Socioeconomic Factors Associated with Food Insecurity Rates in Virginia

Abstract

Food insecurity is a public health issue that varies across local communities. In this study, we aim to investigate whether regional food insecurity rates across 133 Virginia counties and county-equivalents are more closely related to financial constraints or infrastructural barriers such as limited transportation. Using government data from the U.S. Census Bureau and USDA sources, as well as data from Feeding America, we constructed a multiple linear regression model to evaluate the associations of median household income, poverty level, cost per meal, metro status, unemployment rate, vehicle access, and local food availability. Our final model showed that local economic conditions, including unemployment, poverty level, and median household income, were the clearest predictors of food insecurity in Virginia. Furthermore, a lack of access to personal transportation may also be related to food insecurity, although the relationship was weaker than the economic factors in our model.

1 Introduction

Food insecurity remains a critical public health crisis, acting as a primary indicator of both community food access and economic vulnerability (Feeding America, 2018). Research has shown that food insecurity is closely tied to neighborhood food environments and health disparities, with communities facing structural barriers experiencing disproportionately worse outcomes (Odoms-Young, 2024). According to recent data, 13.7% of U.S. households (roughly 18.3 million) experienced food insecurity in 2024 (Hales, n.d.). This includes approximately 7.3 million children who lived in households where both children and adults suffered from inadequate food access (Hales, n.d). These figures highlight that despite overall national economic growth, a significant number of households still face barriers to food access.

Virginia provides a useful setting for studying these local differences because its 133 counties and county-equivalents exhibit significant variations in median household income, poverty level, cost per meal, metro status, unemployment rate, vehicle access, and local food availability. It is precisely these regional differences that help us analyze and explain why food insecurity rates are higher in some counties than in others. In this project, we aim to examine how these barriers appear locally in Virginia counties. We are specifically focused on why certain counties have higher insecurity rates than others. This motivated us to look at three specific questions:

1. Is lower transportation access associated with higher food insecurity across Virginia counties and county-equivalents?
2. Are economic conditions such as median household income and unemployment rate associated with food insecurity rates across Virginia counties and county-equivalents?
3. Does a higher poverty level relate to higher food insecurity across Virginia counties and county-equivalents?

We are motivated to explore these questions to identify gaps in the food supply network, especially for residents lacking personal vehicles, and to distinguish between infrastructural barriers and direct financial constraints. By quantifying whether transportation access, income, unemployment, or poverty are most strongly associated with food insecurity, our goal is to provide insights necessary for targeted resource distribution.

2 Data and Methods

Our dataset consists of 133 observations, with each row representing a Virginia county or county-equivalent. The data were compiled from three primary sources: the U.S. Census Bureau, the USDA Economic Research Service, and a reputable nonprofit organization called Feeding America. The response variable was the food insecurity rate, defined as the estimated percentage of the county's population living in a state of food insecurity. The explanatory variables included median household income, poverty level, cost per meal, metro status, unemployment rate, vehicle access, and local food availability.

Before model fitting, we conducted exploratory analyses on the VA_FoodAccess dataset to explore the distribution and relationships among key variables. The histogram of food insecurity rate showed an approximately symmetric, bell-shaped distribution with a mean of 11.1% (SD = 3.4%), ranging from 3.8% (Loudoun County) to 20% (Petersburg City). The absence of strong skew or extreme outliers suggested that no transformation of the food insecurity rate was necessary before fitting the regression model. In addition, we used scatterplots, boxplots, and the correlation matrix to examine the associations between the response variable and the exploratory variables.

Next, we used a multiple linear regression. For easier interpretation and analysis, we recoded several predictors as categorical variables, including metro status (metro vs. nonmetro), poverty level based on poverty rate (low, medium, high), and farmers market availability (none, low, medium, and high). We did not include FI_Persons and Ann_Budget_Shortfall because they were directly derived from the response variable, FI_Rate, which could lead to a circular relationship and an overly optimistic model performance. After choosing the final model, we checked residual plots, a Q-Q plot, multicollinearity measures (VIF values), and influence diagnostics. We also used a 10-fold cross-validation method to assess the model's consistency in predicting food insecurity rates across Virginia counties and county-equivalents.

