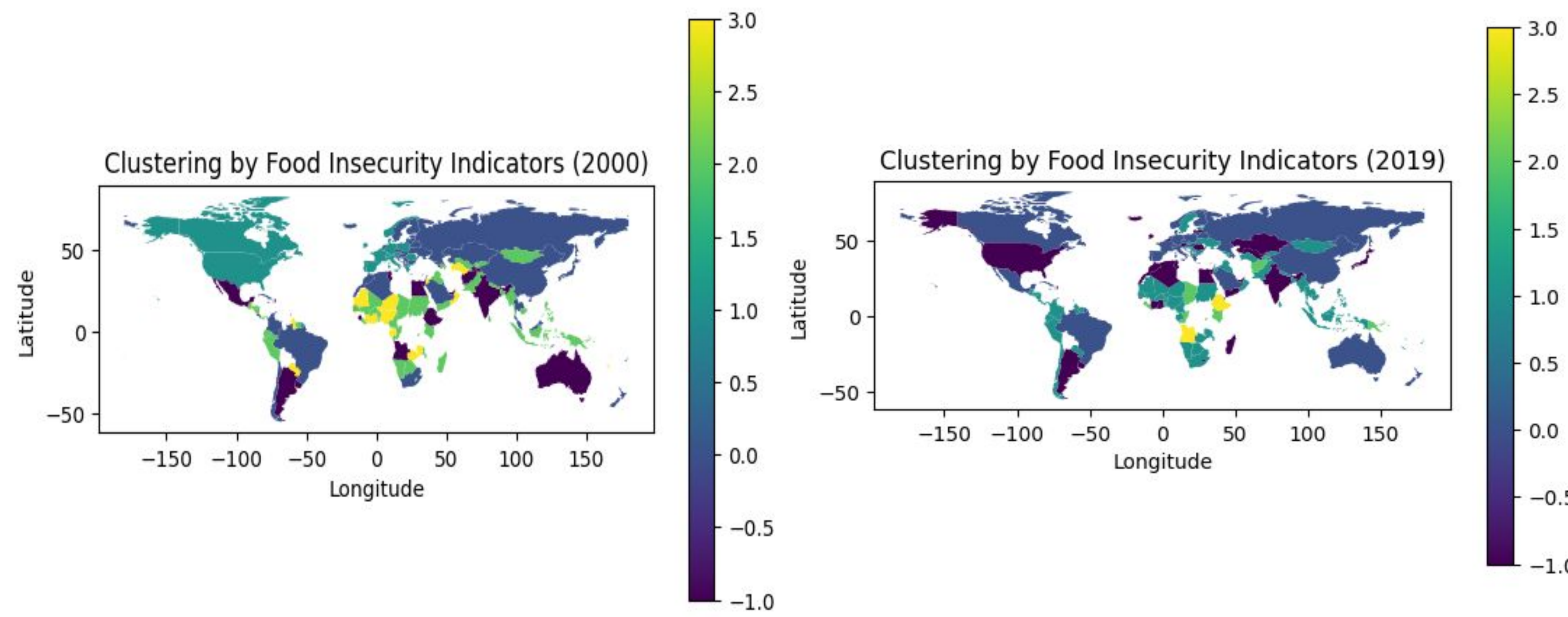

Figure 2: DBSCAN Clustering of Countries by Food Security Indicators

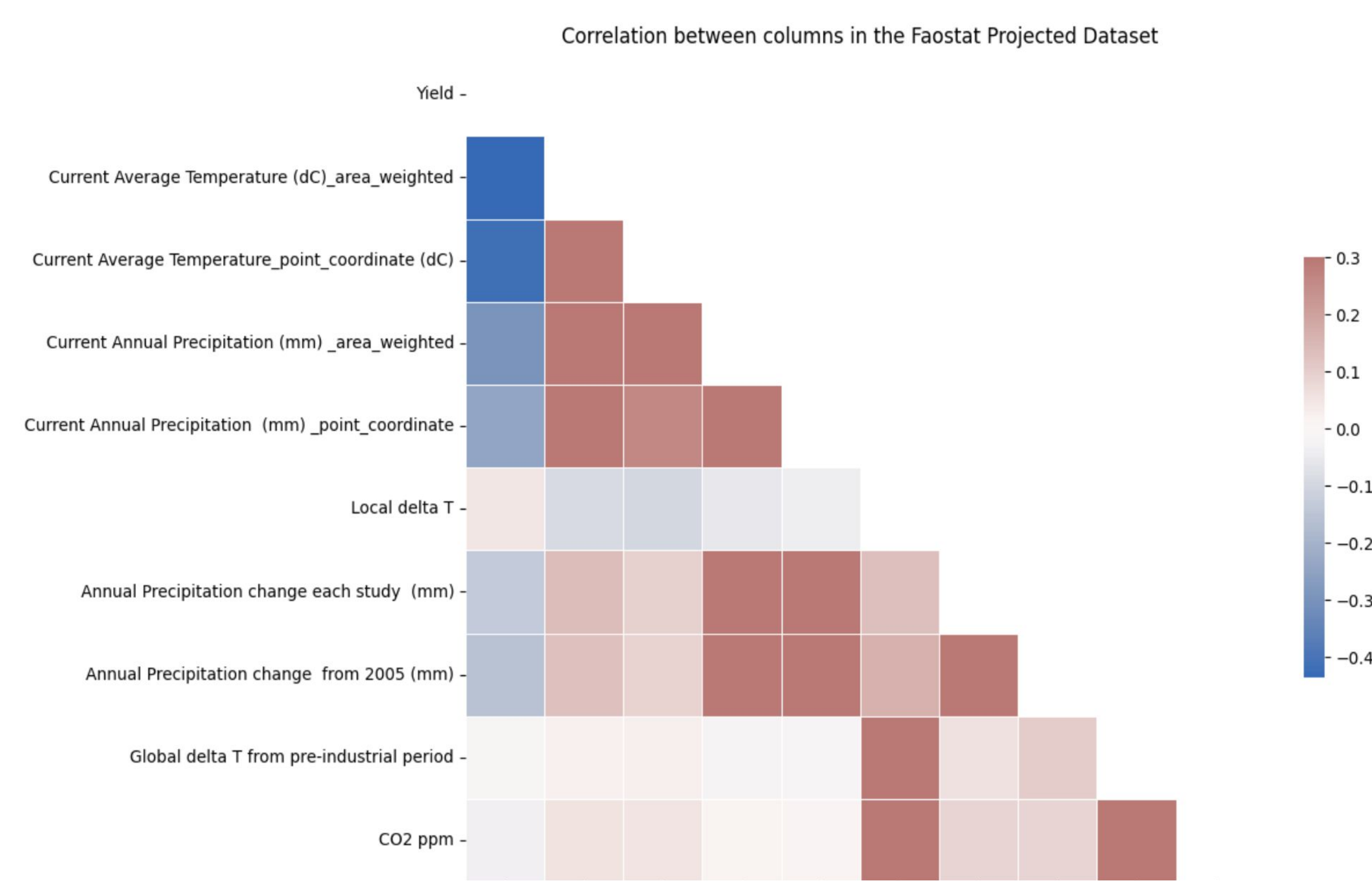

Figure 3: Correlation Matrix Between Features

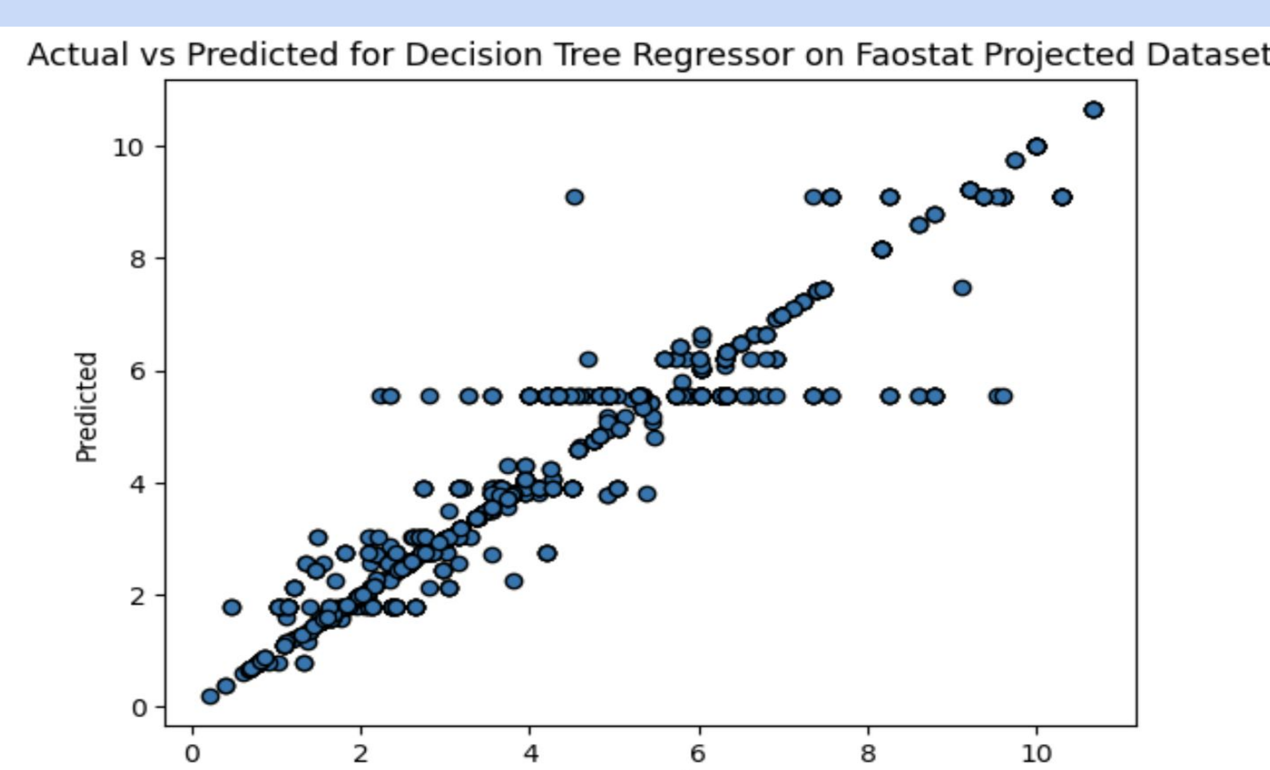

Figure 4: Actual vs. Predicted Crop Yield

6 | ANALYSIS

Clustering Analysis:

Overall, the resulting clusters were fairly indicative of how food insecure each country was: clusters 2 and 3 generally included countries with higher food insecurity than clusters 0 and 1. It is interesting to note the countries whose clusters changed between 2000 and 2019 but whose total food insecurity was fairly stable. This suggests that there are several different ways for a country to end up with a similar amount of insecurity.

Decision Tree Analysis:

MSE	0.4536
RMSE	0.6735
Cross-validation score	0.4262

Figure 5: Decision Tree Regressor Prediction Confidence

The small MSE measures the difference between actual and predicted yield, indicates a strong grasp of the relationship between variables. The RMSE is in a respectable range. The low cross-validation score boosts confidence in the model's ability to generalize to new data.

Hypothesis Tests:

- There is a correlation between annual precipitation and crop yield in the dataset.
- There is a significant correlation between CO_2 ppm and crop yield in the dataset.
- There is a significant difference in crop yield between countries that use irrigation and countries that don't use irrigation.

7 | DISCUSSION

Significance: Highlights climate change's impact on food insecurity, guiding action to mitigate its effects on global food systems.
Ramifications: Informs policy-making and agricultural practices for resilient and sustainable food production.
Limitations: Relies on historical data, limited scope to specific crops/regions, and additional factors warrant further investigation.