PREDICTING FOOD INSECURITY IN THE UNITED STATES BY FIPS CODE COUNTY

Daniel Kim
Advised by Professor Jonathan David Reuning-Scherer
Yale University
New Haven, CT, 06510
d.j.kim@yale.edu

ABSTRACT

This report aims to dissect the landscape of food deserts within the United States of America, a significant issue affecting many subsections of the population. Utilizing geographic divisions by FIPS Code counties, techniques such as Principal Components Analysis and Cluster Analysis help understand traits that relate to food insecurity and confirm ideas that existing literature on the nuances of the subject suggest. Moreover, the use of predictive modeling including techniques such as linear regression, ridge regression, and random forests forecasts food insecurity on the county level. The linear regression model provides the highest level of interpretability and accuracy; there appears to be a significant linear relationship between indicators of poverty in FIPS County with Food Insecurity Rates. This points to a relationship between factors indicative of poverty and higher rates of food insecurity among counties.

1. INTRODUCTION

Food insecurity is defined as "a lack of consistent access to enough food for every person in a household to live an active, healthy life" by the United States Department of Agriculture [1]. The issue of adequate food security has and continues to be a pressing issue in the United States [2]. The U.S., widely considered a prominent leader amongst first-world countries, still has had over 13 million households experiencing some level of food insecurity even before the onset of the COVID-19 pandemic [3]. This makes up for more than 38 million individuals with inadequate access to food, entrenched in cycles of food deserts and unable to meet their basic needs.

This report works to identify the communities that may fall within frequent food insecurity criteria and understand the factors that contribute to the patterns of inadequate food supply. With the use of several different models, this report quantifies the elements impacting food insecurity with the hope of predicting future levels of food shortfall; this can have implications on the ability to take preventative measures and allocate future resources towards particular counties with higher proportions of households in food need. This is particularly important as besides just being a human rights issue, food insecurity has implications on chronic disease and obesity inequities, particularly with the way the United States' food banks are structured currently [4].

2. DESIGN AND PRIMARY QUESTIONS

This paper aims to forecast what traits are the most indicative of the greatest variance in food insecurity across the United States using dimensionality reduction techniques such as Principal Components Analysis. A conceptual understanding of the factors affecting food insecurity and where the issue might have a more severe impact is useful information for policymakers in combating food insecurity, especially in the design of future policy [5]. Moreover, with the help of Cluster Analysis at the state level, counties with similar characteristics impacting food insecurity are grouped together, with the attempt to understand whether geographical proximity mirrors a pattern in food insecurity [6].

Machine learning techniques are put into practice in order to accurately predict future counties that may increase in severity of food insecurity. These techniques include linear regression, ridge regression, and random forest modeling. Features can then be fed into custom-made machine learning algorithms, where inputs are taken to deliver an output of the proportion of the population affected by food insecurity in a given area. The models are helpful in determining how critical an area would be given a set of features trained to the given area. While focusing on modeling food insecurity rates for the year 2019, the hope is that the framework of these models can be applied in future years to predict and preemptively aid counties that may emerge as areas of pressing food security need.

3. DATA

The food insecurity datasets were provided by Feeding America, a non-profit organization in the United States focused on domestic hunger relief. The datasets were pulled from their *Map the Meal Gap* campaign which is conducted annually to improve understanding of how food insecurity and food costs vary at the local level [7]. As there is a two year gap between analysis of food insecurity and the release of reports, this report focuses on data for 2019. This was combined with county-level data on educational attainment provided by the United States Department of Agriculture's Economic Research Service using an inner join. The datasets merged include information based on Poverty, Unemployment and Educational information, with a focus on the variables of Unemployment Rate, Proportion of Disabled Workers, Cost per Meal, Percent of all People in Poverty, Logarithm of Median Household Income, and the Percentage of People with Less than a High School Diploma. Here is a description of the said variables:

- Unemployment Rate is the percentage of the labor force that is actively seeking work but is unemployed.
- The Proportion of Disabled Workers is the number of disabled workers per FIPS Code County divided by the total population in the respective county.
- Cost per Meal is the average dollar amount spent on food per meal by food-secure individuals.
- The Percent of all People in Poverty is the estimated percent of all people of all ages in poverty in 2019.
- The Logarithm of Median Household Income is the logarithm of the median household income per FIPS County in 2019.
- The Percentage of People with Less than a High School Diploma is the proportion of residents per FIPS County that did not graduate from high school from the years 2015-2019. All of the variables are quantitative.

The response variable is the Household Food Insecurity Rate for the year 2019, which is expressed as a proportion; a household is declared to be in food insecurity if the respondents respond affirmatively to three or more questions from the Core Food Security Module in the December Supplement of the CPS for the years 2009-2019. Examples of such questions include "In the last 12 months, did you ever eat less than you felt you should because there wasn't enough money for food?" and "In the last 12 months, did you lose weight because you didn't have enough money for food?" [7].

The data was organized geographically by county-level FIPS Code, which is a five-digit Federal Information Processing Standards code that uniquely identifies counties and county-equivalents in the United States, certain U.S. possessions, and certain freely associated states.