3 Results

Overall, the final model containing median household income, unemployment rate, percentage of households without a vehicle, and poverty level was deemed useful for predicting food insecurity rates across Virginia counties and county-equivalents. The adjusted R-squared value of the final model was 0.873, which means that approximately 87.3% of the variation in food insecurity rates across counties was explained by the predictors in the model. The global F-test ($p < 2.2e\text{-}16$) was also statistically significant, indicating that the model was useful for explaining variations in food insecurity at the county level. Cross-validation further supported the stability of the model within the range of the observed data.

Table 1: Final regression model for county food insecurity rate

Predictor	Estimate (β)	p-value
Median household income	-0.000073	< 0.001
Unemployment rate	1.448	< 0.001
Vehicle access (% Households Without a Vehicle)	0.0000095	0.071
Medium poverty level	0.159	0.734
High poverty level	1.779	0.003

Note. Estimates are interpreted on the percentage-point scale for food insecurity rate.

The results above show that median household income and unemployment rate were the strongest quantitative predictors. Holding the other variables constant, a $1,000 increase in median household income was associated with a decrease of approximately 0.073 percentage points in food insecurity rate. In contrast, every 1 percentage point increase in the unemployment rate was associated with an increase of approximately 1.45 percentage points in food insecurity rate. These results suggest that counties with higher levels of economic hardship tended to have higher rates of food insecurity.

Poverty level also helped explain differences across counties. After controlling for the other variables, counties in the high-poverty group had food insecurity rates approximately 1.78 percentage points higher than counties in the low-poverty group. No clear difference was found between the medium-poverty and low-poverty groups. The percentage of households without a vehicle had a positive coefficient, but the evidence was weaker than other predictors. For this reason, we treated this result as suggestive rather than conclusive.

To check the model's stability, we performed 10-fold cross-validation using the caret package in R. Specifically, we incorporated this technique by splitting our data of 133 counties and county-equivalents into 10 subsets or folds, training on nine, and testing on the remaining final fold for 10 iterations. These results were then averaged out, ensuring that each fold served as both a training and testing set exactly once. The consistency across all 10 folds, shown through R-squared values ranging from 0.766 to 0.961, suggests that the model performed reasonably well across different subsets of counties and was not dependent on only a few influential observations. We also found an RMSE of 0.012, indicating that our model's predictions differed from the observed food insecurity rates by approximately 1.22 percentage points on average.

4 Conclusion/Discussion

The final prediction model for food insecurity rate across Virginia counties and county-equivalents is: **FI_Rate = 10.46 - 0.000073(Med_Income) + 1.448(Unemp_Rate) + 0.0000095(Pct_No_Vehicle) + 0.159(Medium_Poverty) + 1.779(High_Poverty)**. Overall, the results were consistent with our research hypotheses. Counties with lower income, higher unemployment, and high poverty levels tended to have higher food insecurity rates. Therefore, economic hardship is strongly associated with county-level food insecurity. The percentage of households without a vehicle also showed a positive relationship with food insecurity, although this result was weaker than the economic variables and should be interpreted more carefully. To illustrate prediction, we estimated the food insecurity rate for a hypothetical county with median household income of $45,000, unemployment of 8.2%, and 5.1% of households without a vehicle, assuming the county is in the high-poverty group (Medium_Poverty = 0, High_Poverty = 1): **FI_Rate = 10.46-0.000073(45000) + 1.448(8.2) + 0.0000095(5.1) + 0.159(0) + 1.779(1) = 20.8%**. The 90% prediction interval for the food insecurity rate of a county with these characteristics was between 19.26% and 24.58%. This interval includes Petersburg's observed rate of about 20%, suggesting that the model captures a realistic level of food vulnerability. Although annual food budget shortfall was relevant descriptively, it was not retained in the final predictive model because it was too closely tied to the response. Including it would have made the model less helpful for understanding which outside county-level factors are related to food insecurity. For this reason, the final model focused more on broader economic and access-related variables. This analysis has some limitations. Since the data are observational, the results should be interpreted as associations, not proof of cause and effect. The model also works at the county level, so it may miss important differences between households within the same county. In addition, the model underpredicted food insecurity for counties such as Emporia and Norton, which suggests that some local factors were not fully captured. These could include food access infrastructure, service availability, transportation options, or regional policy differences.

