Team 4. ANALYTICS PROJECT:

POTENTIAL FACTORS ON US CANCER RATES

https://github.com/cassidyschul/American_cancer_stats.git

TEAM MEMBERS

- Abel Dumecha

- Cassidy Schultheis

- Jessica Snowden

- Mina Agyen

HYPOTHESIS

- Cancer rates are increasing--both for the number of incidents as well as the mortality rates.

- **Gender** impacts cancer rates and types of cancer
- **Physical location** impacts cancer rates
- **Age** impacts cancer rates
- **Race** impacts cancer rates and types of cancer

DATA SELECTION

Centers for Disease Control's US Cancer Statistics (1999 - 2020)

https://www.cdc.gov/cancer/uscs/dataviz/download_data.htm

Variables Selected

- Age
- Count (number of individuals with cancer)
- Event Type (i.e. Incidence and Mortality)
- Population of reporting area (i.e. State)
- Race

- Rate
- Sex
- Site (i.e. Type of Cancer)
- Year

HYPOTHESIS 1

CANCER RATES ARE INCREASING—
BOTH FOR THE NUMBER OF INCIDENTS AS WELL
AS THE MORTALITY RATES.

HYPOTHESIS

CANCER RATES ARE INCREASING--BOTH FOR THE NUMBER OF INCIDENTS AS WELL AS THE MORTALITY RATES.

<u>Null-Hypothesis</u>

Cancer rates are not changing year over year.

<u>Alternative Hypothesis</u>
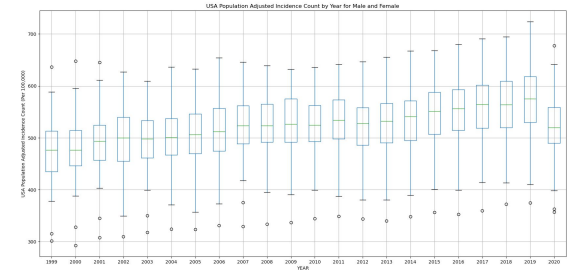
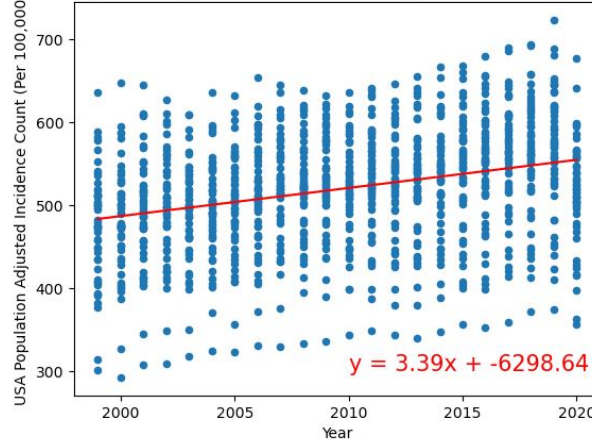Cancer rates are changing year over year.

<u>Visualization:</u> USA Population Adjusted Incidence Count (Per 100,000) by Year for Females and Males

- Includes all races, ages, and states

Observations:

- Incidence rates are slightly increasing year over year

- This may be be to new detection methods and technologies

- Drop in 2020 due to COVID

- The correlation between year and incidence count is 0.31 indicating a very weak positive correlation

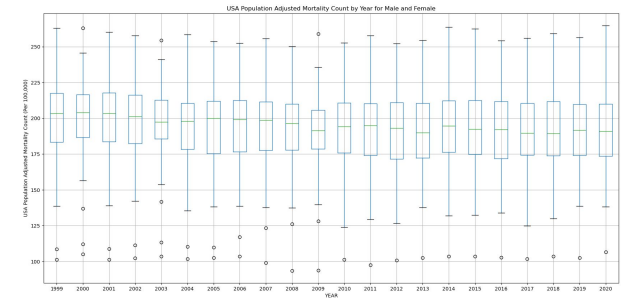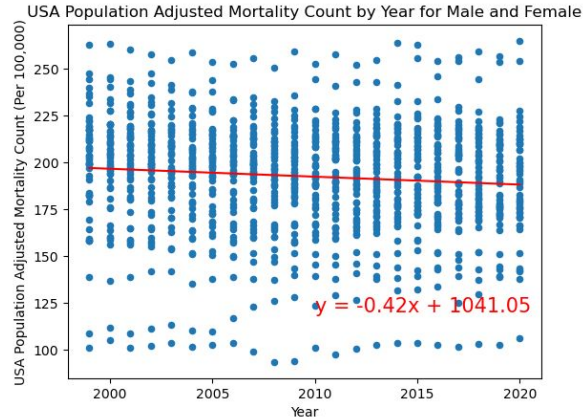USA Population Adjusted Incidence Count by Year for Male and Female