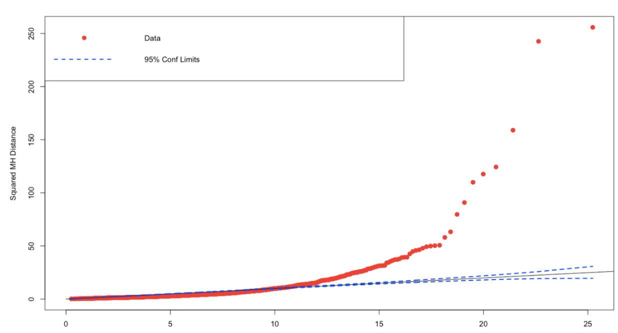

Figure 1: Chi-Square Quantiles for Food Insecurity Data

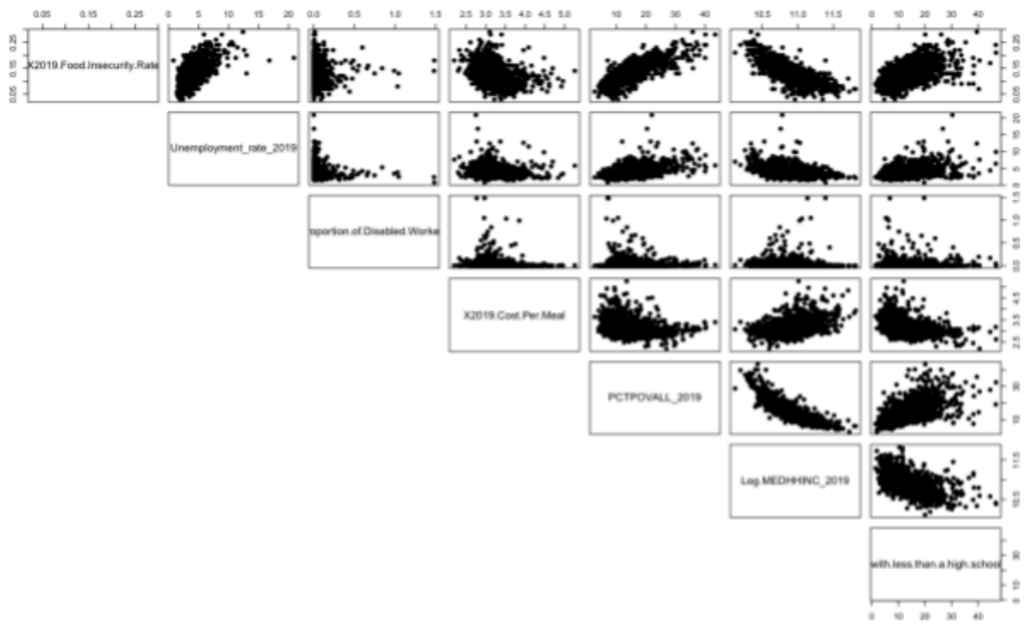

Figure 2: Matrix Scatterplot of Explanatory Variables

There exist 3141 unique FIPS Code United States counties in the final dataset since U.S. possessions and freely associated states were not included in the analysis.

Examining a chi-square quantile plot test in Figure 1, which uses the six variables detailed prior, the data does not appear to have a multivariate normal distribution.

4. PRINCIPAL COMPONENTS ANALYSIS

Principal Components Analysis is a technique of dimensionality reduction designed to maintain as much information in the dataset as possible. One core assumption of Principal Components Analysis is linearity in the dataset: essentially that the variables exhibit relationships amongst themselves.

While it is true that some of the variables did not have the strongest correlations with one another—Proportion of Disabled Workers in particular seems to have particularly low correlations with the other variables—there seems to be enough correlation between the other variables to proceed with Principal Components Analysis based on the results in Figure 3.

Moreover, it seems like the relationships are linear for the most part, when examining the graphs in Figure 2, rather than simply being random patterns. In particular, there seems to be a correlation between the Percentage of Adults with Less than a High School Diploma from 2015-2019 and the Percent of People of All Ages in Poverty in 2019.

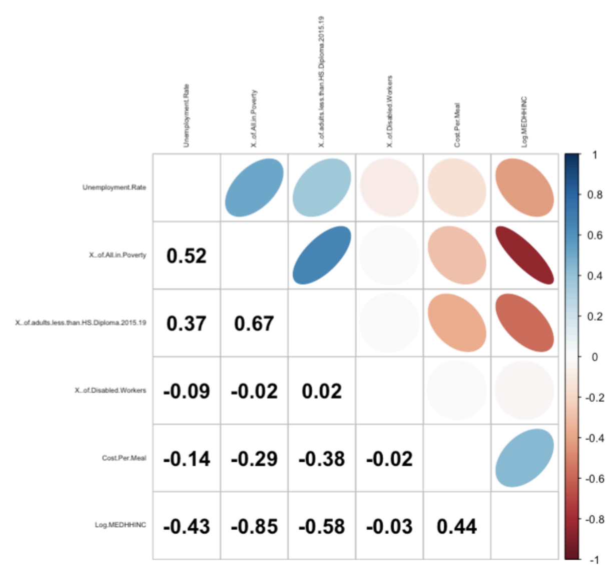

Figure 3: Correlation Matrix of Variables

Moreover, there seems to be a strong negative correlation between the Logarithm of Medium Household Income and the Percent of People of all ages in Poverty in 2019. This allows for enough confidence to proceed with Principal Components Analysis.

Using the *princomp* function in R, Principal Components Analysis on the aforementioned variables is performed. Below are the results from said analysis:

Table 1: Results from Principal Components Analysis

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6
Eigenvalue (SD^2)	2.87	1.05	0.92	0.55	0.49	0.13
Proportion of Variance	0.4781	0.1752	0.1527	0.0909	0.0818	0.0212
Stand. Deviation	1.6937	1.0253	0.9572	0.7386	0.7008	0.3568
Cum. Proportion	0.4781	0.6533	0.8060	0.8969	0.9788	1.0000

Table 1 shows that in order to explain 65.3% of the variance in the data, the first two principal components would be needed. Moreover, the first two eigenvalues are the only values greater than one. However, it appears that the first principal component is much more important than the second, as it makes up 47.8% of the variance in comparison to the second component's 17.5%. To explain 80.6% of the variance in the data, the first three principal components would be needed.