In future work, we would consider additional predictors such as SNAP participation, grocery access, housing burden, and public transportation. It would also be useful to test whether the same patterns appear outside Virginia. Overall, the results suggest that food insecurity across Virginia counties and county-equivalents is strongly associated with economic vulnerability, especially income, unemployment, and poverty level.

5 References

Feeding America. Map the meal gap [Data set]. ArcGIS Hub. https://hub.arcgis.com/search?collection=dataset&layout=grid&tags=feeding%2520america

Feeding America. (2025). How we got the map data: Map the Meal Gap methodology. https://www.feedingamerica.org/research/map-the-meal-gap/how-we-got-the-map-data

Hales, L. J. (n.d.). Food security in the U.S. - Key statistics & graphics. USDA Economic Research Service. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/key-statistics-graphics

Odoms-Young, A. (2024, March). Food insecurity, neighborhood food environment, and health disparities: State of the science, research gaps and opportunities. The Journal of Nutrition, 154(3). https://www.sciencedirect.com/science/article/pii/S000291652366352X

Semega, J., Kollar, M., Creamer, J., & Mohanty, A. (2019). Income and Poverty in the United States: 2018 (Report No. P60-266). U.S. Census Bureau. https://www.census.gov/content/dam/Census/library/publications/2019/demo/p60-266.pdf

U.S. Census Bureau. 2018 State and county poverty estimates [Data set]. https://www.census.gov/data/datasets/2018/demo/saipe/2018-state-and-county.html

U.S. Census Bureau. Annual estimates of the resident population: 2018 [Data set]. https://www.census.gov/programs-surveys/popest/data/data-sets.2018.html

U.S. Census Bureau. (n.d.). American Housing Survey (AHS) Methodology. Retrieved from https://www.census.gov/programs-surveys/ahs/about/methodology.html

U.S. Census Bureau. Physical housing characteristics for occupied housing units (Table S2504) [Data set]. American Community Survey 5-Year Estimates. https://data.census.gov/table/ACSST5Y2023.S2504?q=S2504&g=040XX00US51$0500000

U.S. Hunger. (2025, February 14). It's not just hunger: Transportation's role in food insecurity. https://ushunger.org/blog/transportation-food-insecurity/

USDA Economic Research Service. County-level data sets [Data set]. https://data.ers.usda.gov/reports.aspx?ID=4038Population

USDA Economic Research Service. Food environment atlas - Local foods [Data set]. Virginia Open Data. https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads

USDA Economic Research Service. Rural-urban continuum codes [Data set]. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes

Zach. (2020, November 4). K-Fold Cross Validation in R (Step-by-Step). Statology. https://www.statology.org/k-fold-cross-validation-in-r/

