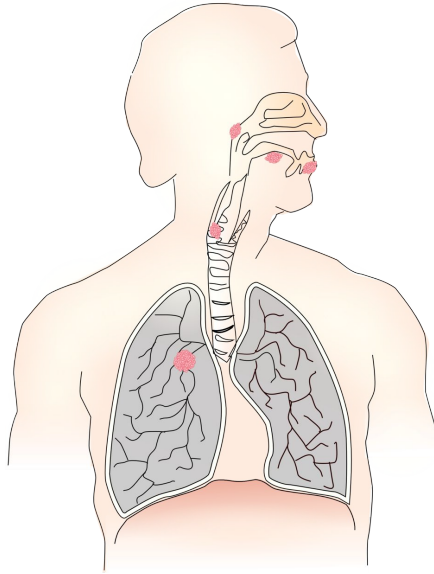


# **U.S. Chronic Disease Indicators: Lung and Bronchus Cancer**



By: Cassia Yoon, Faith Hall, Kyle Dalton, Thomas Keene

# Project Overview

The aim of our project is to uncover patterns of lung and bronchus cancer. We examined relationships between demographic categories (race/ethnicity and gender), place of residence, and diagnoses of cancer derived from the data.

Our data was pulled from The Centers for Disease Control website. Data was broken into 4 year time frames. In order to have a data set we could manage, we filtered the data using the most recent available time frame (2015-2019).

# Process and Project Visualizations

After downloading the CSV, we created a SQLite database, and hosted the database using Flask API so that we could access the data as a JSon.

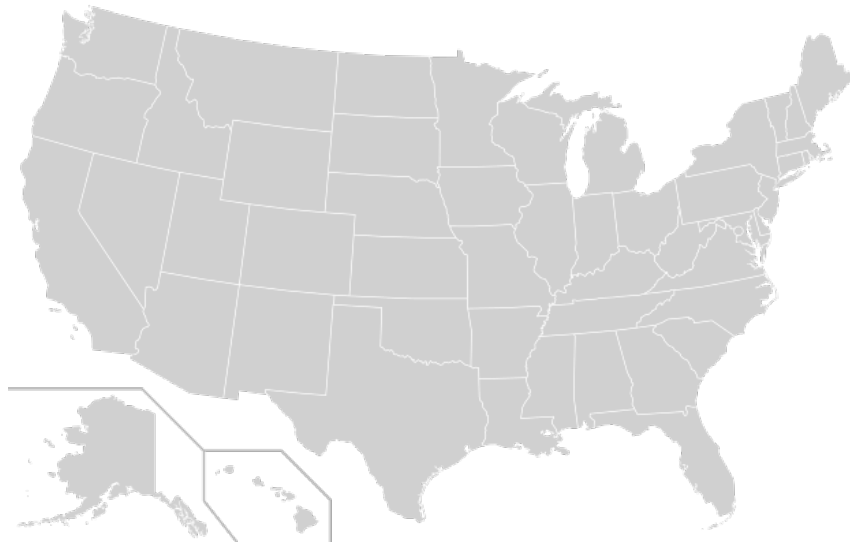
After filtering our data we created several interactive visualizations to help the user gain a quick understanding of the incidences of lung and bronchus cancer by state, race/ ethnicity and gender.

# Research Questions

- What are the incident trends per state?
- What are the incident trends by race/ethnicity?
- What are the incident trends by gender?

# Analysis 1

What are the incident trends per state?



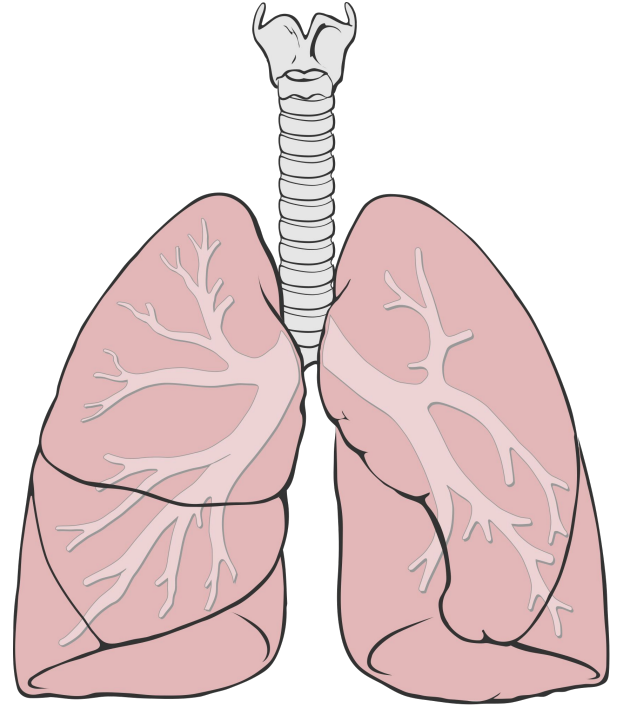
# Analysis 1 Conclusion

While this map does a good job of displaying the number of cases by state it does not necessarily tell you why one state may have more cases than another. Nor does it correlate the data to population.