Examining Figure 4, a Scree Plot showing the variances plotted against the different principal components, there seems to be a significant elbow after the second principal component. It seems safe to proceed establishing a cutoff point at the second principal component for the purposes of this analysis. Note that the data is not of a multivariate distribution so parallel analysis is not performed.

Looking at the loading coefficients for the first two components in Table 2, which are of primary focus, Component One has positive correlations with Unemployment Rate, Percent of All People in Poverty, and Percent of Adults with less than a High School Diploma.

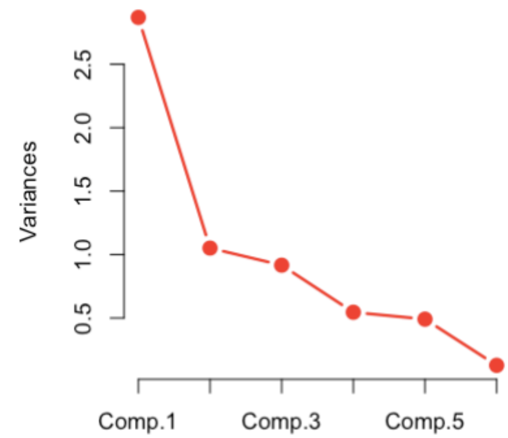

Figure 4: Scree Plot of Principal Components Analysis on Food Insecurity Data

It has negative correlations with the variables Proportion of Disabled Workers, Cost Per Meal, and Logarithm of Median Household Income.

Table 2: Loading Coefficients of Principal Components Analysis

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6
Unemployment Rate	0.37	0.40	0.43	0.68	0.23	0.08
Prop. of Disabled Workers	-0.01	-0.82	0.56	0.12	0.05	-0.02
Cost per Meal	-0.30	0.40	0.69	-0.50	-0.08	0.15
% of All in Poverty	0.54	0.06	0.15	-0.24	-0.30	-0.73
Log Median Household Income	-0.52	0.08	0.01	0.12	0.55	-0.64
% of Adults w/ Less than High School Diploma	0.46	-0.06	-0.04	-0.45	0.74	0.17

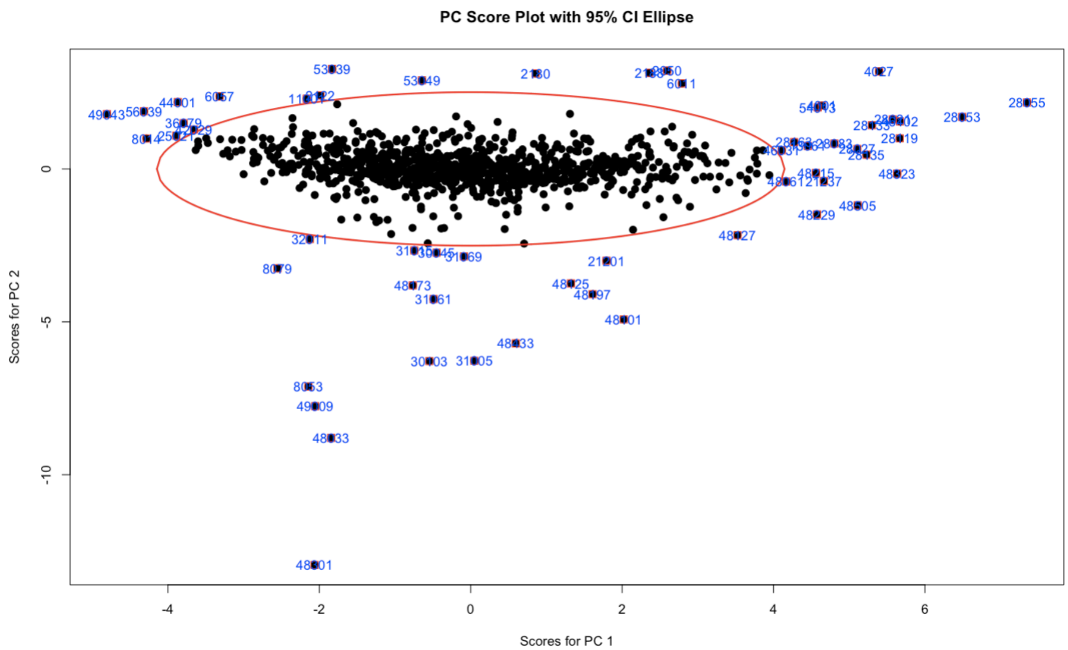

Figure 5: Principal Components Score Plot with 95% Confidence Interval Ellipse

The values with the largest magnitude, Percent of All People in Poverty and Logarithm of Median Household Income point to the belief that Component One is picking up on the variation in penury within different FIPS Codes in the United States. Component Two, which has positive correlations with Unemployment Rate and Cost Per Meal and a strong negative correlation with the Proportion of Disabled Workers. It appears that the second component is picking up on variation in the joblessness in different areas of the United States.

