Life Expectancy Trends in Alaska Counties (1982-2019)

Authors: Maxymilian Stefanski, Aditya Kutty, Jaemo Hyun

Overall Theory
- Life Expectancy in Alaska has increased over time but varies
- Independent Variable:
  - Year (Range from 1982- 2019)
  - Operationalization: Using the "year" column in the dataset
- Dependent Variable:
  - Life Expectancy
  - Operationalization: Using the column "ex" - is for the expected remaining years of life

Our Hypotheses
- Null Hypothesis: We would not expect there to be any relationship between the life expectancy and the year
- Alternative Hypothesis: We would expect the IV year to have a strong positive and direct correlation with the DV life expectancy

Datasets Used
- We used the US County Life tables created at Berkeley by Barbieri, Magali, Winant, and Celeste
- This dataset contains information about the average life expectancy, gender of various countries at various ages over a period of time for a number of states over a number of years

Proposed Control Variable
- One possible control variable would be the climate change. Climate change has been occurring in Alaska at quite a fast pace, and it can potentially have an impact on the life expectancy of Alaska. Due to our data starting in 1982 there have been metrics in place to see data regarding Climate Change and how it could impact life expectancy.

Descriptive Statistics
- This correlation however, had an R-squared value of 0.009313 which suggests low correlation between life expectancy over time
- This correlation between year and life expectancy had a p-value of 5.63×10^{-9}, which suggests had an extremely high significance
- The coefficient of this correlation was 2.973×10^{-2}, which suggests a very weak positive relationship between life expectancy and year

Conclusion
- The results from our regression analysis lead us to reject the alternative hypothesis and affirm the null hypothesis that there is no relationship between life expectancy and year
- A logical next step would be to explore climate data based on county to explore how environmental changes can impact life expectancy.


```
Residuals:
    Min      1Q  Median      3Q     Max
-19.5188 -9.4971  0.0228  9.4419 19.5273

Coefficients:
                    Estimate Std. Error  t value Pr(>|t|)
(Intercept)        1.999e+03  2.362e-01 8463.287  < 2e-16 ***
AK_b_county_lt$ex  2.973e-02  5.096e-03    5.833 5.63e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.95 on 8662 degrees of freedom
Multiple R-squared:  0.003913,  Adjusted R-squared:  0.003798
F-statistic: 34.03 on 1 and 8662 DF,  p-value: 5.63e-09
```
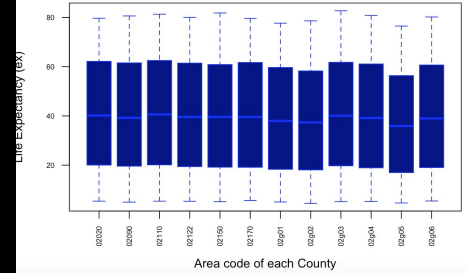
Scatter Plot with Line of Regression
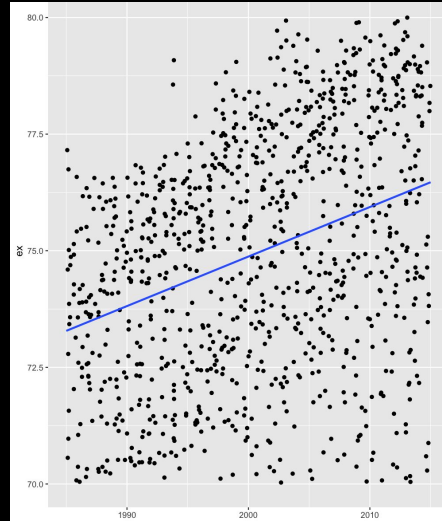
"Each county is defined by Area Code numbers"

Life Expectancy by County over average of 1982-2019

Life Expectancy by Year in Alaska Counties Average combined