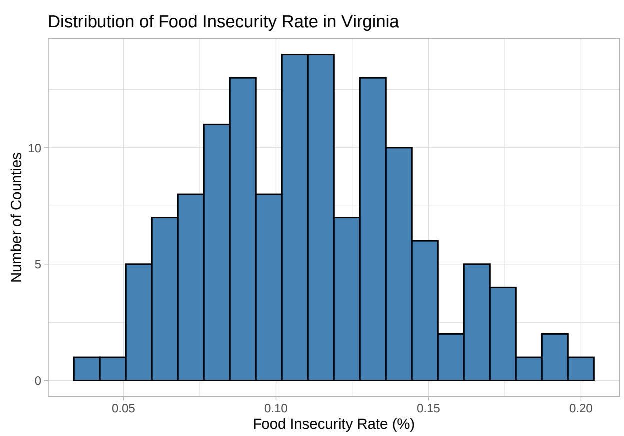
Generative AI Statement

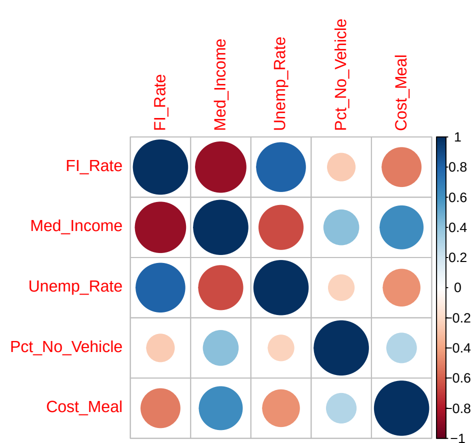
Generative AI tools were used to help revise grammar, wording, and clarity of expression. Additionally, Gemini was used to help refit the appendix sections and help debug. All other aspects of this work were independently completed and reviewed by the student authors.

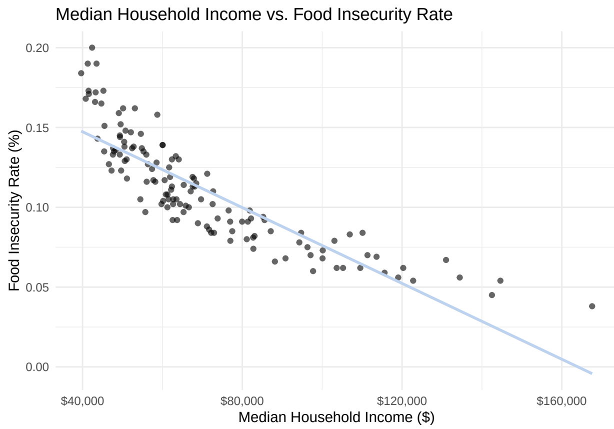
Appendix A: Data Dictionary

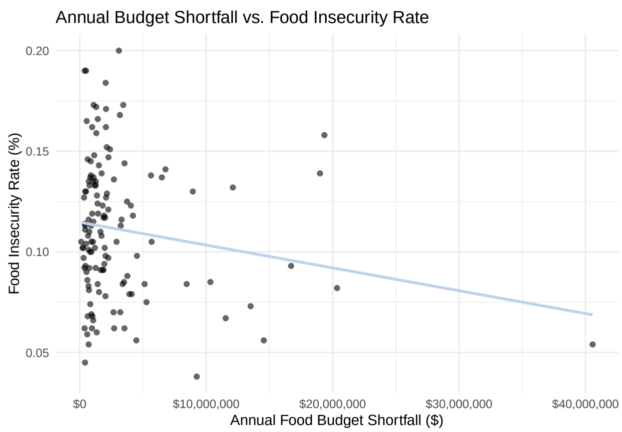
Variable Name	Abbreviated Name	Description
Food Insecurity Rate	FI_Rate	Percentage of county population estimated to be food insecure (quantitative)
Annual Food Budget Shortfall	Ann_Budget_Shortfall	Estimated total yearly dollar gap between food needs and available resources in the county (quantitative)
Cost Per Meal	Cost_Meal	Estimated average cost of a single meal in the county in USD (quantitative)
Food Insecure Persons	FI_Persons	Total number of food insecure individuals in county (quantitative)
Unemployment Rate	Unemp_Rate	County-level unemployment rate as a percentage (quantitative)
Median Household Income	Med_Income	Median household income in USD for the county (quantitative)
% Households Without a Vehicle	Pct_No_Vehicle	Percentage of occupied housing units with no vehicle available (quantitative)
Metro vs. Non-metro Status	Metro_Status	Whether the county is classified as metro or non-metro (qualitative)
Poverty Rate	Poverty_Rate	Categorized poverty rate: Low (<7%), Medium (7–15%), High (>15%) (qualitative)
Farmers Market Availability	FM_Availability	Number of farmers markets categorized as: None (0), Low (1–5), Medium (5–10), High (>10)