Point to note, the cases displayed on the overall map are diagnosis per per 100,000. So, in the case of Minnesota we have 3777 cases per 100,000 roughly equal to 213,023 cases between 2015-2019.

# Analysis 2

What are the trends by race/ethnicity?



# Analysis 2 Conclusion

- The data for the United States (overall) shows that the highest incidence of lung and bronchus cancer is recorded for the ethnicity category “White, non-Hispanic”.
- In descending order, the lung and bronchus incidence rates for the United States (overall) are as follows:
  - White, non-Hispanic
  - Black, non-Hispanic
  - Hispanic
  - Asian or Pacific Islander
  - American Indian or Alaska Native



## Analysis 2 Conclusion

- While most states reported data for all ethnicities, several states only reported data for the ethnicities in which data had been collected over the specified time period.
- As a result, it is somewhat difficult to draw a complete picture of the state-to-state variation of cancer incidence within the U.S.
- For the states that did report data for all ethnicities, the trends tend to mimic the general ethnic population breakdown of the state.
- It is noted that some states (such as Illinois) did not report data by ethnicity at all, and instead just reported overall incidence data stratified by gender only.

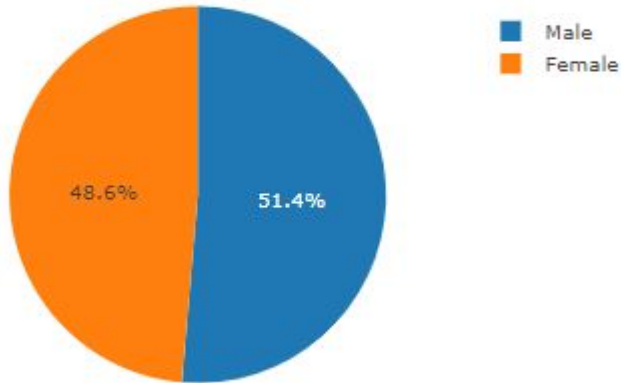
## Analysis 3

What are the incident trends by gender?

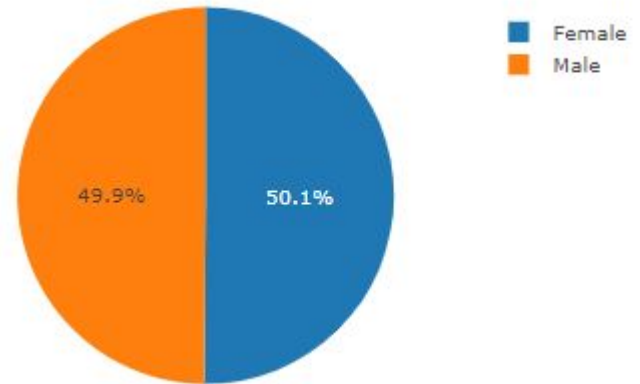


# Analysis 3 Conclusion

Lung Cancer cases across America  
from 2015-2019 by gender



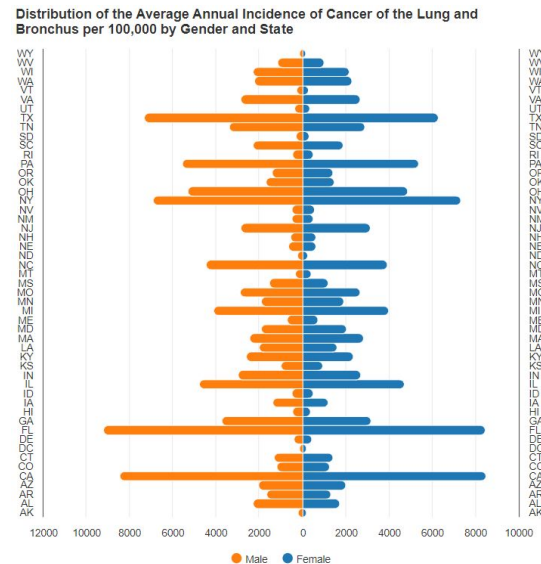
Lung Cancer cases in Minnesota  
from 2015-2019 by gender



## Analysis 4: Gender continued

Is there a significant difference in lung cancer incidence between males and females?