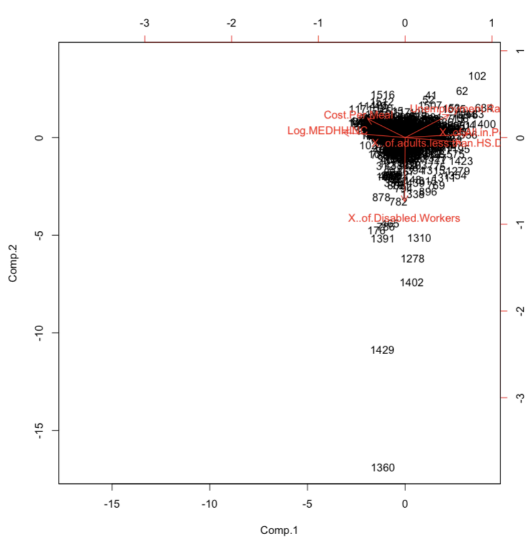

Figure 6: Biplot of First Two Principal Components

Figure 6 is the biplot of the first two principal components, the bottom and left-side axes are the principal component scores for the first and second components, respectively. The right-side and top axes are the loading coefficients for the principal components. There is more spread in the direction of Component 1 than in Component 2. The heavy clustering indicates that the different U.S. FIPS Codes tend to be mostly similar across the six variables.

Finally, a confidence ellipse plot is provided in Figure 5. The Confidence Interval Ellipse is particularly useful when the dataset has a multivariate distribution to test for outliers in the first two components. Though this dataset does not have a multivariate distribution, the plot can still be helpful looking at significant outliers and examining the landscape of variation within the components. Looking at the outliers in Figure 5, there seems to be a fair number of outliers in the direction of both principal components. In the direction of the first component, which is more significant, there are outliers in FIPS Codes such as 28133, 28135, and 28053 which correspond to counties in Northwestern Mississippi. These counties seem to have higher percentages of all people in poverty and unemployment rates than other counties in the United States.

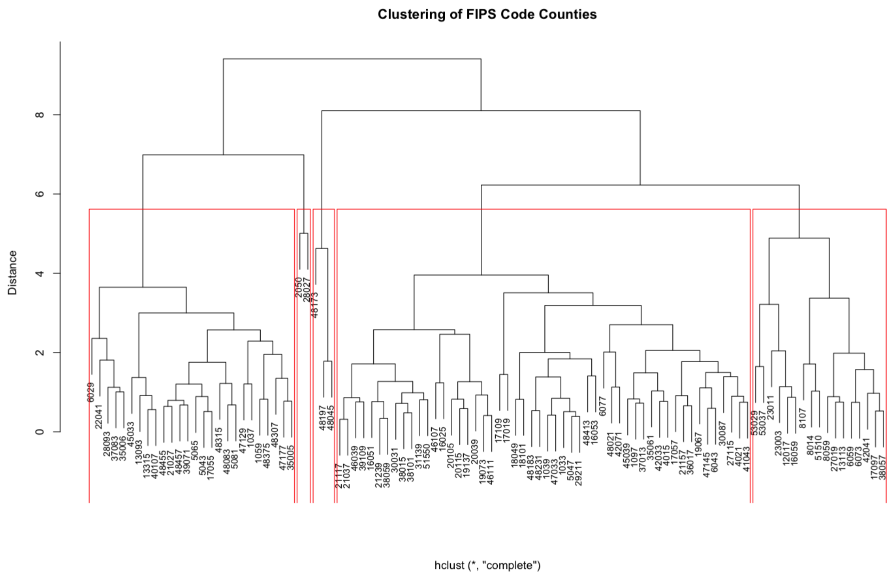

Figure 7: Clustering Based on Euclidean Distance and Complete Agglomeration

In the direction of the second component, there are outliers in FIPS Codes such as 48033, 48301, and 48173, which correspond to locations in West Texas. These counties have, for instance, have a higher proportion of disabled workers and lower cost of meals.

4. CLUSTER ANALYSIS

Cluster Analysis is an unsupervised machine learning method to identify and group similar data points into respective categories. To begin, 100 different FIPS Code counties are randomly sampled out of the 3141 counties to ensure interpretability of the clustering. Note that because this only comprises approximately 3 percent of the total data, cluster analysis is mainly being used to evaluate whether FIPS Code counties geographically near one another fall under similar cluster categories. The same variables utilized for Principal Components Analysis are used for this Cluster Analysis, with the data standardized to have the same mean and standard deviation. Standardizing the variables ensures that the differing ranges of the original variables won't be an issue when attempting to perform cluster analysis so that certain variables don't hold more weight than others.

The first dendrogram produced uses Euclidean distance as a metric to measure similarity between the different counties and the Complete Agglomeration method to define cluster distance. The Euclidean Distance Metric gives equal weight to each of the seven different variables when used to determine the distance between counties. The Complete Agglomeration method defines the cluster distance between two clusters to be the maximum distance between their individual components. At every stage of the clustering process, the two nearest clusters are merged into a new cluster.