$y = 3.39x + -6298.64$

USA Population Adjusted Incidence Count by Year for Male and Female

Visualization: USA Population Adjusted Mortality Count (Per 100,000) by Year for Females and Males

- Includes all races, ages, and states

Observations:

- Mortality rates are slightly decreasing

- The correlation between year and incidence count is 0.09 including either no correlation or an extremely weak negative correlation

- This may be due to new research treatments available

USA Population Adjusted Mortality Count by Year for Male and Female

y = -0.42x + 1041.05

USA Population Adjusted Mortality Count by Year for Male and Female

HYPOTHESIS 1 - CONCLUSION

HYPOTHESIS IS TRUE

CANCER RATES SLIGHTLY INCREASING WHILE MORTALITY RATES HAVE NO CHANGE OR ARE SLIGHTLY DECREASING BASED ON BOTH BOX PLOTS, SCATTER PLOTS WITH REGRESSION LINES AND CORRELATION COEFFICIENTS

HYPOTHESIS 2

GENDER DOES HAVE AN IMPACT ON CANCER
RATES AND TYPE

HYPOTHESIS

GENDER DOES HAVE AN EFFECT ON CANCER RATES-BOTH FOR THE NUMBER OF INCIDENTS AS WELL AS THE MORTALITY RATES.

<u>Null-Hypothesis</u>

Gender does not have an effect on cancer rates.

<u>Alternative Hypothesis</u>
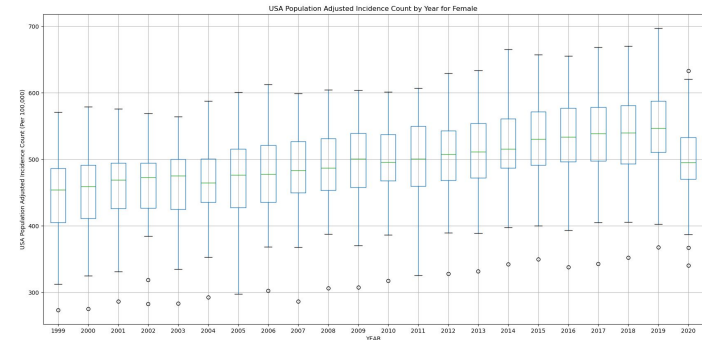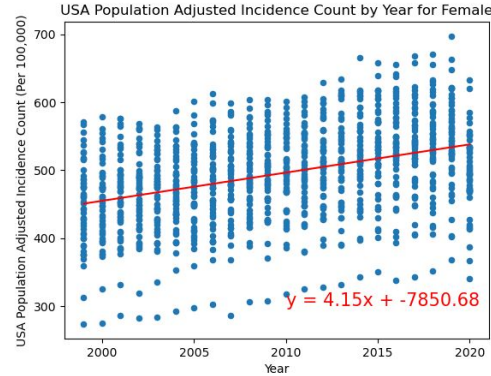
Gender does have an effect on cancer rates.

<u>Visualization:</u> USA Population Adjusted Incidence Count (Per 100,000) by Year for Females

- Includes all races, ages, and states

Observations:

- Incidence rates are slightly increasing year over year for females

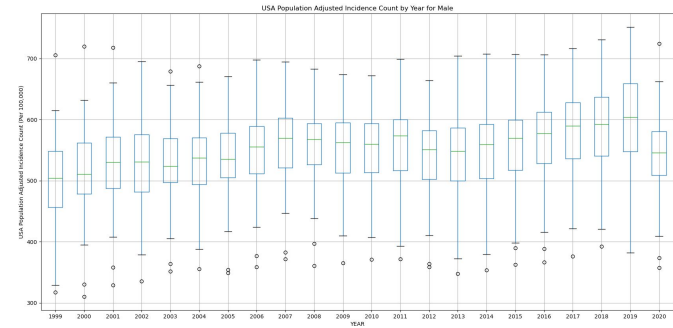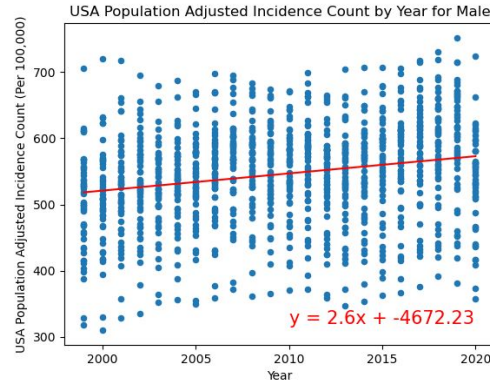- The correlation between year and incidence count is 0.39 indicating a weak positive correlation

USA Population Adjusted Incidence Count by Year for Female

$y = 4.15x + -7850.68$

USA Population Adjusted Incidence Count by Year for Female

Visualization: USA Population Adjusted Incidence Count (Per 100,000) by Year for Males

- Includes all races, ages, and states

Observations:

- Incidence rates are slightly increasing year over year for males

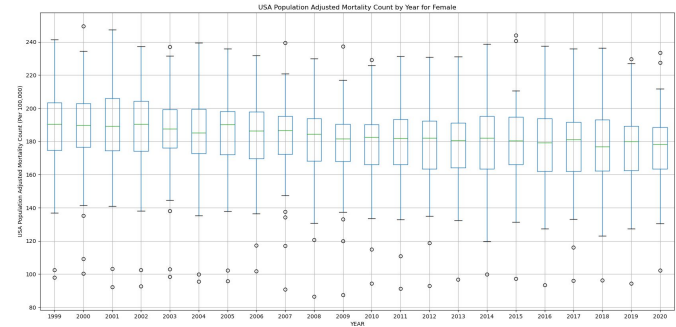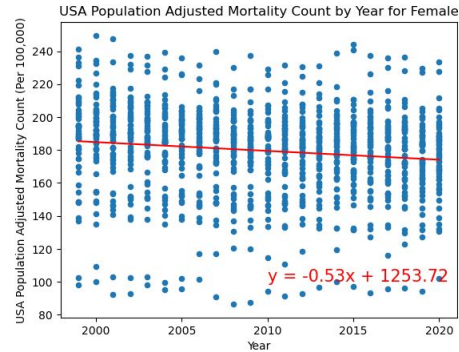- The correlation between year and incidence count is 0.22 indicating a very weak positive correlation

USA Population Adjusted Incidence Count by Year for Male

$y = 2.6x + -4672.23$

USA Population Adjusted Incidence Count by Year for Male

Visualization: USA Population Adjusted Mortality Count (Per 100,000) by Year for Females

- Includes all races, ages, and states

Observations:

- Mortality rates are slightly decreasing year over year for females

- The correlation between year and incidence count is 0.13 indicating a very weak negative correlation

USA Population Adjusted Mortality Count by Year for Female

$y = -0.53x + 1253.72$

USA Population Adjusted Mortality Count by Year for Female

Visualization: USA Population Adjusted Mortality Count (Per 100,000) by Year for Males

- Includes all races, ages, and states

Observations:

- Mortality rates are slightly decreasing year over year for males

- The correlation between year and incidence count is 0.06 indicating no or a very weak negative correlation