- Negative stacked bar chart
  - Shows overall picture of male and female average annual incidence by state.
- Used a new Javascript library (Highcharts)



# Analysis 4 Conclusions

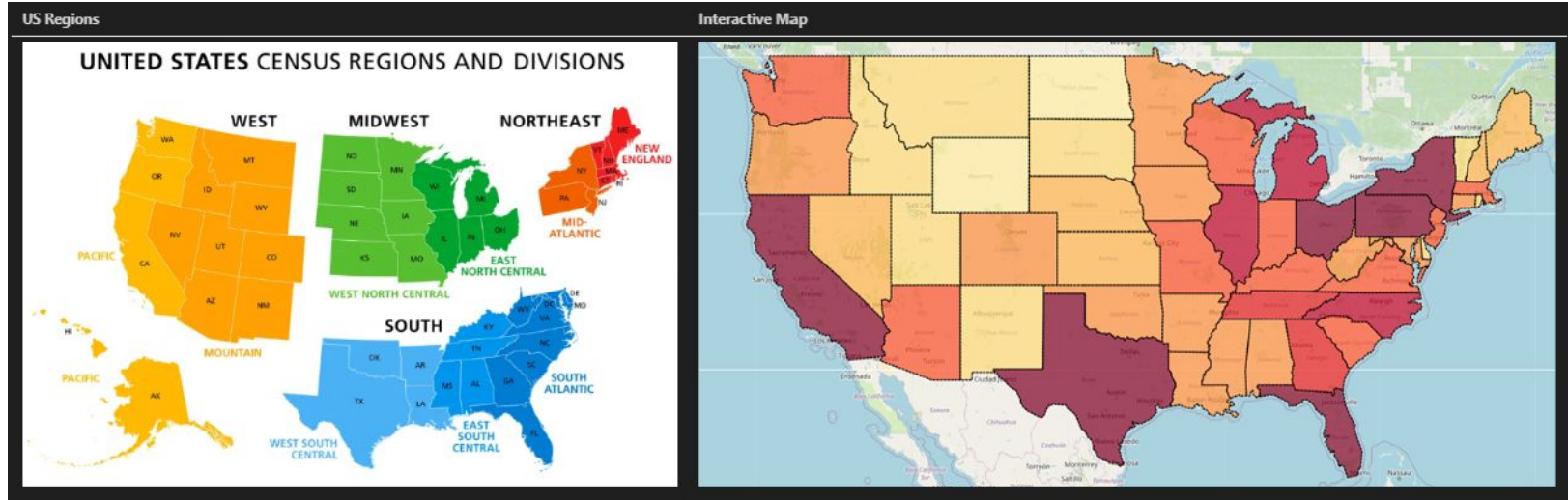
- The negative stacked bar chart showed even distributions between genders across the states.
- Independent T-Test:
  - $pvalue=0.7792577059693152$
  - $Pvalue > 0.05$
  - Therefore, the difference between the male and female data samples is not significant.

```
stats.ttest_ind(female_samples, male_samples, equal_var=False)
```

✓ 0.0s

```
Ttest_indResult(statistic=-0.2810449172608385, pvalue=0.7792577059693152)
```