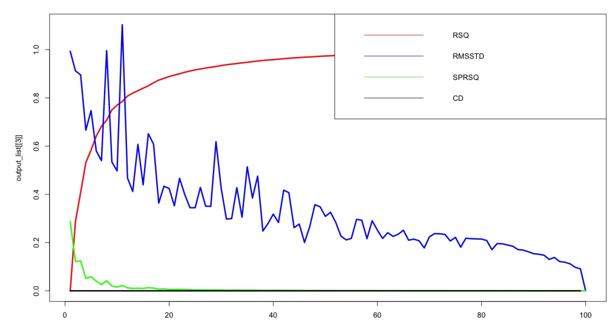

Figure 8: Cluster Evaluation Graph

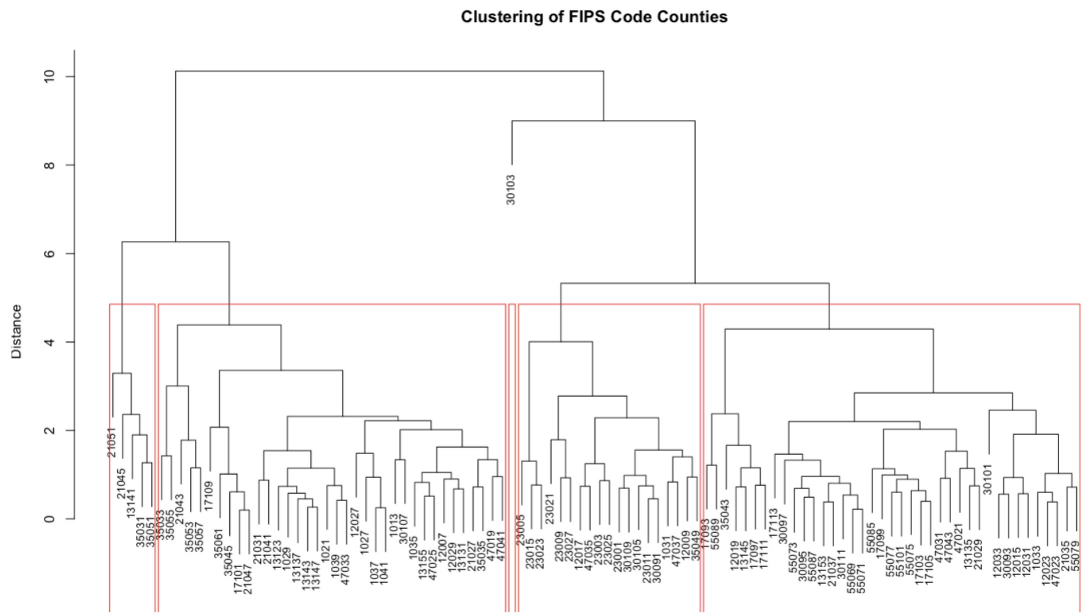

Figure 9: Cluster Analysis: Based on 10 Randomly Sample Counties and their Nine Closest Respective Counties

Looking at the produced dendrogram, there appear to be approximately five groups. To confirm this, metrics such as Root-Mean-Square Standard Deviation, R-Squared, Semi-partial R-Squared, and Cluster Distance are examined. The Root-Mean-Square Standard Deviation is the average standard deviation for however many variables there are. R-squared is the total sum of squares between clusters divided by the total sum of squares total. Semi-partial R-Squared measures the relative change in within-clusters sum of squares and Cluster Distance measures the determined similarity between elements.

From Figure 8, It looks like there are around 5 cluster groups. While the RMSSTD and CD lines offer little information, the points where the RSQ and SPRSQ curves start to level out is around 5 cluster groups. Counties that are geographically located near one another tend to mirror one another in the characteristics predicting food insecurity.

For instance, FIPS Code Counties 48197 and 48045 correspond to Hardeman and Briscoe counties in Texas respectively, both located in rural Texas and with similar values in the explanatory variables. Their food insecurity rates are 0.17 and 0.16 respectively. In another cluster, FIPS Code Counties 38101 and 38015 are closely matched. Representing Ward and Burleigh counties in central North Dakota, these FIPS Codes have food insecurity rates of 0.06 and 0.05 respectively.

To further supplement this analysis, 10 random FIPS Code Counties were sampled and for each respective county, the nearest nine counties were included in a cluster analysis. The counties are 13143 (Haralson County, GA), 12023 (Columbia County, FL), 30101 (Toole County, MT), 21037 (Campbell County, KY), 35061 (Valencia County, NM), 55077 (Marquette County, WI), 23003 (Aroostook County, ME), 47031 (Coffee County, TN), 01013 (Butler County, AL), 17111 (McHenry County, IL). Based on Euclidean Distance and Complete Agglomeration clustering, these 100 counties are grouped; the results are displayed in Figure 9. Similar to the previous results, it does appear that counties neighboring one another seem to cluster appropriately to what the literature has anticipated. This pattern can be visually observed in Figure 10, where neighboring counties tend to be similarly shaded. There are exceptions, however, especially when there are large wealth divides near adjacent counties

Areas of high food insecurity tend to be clustered with one another while areas of low food insecurity are similarly near one another. This is confirmed by the literature stating that “low–food

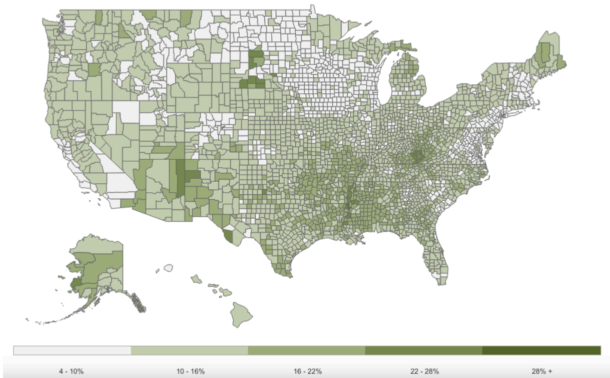

Figure 10: U.S. Map of Food Insecurity Rate

insecurity peer groups are often right beside each other" [6]. While not a hard and fast rule, geographical proximity seems to cluster areas of similar levels of food insecurity.

5. PREDICTION MODELS

In order to map the landscape of food insecurity, several different prediction models are used and evaluated. In this report, the three different techniques utilized are Linear Regression, Ridge Regression, and Random Forest Modelling. Detailed below are the assumptions and methods behind the models, along with their results.

