### Income Inequality Multicollinearity Process

- The way we calculated Income-Inequality for each zipcode was doing the calculation 80th percentile – 20th percentile.

-When running the regression crimes2014 ~ IncomeInequality, we got an unexpected coefficient. This may be because we didn’t’ calculate income\_inequality correctly. Past researchers have used the Gini coefficient to calculate income inequality. Our motivation to use IncomeInequality is because past research has shown that this is a strong factor in predicting crime (insert author name here).

- For further research, we should use the Gini coefficient/index next time

-One problem with our method of calculating Income\_Inequality was because our method does not account for the distribution of the other incomes, it only looked at the 80th and 20th percentile point of the income. In addition, the 20th and 80th percentile was calculated within a range instead of the accurate pin-pointed 20th and 80th percentile (instead more details here, along with data source).

-In order to carry on with the research, we did not look at the Income Inequality variable

###based on my analysis in modeling.R, code lines 11 – 60, I conclude that only the variable PovertyLine is significant and adequate enough to capture the idea of how income affects crime. (Meaning adding other variables, such as MedIncome, Unemployment, etc. does not add more to the model if PovertyLine is already a variable in there).