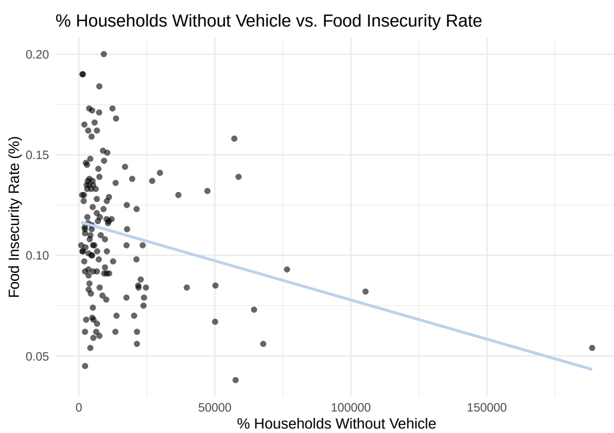
Appendix B: Exploratory Data Analysis

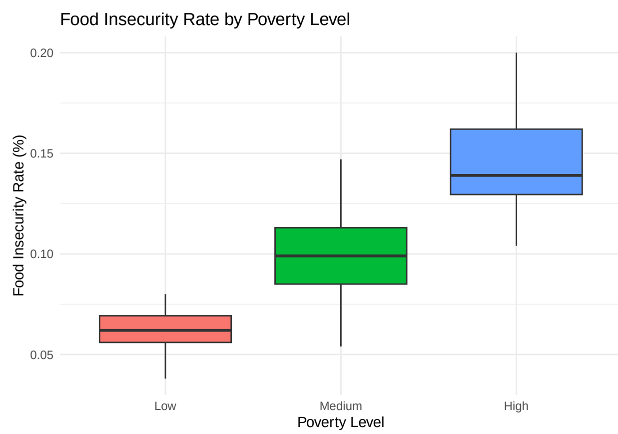


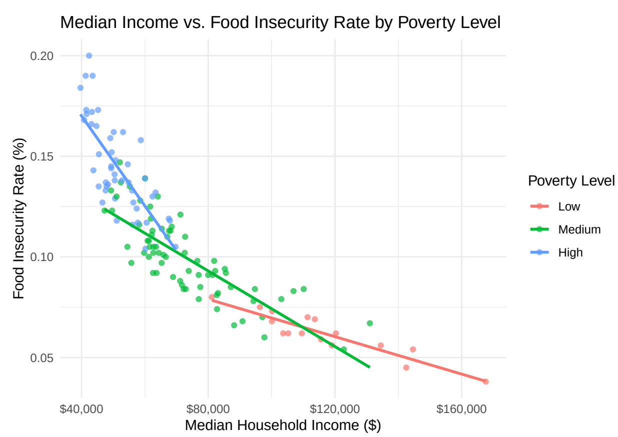



Median Household Income vs. Food Insecurity Rate
Food Insecurity Rate (%)
0.20
0.15
0.10
0.05
0.00
$40,000
$80,000
$120,000
$160,000
Median Household Income ($)

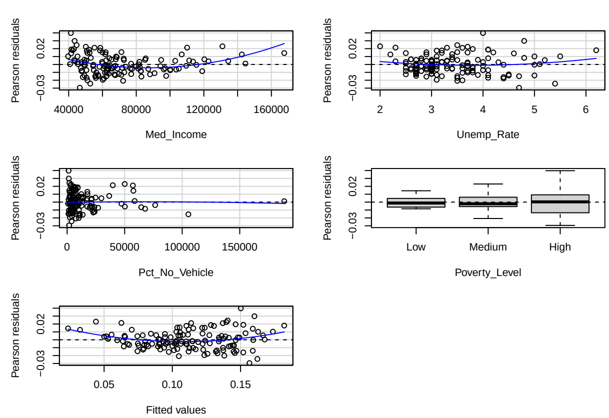

Annual Budget Shortfall vs. Food Insecurity Rate
Food Insecurity Rate (%)
0.20
0.15
0.10
0.05
$0
$10,000,000
$20,000,000
$30,000,000
$40,000,000
Annual Food Budget Shortfall ($)