5.1 Linear Regression

The first model used is a Linear Regression model to predict the metric of food insecurity rate for the year 2019.The food insecurity estimates coincide with the geographical boundaries set in 2019. Linear Regression is a linear approach to model the relationship between a scalar response variable and, in this case, multiple explanatory variables. First, the relationships between each of the potential predictor variables and the response variable are plotted.

From Figure 11, there seems to exist linear relationships between each of the predictor variables with the 2019 Food Insecurity Rate, with the exception of perhaps Proportion of Disabled Workers.

A linear regression using Python's *sklearn* package was created.

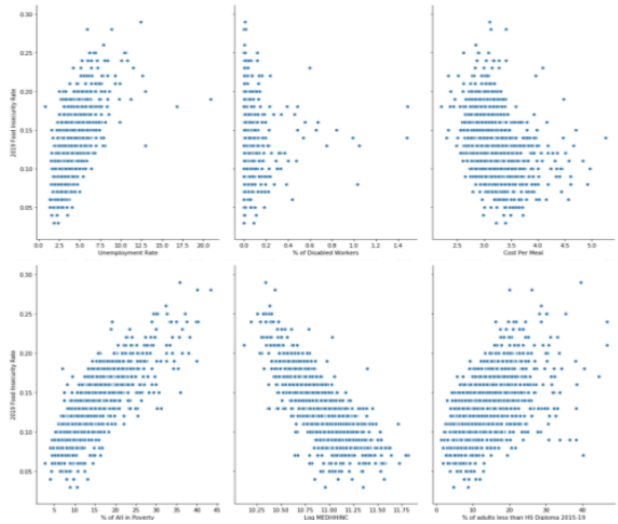

Figure 11: Predictor Variables Versus 2019 Food Insecurity Rate

The regression was trained on an 80%-20% Train-Test split, randomly selecting 80% of the FIPS Code counties to train on and evaluating based on the remaining 20% of the counties.

To evaluate the necessity of each of the predictors, T-values of each of the coefficients were calculated using a Matrix Algebra analysis of the *sklearn* linear regression. Table 3 shows that all of the coefficients seem to hold significance; the highest p-value comes in at 0.002 for the Proportion of Disabled Workers, which still falls under a 0.05 bound for significance commonly used.

Table 3: Statistical Analysis of Linear Regression Coefficients

	Coefficients	Standard Errors	t values	Probabilities
0	0.5850	0.026	22.837	0.000
1	0.0054	0.000	25.148	0.000
2	0.0101	0.003	3.168	0.002
3	-0.0065	0.001	-6.739	0.000
4	0.0023	0.000	22.640	0.000
5	-0.0454	0.002	-19.437	0.000
6	0.0005	0.000	9.369	0.000

To further prove the legitimacy of the explanatory variables, backwards stepwise regression is performed. The results point towards keeping all of the predictors in this particular model. Backward elimination starts with all of the predictors in the model and calculates the F-statistic as one of the regressors is removed one at a time. The predictor with the smallest F-statistic is removed from the model and this is continued until the smallest F-statistic is greater than a preselected cutoff value of F. This helps ensure, particularly when a model does not start with too many variables, that the best combination of regressors is selected. In this case, backward elimination resulted in keeping all six of the variables.

This seems to agree with the literature surrounding factors that impact food security across the United States. For instance, it is stated that “factors that lead to food insecurity include human capital,” including “lower education levels" [8]. Moreover, “income shocks, income volatility, and job loss” are all connected with higher rates of food insecurity in the United States [9]. Also, households with lower incomes are consistently found to be “more likely to be food insecure” [9]. The resulting linear regression model takes the form of (1)

$$Food\ Insecurity\ Rate = \alpha + \beta_1 UR + \beta_2 PDW + \beta_3 CPM + \beta_4 PPA + \beta_5 LMHI + \beta_6 LHSD + \epsilon \qquad (1)$$

where *UR* is unemployment rate, *PDW* is Proportion of Disabled Workers, *CPM* is Cost per Meal, *PPA* is Percent of all People in Poverty, *LMHI* is the Logarithm of Median Household Income, and *LHSD* is the proportion of people with less than a high school diploma. The coefficients for the equation are listed in Table 4.

Table 4: Linear Regression Model Coefficients

α	0.5788377942756636
β_1	0.00489582
β_2	0.01864759
β_3	-0.00766285
β_4	0.00232317
β_5	-0.04440106
β_6	0.00056139

When evaluated on the test set, a Mean Squared Error value of 0.000402892566 and Root Mean Squared Error value of 0.020072183893 is obtained. An R-Squared value of 0.7361560701424712 is achieved. This is the proportion of the variance in the response variable that can be explained by the predictor variables in the model. Equation (2) is the formula for for R-Squared.

$$R^2 = \frac{SSR}{SST} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} \qquad (2)$$

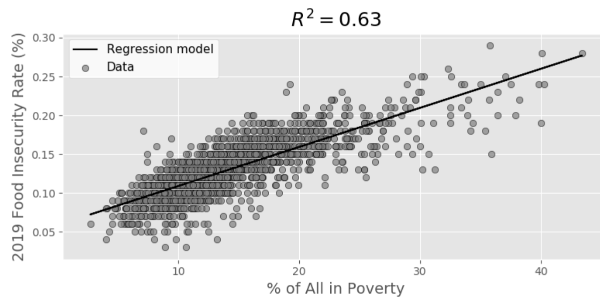

Figure 12: Percentage of All in Poverty Versus Food Insecurity Rate

An R-squared value of 0.63 is achieved using just the Percentage of All People in Poverty as an explanatory variable. Comprising more than 85.6% of the proportion of the variance in Food Insecurity Rate explained, this indicates that the Percentage of All People in Poverty is an important predictor in this model. More importantly discerned from Figure 12, however, is that the data shows that despite being noisy and high in variability, it can have a significant trend.