USA Population Adjusted Mortality Count by Year for Male

$y = -0.32x + 840.2$

USA Population Adjusted Mortality Count by Year for Male
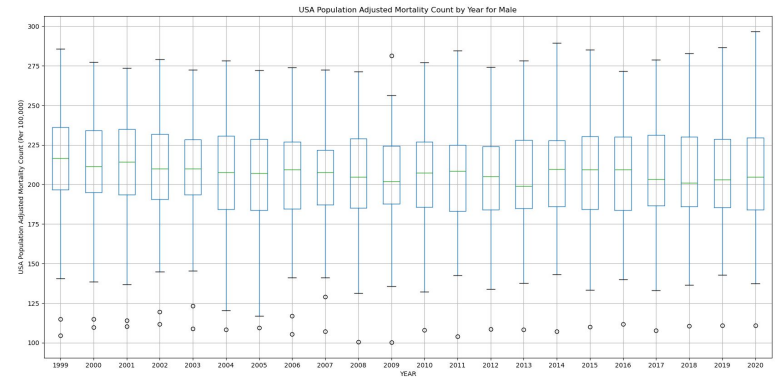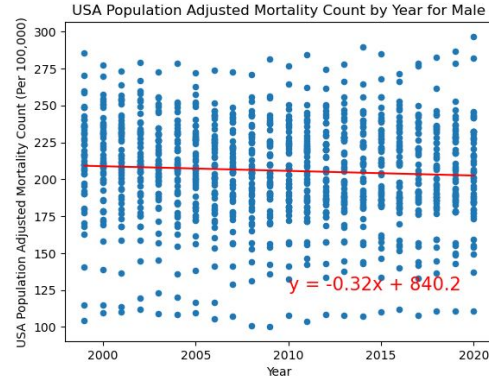
<u>Visualization:</u> Average Incident Count for Each Cancer Site Include Males and Females

- Includes all races, ages, and states

Observations:

- T statistic=-0.341,

- p value=0.735

- This indicates there is not a significant difference between female and male incidence averages for each cancer site and the null hypothesis should be accepted


Average Incidence Count For Each Cancer Site Including Females

Average Incidence Count For Each Cancer Site Including Males
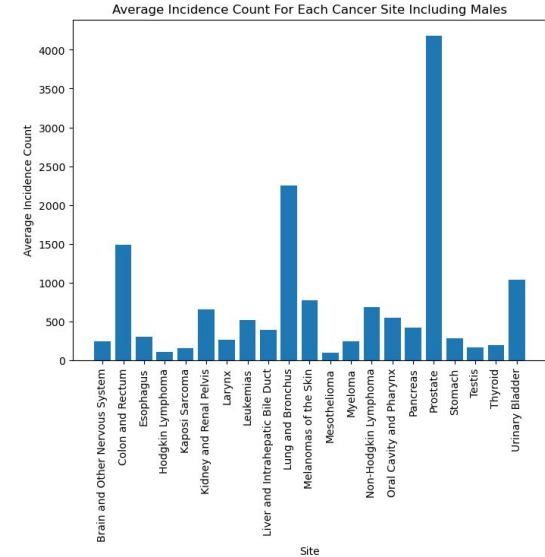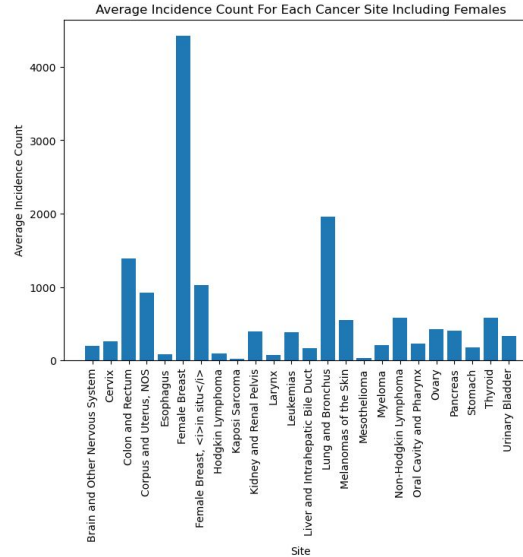
Visualization: Average Mortality
Count for Each Cancer Site Include
Males and Females

- Includes all races, ages, and states

Observations:

- T statistic=-0.417

- p value=0.679

- This indicates there is not a
significant difference between
female and male mortality
averages for each cancer site and
the null hypothesis should be
accepted
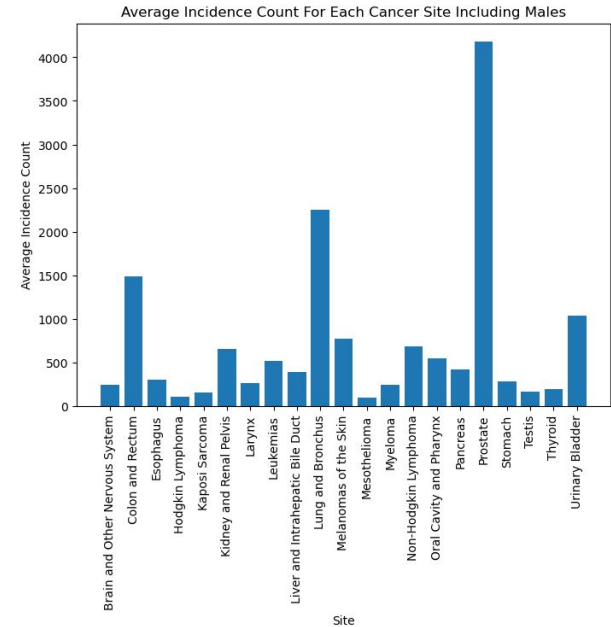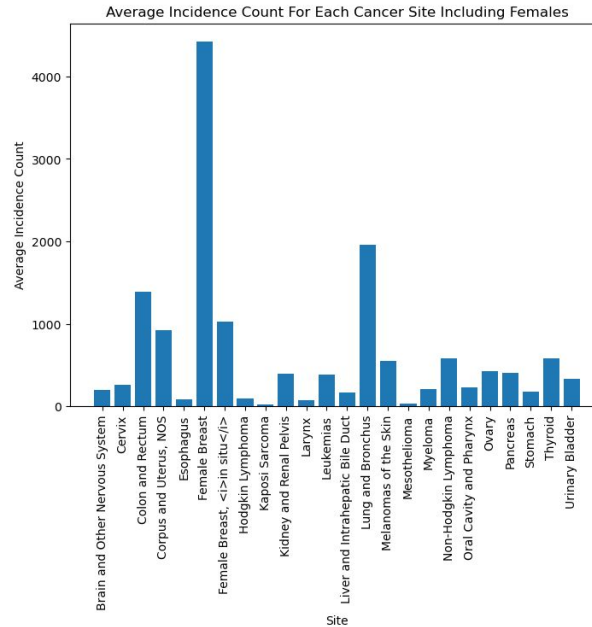
HYPOTHESIS IS FALSE

GENDER DOES NOT HAVE AN EFFECT ON CANCER RATES BASED ON THE BOX PLOTS, SCATTER PLOTS WITH REGRESSION LINES AND WELCH'S T-TEST VALUES

HYPOTHESIS 3

DOES LOCATION HAVE AN IMPACT ON CANCER RATES?

HYPOTHESIS

LOCATION DOES HAVE AN EFFECT-BOTH FOR THE NUMBER OF INCIDENTS AS WELL AS THE MORTALITY RATES.

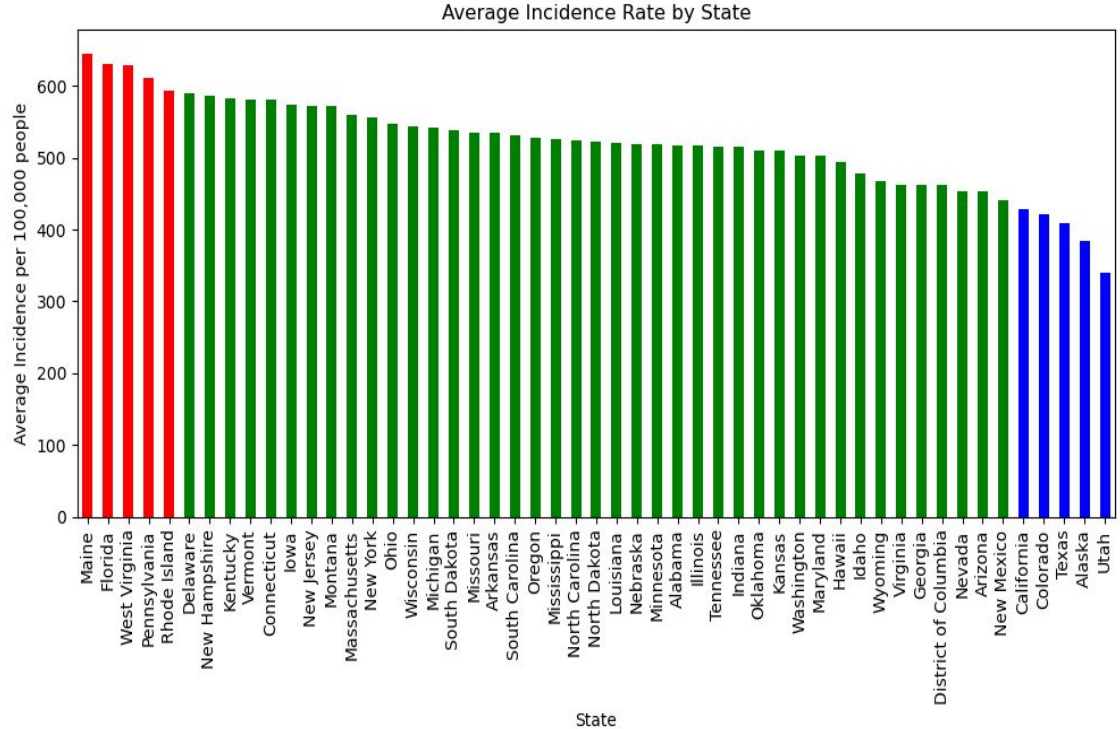| Null-Hypothesis	Alternative Hypothesis
Location does not have an effect on cancer rates.	Location does have an effect on cancer rates.

