Lung Cancer: Death & Demographics

Overview of Analysis:

Lung cancer is one of the six leading causes of death according to the World Health Organization (WHO). We are investigating mortality data on lung cancer in the United States, focusing on disparities within demographics to determine viable relationships and/or correlations between these variables.

Demographics include:
1. Race
2. Poverty Rate
3. Sex
4. Age
5. Geography

General Question:

Is there is a correlation between a person's demographics in the United States and their chances of dying of lung cancer?

Core Hypotheses:

1. There is a correlation between race and lung cancer mortality.

Statistical Analysis

Race & Lung Cancer Mortality

- **Hypothesis:**
    - **Null Hypothesis(H0):** There is no relationship between race and lung cancer mortality in 2015
    - **Alternative Hypothesis(H1):** There is a relationship between race and lung cancer mortality in 2015
- **Statistical Testing:** Chi-Squared Goodness of Fit Test
- **Results of Test**: The Chi-Squared value (10042.59) exceeded the critical value (16.27) with a p-value of .001, thus we reject the null hypothesis and accept the alternative hypothesis.
- **Conclusion:** There is a highly statistically significant relationship between race and lung cancer mortality in the United States in 2015.

Major Findings:

- Is there a trend in the number of cancer deaths across the United States?
    - **Results of Analysis:** As shown in the line graph, there has been an overall decrease in lung cancer mortality in the U.S. from 2006-2015.
- Is there a difference in number of cancer deaths between the sexes?

Lung Cancer: Death & Demographics

- ○ **Results of Analysis:** The bar chart shows that men die of lung cancer at a higher rate than women in the U.S. from 2006-2015. The number of deaths per year is declining for both, but at a higher rate for men.
- What is the relationship between age and lung cancer mortality?
  - ○ **Results of Analysis:** When viewing the number of deaths per age rage as in the provided bar chart, you can see that over 2006-2015 in the United States, most deaths occur when people are in their 70's, and no one under the age of 35 has died of lung cancer
- When taking into account lung cancer crude rate (death count/population *100,000), which state has the highest/lowest crude rate within 2015. Is there a difference when taking account population? If so, what can be causing this (such as demographics?)
  - ○ **Results of Analysis:** Results show within the bar graph, West Virginia has the highest lung cancer crude rate in comparison to the District of Columbia in 2015. This raises questions since population is included, what is causing these differences between the states when comparing states with larger populations to smaller population. Which is why we test demographics to see what is causing these differences?
- Is there a correlation/relationship between race and lung cancer mortality?
  - ○ **Results of Analysis:** Per the pie chart, the death rate of the white population of lung cancer in 2015 is proportionally higher than that of any other race. Per the statistical analysis above, there is a highly statistically significant relationship between race and lung cancer mortality in 2015.
- When taking account lung cancer crude rate, is there a correlation/relationship between poverty and lung cancer mortality on a county/state level.
  - ○ **Results of Analysis:** There is a very small R-squared value when you look at poverty and lung cancer mortality across both states and counties. There is a correlation, but it is very weak.

Sources

CDC: https://wonder.cdc.gov/
- Lung Cancer deaths broken down by:
  - ○ State/Counties
  - ○ Race
  - ○ Age
  - ○ Gender

FRED: https://fred.stlouisfed.org/
- Poverty Rate broken down by States & Counties