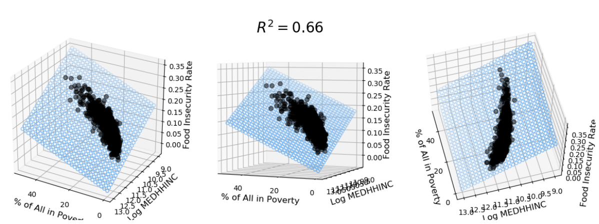

Figure 13: Percentage of All in Poverty & Logarithm of Median Household Income Versus Food Insecurity Rate

The trend indicates that the predictor variable still provides information about the response even though data points fall further from the regression line.

An R-squared value of 0.66 is achieved using the Percentage of All People in Poverty and the Logarithm of Median Household Income as explanatory variables, comprising more than 89.6% of the proportion of the variance in Food Insecurity Rate explained. There seems to be a linear relationship between the chosen explanatory variables and the response variable in this case.

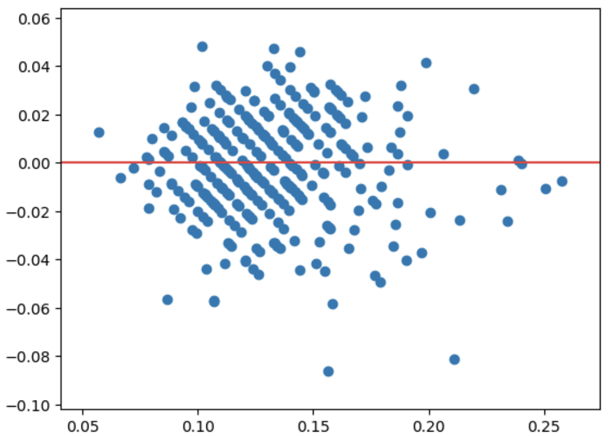

Figure 14: Residual Plot of Linear Regression on Test Set

In Figure 14, the residuals from the test set seem to be normally distributed evenly across the axis, clustering towards the middle of the plot. There does not seem to be any issues in the residuals with nonlinearity or heteroscedasticity. This indicates that despite the R-Squared value, the Linear Regression is valid with a normal distribution of the residuals, which is explained by the unobservable error in our model.

5.2 Ridge Regression

Ridge Regression is a method of estimating the coefficients of multiple-regression models in scenarios where independent variables are highly correlated. Its main benefit is that it prevents overfitting of the data. Ridge Regression utilizes L_2 Regularization. The Ridge Regression coefficients are the ones that minimize (3).

$$\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda\sum_{j=1}^{p}\beta_j^2 \quad (3)$$

where λ is a tuning parameter greater than or equal to zero. This is equivalent to the Residual Sum of Squares plus a shrinkage term. When λ is equal to zero, this equation is equivalent to least-squares linear regression.

In this case, the food insecurity data was trained with a Ridge Regression model with an alpha value of 1. When evaluated on the test set, a Mean Squared Error value of 0.002342736904 and Root Mean Squared Error value of 0.048401827487 is obtained. The R-Squared value from the Ridge Regression model is 0.71155773175429. The metrics for Ridge Regression seem to perform a little worse than that of Linear Regression, as the MSE and RMSE are both higher while the R-Squared value is slightly lower. When testing for various values of alpha, Figure 15 shows that the best value for alpha is 0.001 and 0.01, because they produce the highest R squared values. These are the alpha models that bring the model most approximate to linear regression.

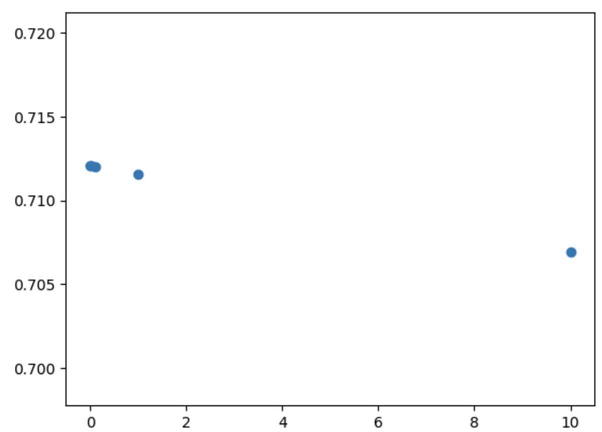

Figure 15: R-Squared Values for Different Alpha Values (0.001, 0.01, 0.1, 1, 10)

5.3 Random Forest Modeling

The last model is a random forest regressor to the data set. Random Forest is a supervised learning algorithm which utilizes an ensemble of decision trees. The Random Forest method is particularly useful because it corrects for decision trees' tendency to overfit to the training set. Random Forest builds multiple decision trees during training and outputs the mean of the classes as the prediction of all the trees. In fact, it utilizes random sampling of training observations when building trees and random subsets of features for splitting nodes.