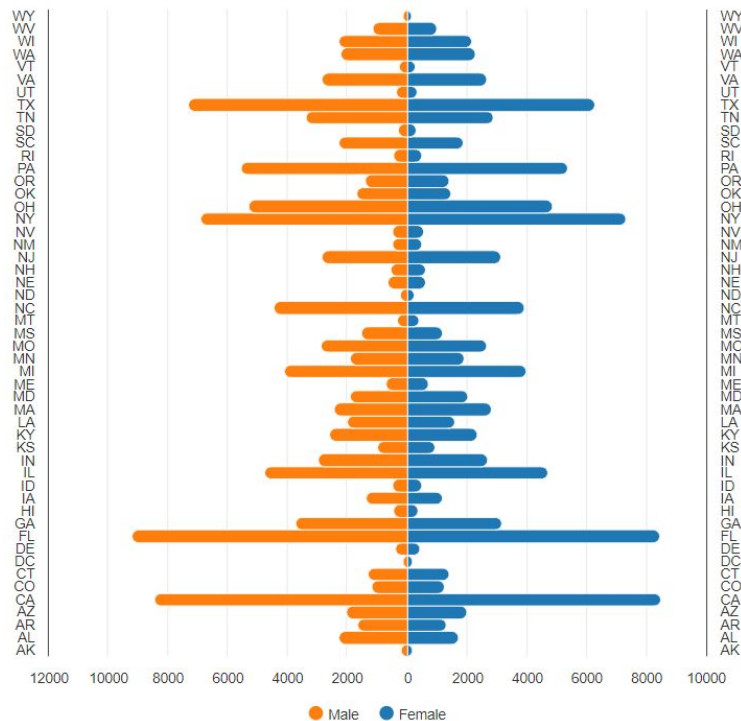
# General Conclusions



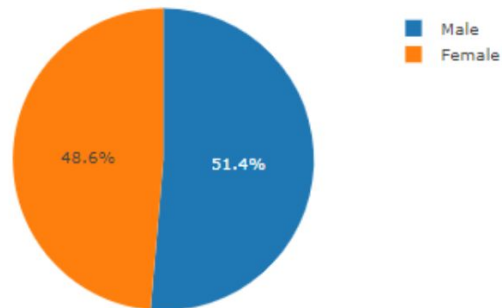
Lung cancer appears to be most prevalent near major bodies of water (The Pacific and Atlantic Oceans, and the Great Lakes region) except for the New England Region.

# General Conclusions

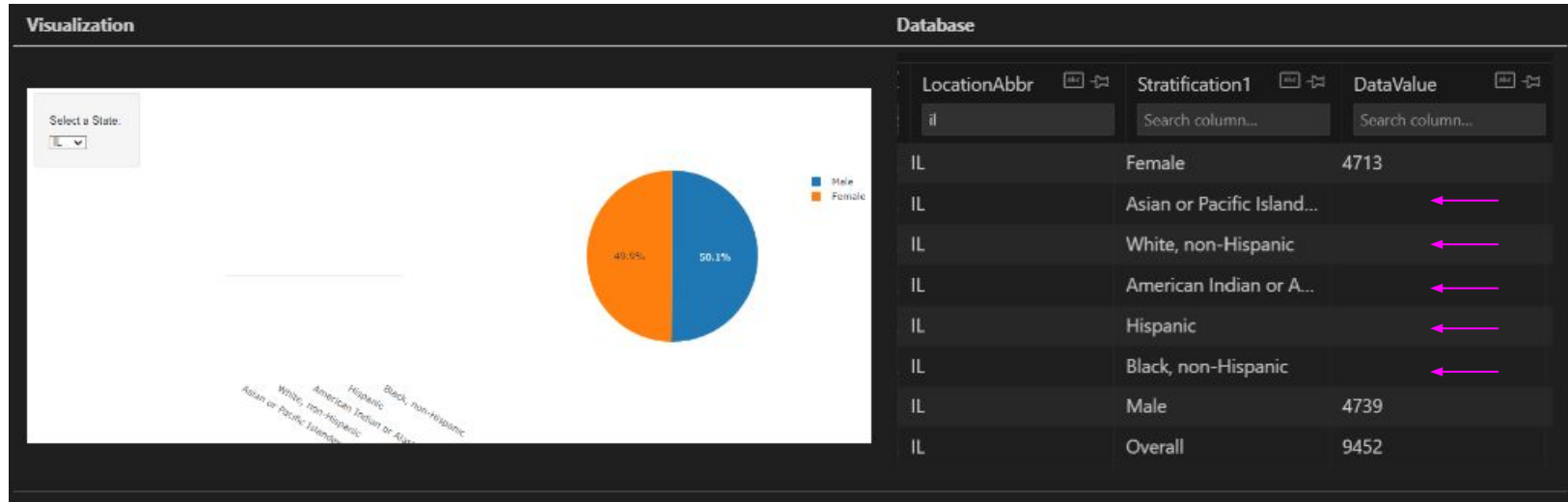
Distribution of the Average Annual Incidence of Cancer of the Lung and Bronchus per 100,000 by Gender and State



Statistical analysis showed no significant difference between male and female incidence of cancer.



# General Conclusions



There are limitations in our data due to missing values. For the future, it would be interesting to cross-reference this cancer dataset with US census data (ethnicity/race, gender, and population), climate data and air-quality data.



# Acknowledgements



We wish to thank our teaching staff:

- Hunter Hollis
- Sam Espe
- Randy Sendek

# Resources

- Cancer dataset:

<https://data.cdc.gov/Chronic-Disease-Indicators/U-S-Chronic-Disease-Indicators-Cancer/u9ek-bct3>

- Visualization example of the above dataset:

[https://nccd.cdc.gov/cdi/rdPage.aspx?rdReport=DPH\\_CDI.ExploreByTopic&isTopic=CAN&isYear=9999&go=GO](https://nccd.cdc.gov/cdi/rdPage.aspx?rdReport=DPH_CDI.ExploreByTopic&isTopic=CAN&isYear=9999&go=GO)