% Households Without Vehicle vs. Food Insecurity Rate
Food Insecurity Rate (%)
0.20
0.15
0.10
0.05
0
50000
100000
150000
% Households Without Vehicle



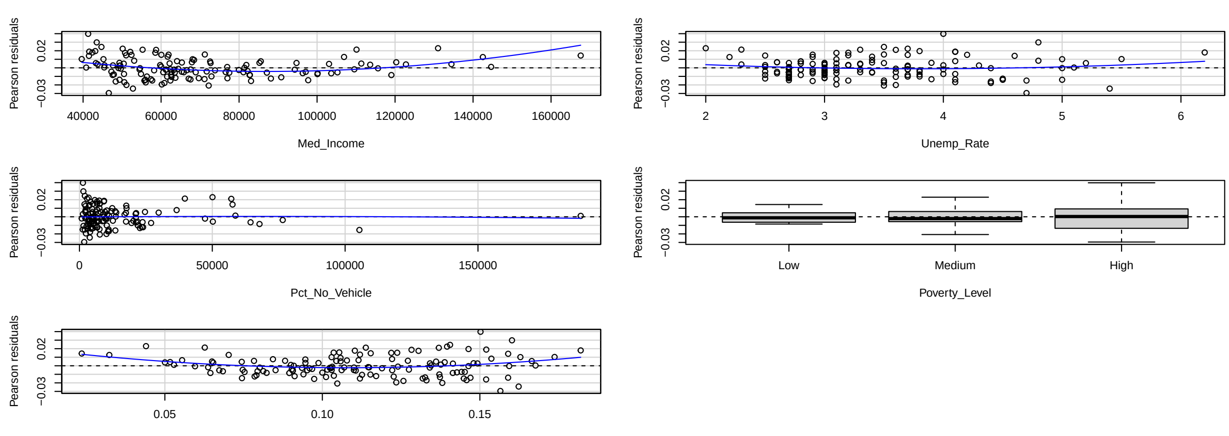


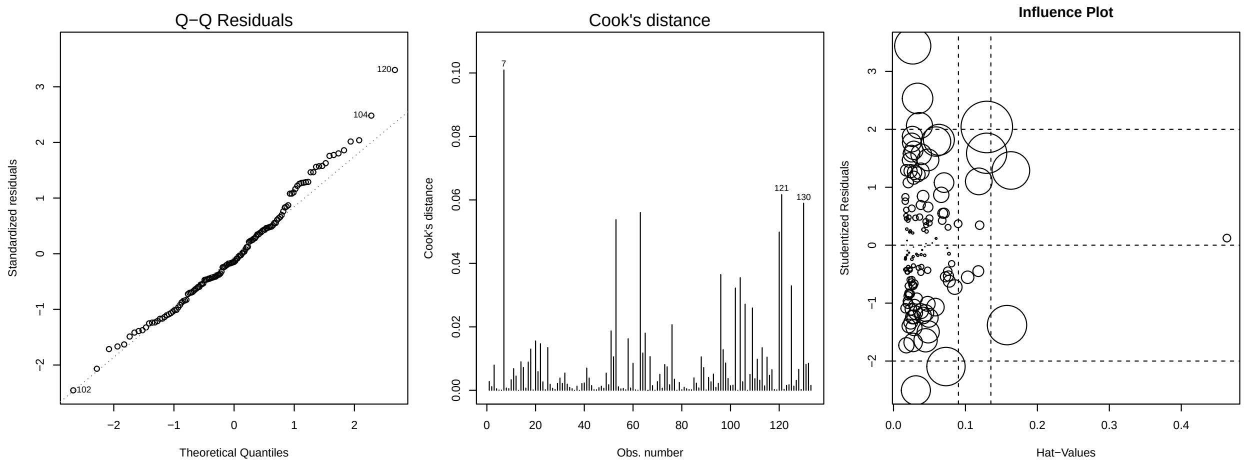
Appendix C: Final Model Output and Plots