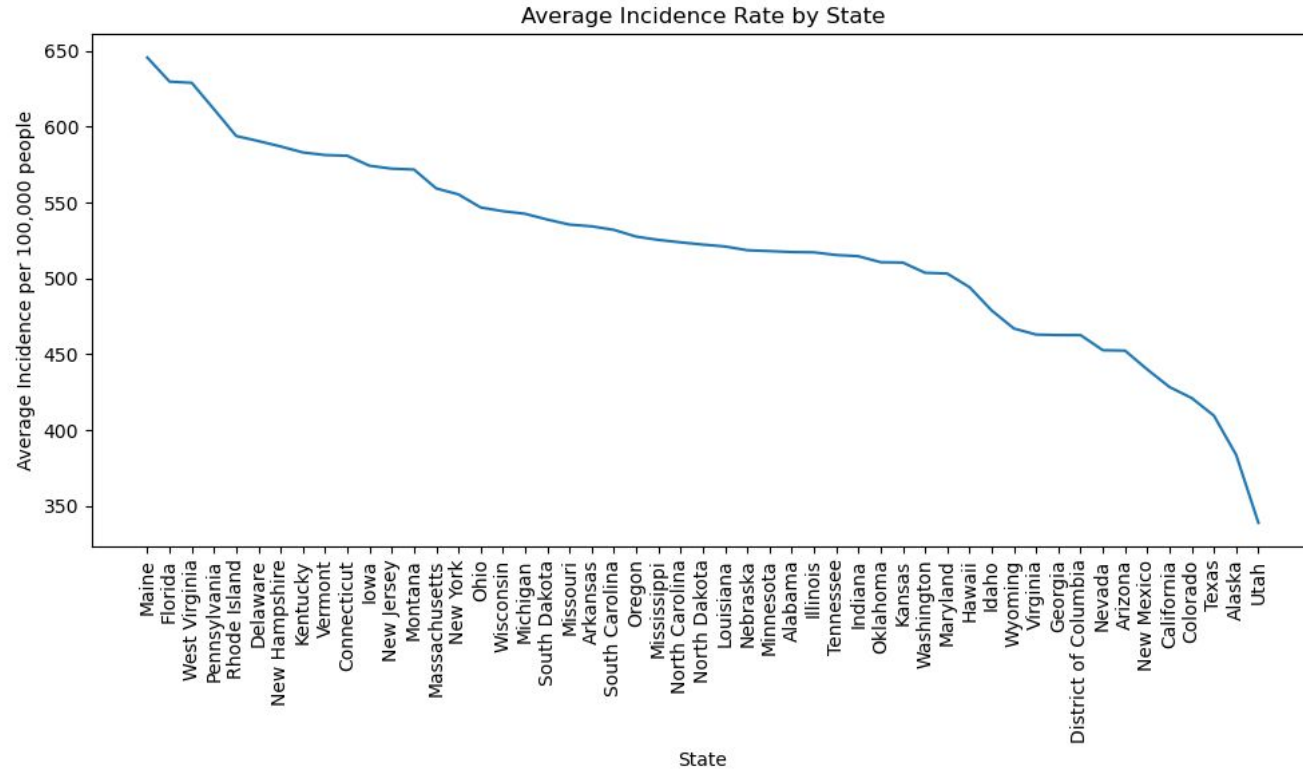Visualization: Average incidence rate per 100,000 ppl for all cancer types in each state

• Includes all races, ages, and states

Observations:

• People living in Maine, Florida, West Virginia, Pennsylvania, and Rhode Island have a high tendency of exposure to cancer compared to those living in California, Colorado, Texas, Alaska, and Utah.


