**Potential Capstone Project Ideas – Lisa Ang**

**1. Do state taxes on cigarettes have an impact on cancer incidence?**

The deleterious effects of tobacco usage on human health have been well-documented since the 1950’s and it is currently the single greatest cause of preventable deaths worldwide. While the association between cigarette smoking and lung cancer is the strongest, smoking is known to increase the risk of contracting many types of cancer, particularly in the kidneys, head, neck, esophagus, pancreas and stomach among others.

Excise taxes on cigarettes in the United States have historically been viewed as a source of income generation for government at all levels, but support for increased taxation on tobacco products is now partly due to its observed effectiveness at decreasing smoking rates. Besides the federal tax rate of $1.01 per pack of 20 cigarettes, all states, the District of Columbia and the territories have implemented additional taxation on cigarettes, but at widely varying rates (from $0.17 in Missouri to $4.35 in New York.) While it is expected that higher taxes on cigarettes should reduce consumption, and thus overall smoking rates, does this correlate to a reduction in cancer incidence or do other confounding factors obscure any effects of this variance in taxation?

The Centers for Disease Control and Prevention (CDC) publishes data on cancer incidence and death by type, year and state, as well as behavioural data on smoking. State tax rates on cigarettes from 1950-present day are freely available. Comparing data from these sources, it is possible to determine the relationship between state tax rates, cigarette smoking and cancer rates.

**2. Is cancer incidence increasing or is diagnosis improving? The impact of Medicaid expansion under the Affordable Care Act on cancer incidence and mortality.**

The NIH National Cancer Institute reports that in 2016, an estimated 1,685,210 new cases of cancer will be diagnosed in the United States and 595,690 people will die from the disease. The most commonly-used indicator of progress in cancer treatment is the measurement of age-adjusted mortality rate. Other measures, such as incidence, are also important but it is not always straightforward to interpret changes in incidence rates. For example, is increased incidence reflecting a real increase in disease or has there been a change in the method of detection?

The Affordable Care Act was signed into law by President Obama in 2010 and significant reforms to health insurance standards and accessibility were enacted on Jan 1, 2014. Under the law, Medicaid eligibility was expanded to include individuals and families with incomes up to 133% of the federal poverty level but many states chose to opt out of the expansion. As people without insurance are less likely to seek medical care, it is expected that this expansion of insurance coverage ought to increase rates of cancer detection independently of actual disease rates. However, over half of the uninsured population resides in states that opted out of Medicaid expansion. By comparing incidence and mortality rates in states that did and did not expand Medicaid, this provides an opportunity to determine if changes in cancer incidence are due to changes in diagnosis or actual changes in disease prevalence.

**3. Do famous baseball players have an impact on baby naming trends?**

There are many websites devoted to tracking the popularity of baby names and it is even possible to even predict a person’s age based on their name. One enduring trend in baby naming over the years is the influence of pop-culture and celebrities on parents’ name choices for their newborns.

Kaggle has very kindly made available Major League Baseball’s complete batting and pitching statistics from 1871 to 2015. Comparing this to Social Security applications from Data.gov, it should be possible to determine if player popularity and batting ability has had an effect on the names given to newborn babies in the US over time.