Using permutation feature importances on the Random Forest Regressor model, the importance of each variable is observed by randomly shuffling each predictor variable seeing the effect on model accuracy. In order to do so, the variables are evaluated for their mean decrease in impurity importance by seeing how effective each of the explanatory variables is at reducing variance when creating decision trees within Random Forest. The results yield:

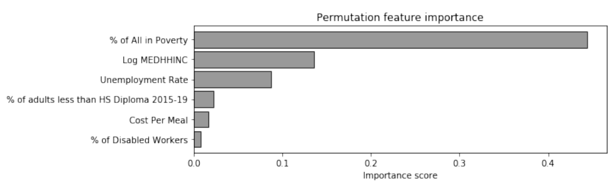

Figure 16: Random Forest Permutation Feature Chart

Similarly, to the linear regression model before, Percent of People of All Ages in Poverty 2019 and the Logarithm of Median Household Income seem to hold the most importance in the Random Forest model. The Percent of People of All Ages in Poverty has an importance score of just over 0.4 and the Logarithm of Median Household Income of just under 0.2. There appears to be a significant drop off after Unemployment Rate.

A Random Forest model is trained on the Food Insecurity dataset with 100 estimators and a max depth of five to prevent overfitting. For demonstrative purposes, Figure 17 shows what one of the decision trees of the Random Forest might look like with a max depth of three, limited to increase interpretability.

When evaluated on the test set, a Mean Squared Error value of 0.000432538481 and Root Mean Squared Error value of 0.020797559489 is obtained.

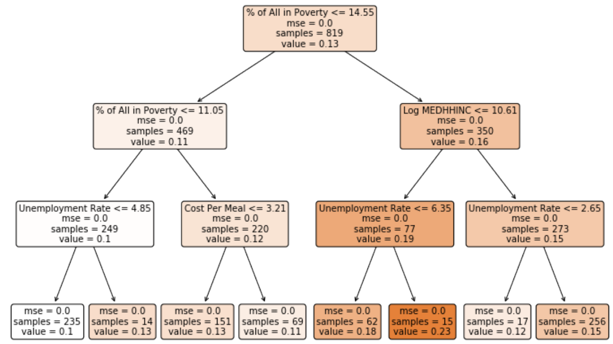

Figure 17: Sample Decision Tree in Random Forest

An R-Squared value of 0.7255548214489 is achieved. The MSE and RMSE values are just over that of the linear regression and the R-Squared value is just under that of the linear regression model. When considering the non-interpretability of Random Forest models along with the slightly higher metrics, the linear regression model is favored.

6. CONCLUSION

This report has yielded multiple forms of statistical analysis to interpret the outlook of food insecurity Principal Components Analysis yielded information about which factors explain most of the variance within food insecurity within the United States: particularly methods of penury and joblessness. Cluster Analysis aligned with existing literature relating geographical proximity to similarities in predictors of food shortfall.

Across the predictive models, the explanatory variables Percentage of All in Poverty and Median Household Income seem to be important predictors in the landscape of food insecurity across the United States. Representative of metrics of financial instability, these characteristics point to the notion that economic distress is correlated with food insecurity. This seems to agree with recent policy as federal income support in the form of stimulus checks and supplemental UI benefits helped boost food insecurity rates [10]. The linear regression model is the most favorable, giving clear indications on the importance of the predictors and also providing the lowest errors.

Barring any major changes in the landscape of food insecurity in the United States, one can utilize these models to make predictions on the level of shortfall by county. This could be particularly critical when implementing preventative measures such as the Supplemental Nutrition Assistance Program (SNAP) to reduce the severity of food deserts, rather than reactionary initiatives.

REFERENCES

[1] Definitions of Food Security. (n.d.). Retrieved December 13, 2021, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/definitions-of-food-security.aspx

[2] Pollard, C. M., & Booth, S. (2019, May 21). Food Insecurity and Hunger in Rich Countries-It is Time for Action against Inequality. Retrieved December 14, 2021, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6572174/

[3] Coleman-Jensen, A., Rabbitt, M. P., Gregory, C. A., & A. (2019). Household Food Security in the United States in 2018. Retrieved December 13, 2021, from https://www.ers.usda.gov/webdocs/publications/94849/err-270.pdf?v=963.1

[4] Decker, D., & Flynn, M. (2018, May 1). Food Insecurity and Chronic Disease: Addressing Food Access as a Healthcare Issue. Retrieved December 16, 2021, from https://pubmed.ncbi.nlm.nih.gov/29703073/

[5] Stropoli, R. (2021, February 24). How Policy Makers Could Address the Food-Insecurity Epidemic. Retrieved December 13, 2021, from https://www.chicagobooth.edu/review/how-policy-makers-could-address-food-insecurity-epidemic

[6] Lou, C. (2019, November 20). Tackling Food Insecurity in Different Places Requires Thinking Beyond Food. Retrieved December 13, 2021, from https://www.urban.org/urban-wire/tackling-food-insecurity-different-places-requires-thinking-beyond-food

[7] U.S. Household Food Security Survey Module: Three-stage Design, with Screeners. (2012, August 23). Retrieved December 16, 2021, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/survey-tools.aspx

[8] Gundersen, C. G., & Garasky, S. B. (2012, October). Financial Management Skills are Associated with Food Insecurity in a Sample of Households with Children in the United States. Retrieved December 12, 2021, from https://pubmed.ncbi.nlm.nih.gov/22955515/

[9] Gundersen, C., & Ziliak, J. (2018, March 01). Food Insecurity Research in the United States: Where We Have Been and Where We Need To Go. Retrieved December 12, 2021, from https://experts.illinois.edu/en/publications/food-insecurity-research-in-the-united-states-where-we-have-been-

[10] Rouse, C., & Restrepo, B. (2021, November 30). Federal Income Support Helps Boost Food Security Rates. Retrieved December 16, 2021, from https://www.whitehouse.gov/cea/written-materials/2021/07/01/federal-income-support-helps-boost-food-security-rates/