## Best research questions = no matter what the answer, it’s interesting. Falsifiable. Why it didn’t work if it doesn’t work

## Dataset: World Development Database

**Question1:** Does a social safety net correlate with mean earnings? How is it correlated? What features are important?

**Question2:**  Is the amount of social safety services correlated with country debt?

**Question 3:** How does moving from coal/natural gas to electric sources impact a country’s vocations? Are threatened bird species lower in subsequent years after a country turns to electric solutions?

**Question 4:** What are factors that correlate with a relative low youth literacy rate?

**Question 5:** Are countries with the same main religion more similar? Countries geographically close to each other? Hot countries as opposed to cold countries? Or countries that have the same exports?

## Dataset: Kiva

**Question 1:** (Sentiment analysis and natural language processing) – what are the key reasons why teams become loaners? Are they also in certain lines of work (e.g. mostly churches)?

**Question 2:** What categories of work (e.g. food production, cereal) of each country have reaped the most benefit?

**Question 3:** How much on average do people loan? What are the characteristics of lenders who loan frequently vs those who loan infrequently?

## Dataset: Kaggle College ScoreCard

**Question1:** Does the income of students coming into college correlate with their income six years after college? (Since this is probably true) is there a certain group of colleges where this gap is lessened?

**Question2:** Are students from urban backgrounds more likely to have a higher SAT math score and SAT verbal score than students from rural backgrounds? Is this still true when variables of race and socio-economic status are taken out?

## FBI Crime Data

**Question 1:** Are crime rates higher in cities that have a significantly lower mean income than the city next to them?

**Question 2:** Are there differences between the types of crime committed in urban areas versus small towns versus rural areas? (What are the causes of this?)

## Boston Housing Data

**Question 1:** How much does location of the house matter in its valuation as opposed to number of rooms?

**Question 2:** What factors correlate with a house in a socio-economically diverse neighborhood as opposed to a socio-economically homogenous one?

## Bureau of Labor Statistics

**Question1:** In the past, America has experienced sudden technological disruptions that have shifted the nature of vocation. How does the percent of unemployment after our most recent disruption compare to the percent of unemployment in past disruptions?

**Question 2:** How likely is it that women who leave the workforce for more than a year will enter it again? Do they do so at the same salary?