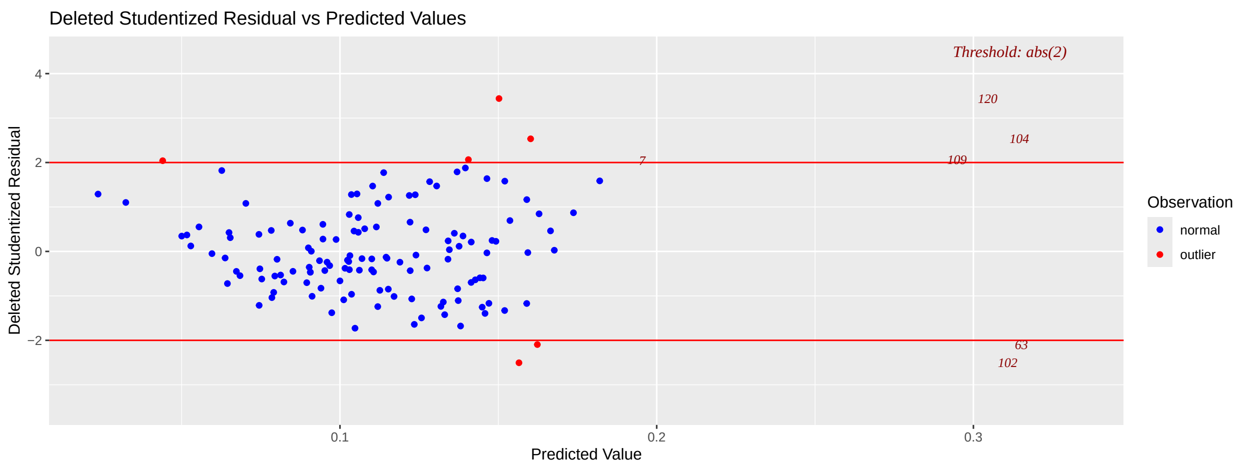



```
Call:
lm(formula = FI_Rate ~ Med_Income + Unemp_Rate + Pct_No_Vehicle +
    Poverty_Level, data = VA_FoodAccess)
```

```
Residuals:
      Min        1Q    Median        3Q       Max 
-0.029514 -0.007996 -0.001738  0.006513  0.039792 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)          1.046e-01  1.325e-02   7.891 1.20e-12 ***
Med_Income          -7.318e-07  8.426e-08  -8.685 1.59e-14 ***
Unemp_Rate           1.448e-02  2.013e-03   7.192 4.89e-11 ***
Pct_No_Vehicle       9.510e-08  5.230e-08   1.818  0.07137 .  
Poverty_LevelMedium  1.588e-03  4.659e-03   0.341  0.73383    
Poverty_LevelHigh    1.779e-02  5.871e-03   3.030  0.00296 ** 
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

Residual standard error: 0.01222 on 127 degrees of freedom
Multiple R-squared:  0.8777,    Adjusted R-squared:  0.8729 
F-statistic: 182.3 on 5 and 127 DF,  p-value: < 2.2e-16


        fit       lwr       upr
1 0.2081876 0.1837049 0.2326703


[1] 127


                   GVIF Df GVIF^(1/(2*Df))
Med_Income     3.872579  1        1.967887
Unemp_Rate     2.143185  1        1.463962
Pct_No_Vehicle 1.240989  1        1.113997
Poverty_Level  3.405290  2        1.358434


[1] 1.797194
```








```
       29         104        111        120        125
 0.1144383  0.4713359 -0.2434958  0.5697855  0.4483765
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.596918 -0.126797 -0.023407  0.008724  0.114438  0.787881
```

```
        RMSE  Rsquared         MAE
1 0.01235953 0.8732429 0.009953152
```