**US County Obesity Rates and Possible Linkages**

This project investigates US county obesity rates and possible economic and demographic linkages using 2016 data from the US Census Bureau, Bureau of Economic Analysis, and County Health Rankings & Roadmaps. All data are at the county level.

**The variables**

|  |
| --- |
| US Census Bureau data:   * SNAP Recipient (%) = (SNAP benefit recipients/population) \* 100 |
| Bureau of Economic Analysis data:   * Income Per Capita |
| County Health Rankings & Roadmaps, Robert Wood Johnson Foundation:   * Obesity (%) * Some College (%) * High School Graduate (%) * Unemployment (%) |

**Hypothesis**

Do obesity rates in US counties have strong correlations with demographical and economic characteristics such as SNAP benefits, education levels, income per capita, and unemployment rates?

**Analysis**

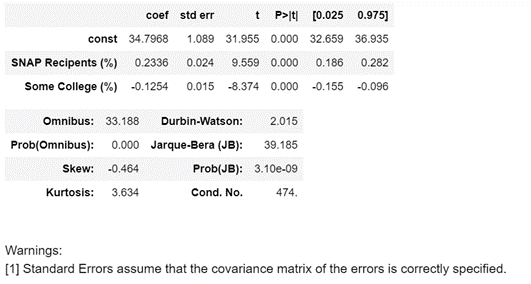
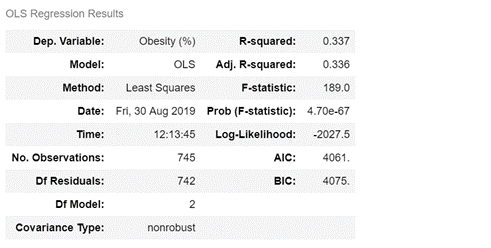
* **Scatter Plots and Trendlines**

The results of scatter plots and trendlines of the relationships of the obesity rates versus the other individual independent variables are as the following. The relationship of the obesity rates versus unemployment rates or SNAP recipient rates is a positive one. The relationship is negative between the obesity rates versus income per capita, some college or high school graduate percent. The high school graduate percent and obesity rate relationship is weaker comparing to the other relationships. The relationships form around the range of 20% - 40% obesity rates.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

* **Multiple Regression**

The results of the first multiple regression model show multicollinearity issues with positive sign on the coefficient of high school graduate rates. One or more variables are removed at times to resolve multicollinearity. The predictive model of obesity rate for this project has two predictors, SNAP recipient rate and some college percent. The significant F-test and P-value of the final model provides evidence for us to conclude that the obesity rate is significantly determined by SNAP recipient rate and some college percentage.



* **Correlation Matrix**

The results of correlation matrix support our aforementioned findings and conclude this project. The map reflects the findings using scatter plots and trendlines. It also shows correlations among independent variables which causes multicollinearity; e.g., positive strong correlations between unemployment rate and SNAP recipient rates and positive correlation between some college percent and income per capita.



**Conclusion**

According to the above findings, there is evidence to conclude that US county SNAP benefits, education levels, income per capita, and unemployment rates are possible linkages to US county obesity rates. Some linkages are positive whereas others are negative and/or weak. The implication, however, suggests against using all variables as predictors of the US county obesity rates together in one model since some of these predictors also have strong correlations among themselves. A model which resolves multicollinearity is presented. However, further tests of the model residuals beyond this project is highly recommended prior to using the predictive model. Exploratory factor analysis and a systematic algorithmic method for finding a suitable model would also be helpful.