Average Incidence Rate by State

Average Incidence Rate by State

Visualization: Average mortality rate per 100,000 ppl for all cancer types in each state
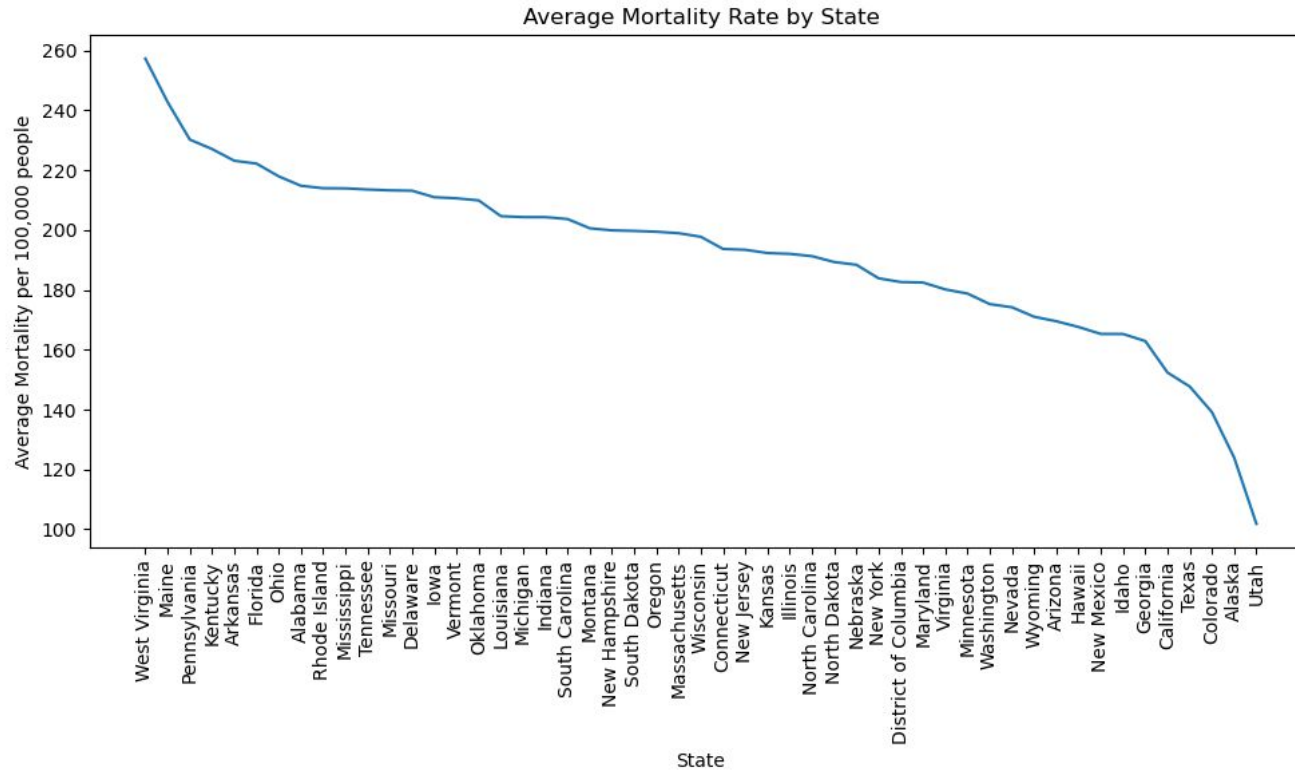
- Includes all races, ages, and states

Observations:

- People living in West Virginia, Maine, Pennsylvania, Kentucky, and Arkansas have a high tendency of mortality by cancer compared to those living in California, Texas, Colorado, Alaska, and Utah.

Average Incidence Rate by State

Average Mortality Rate by State
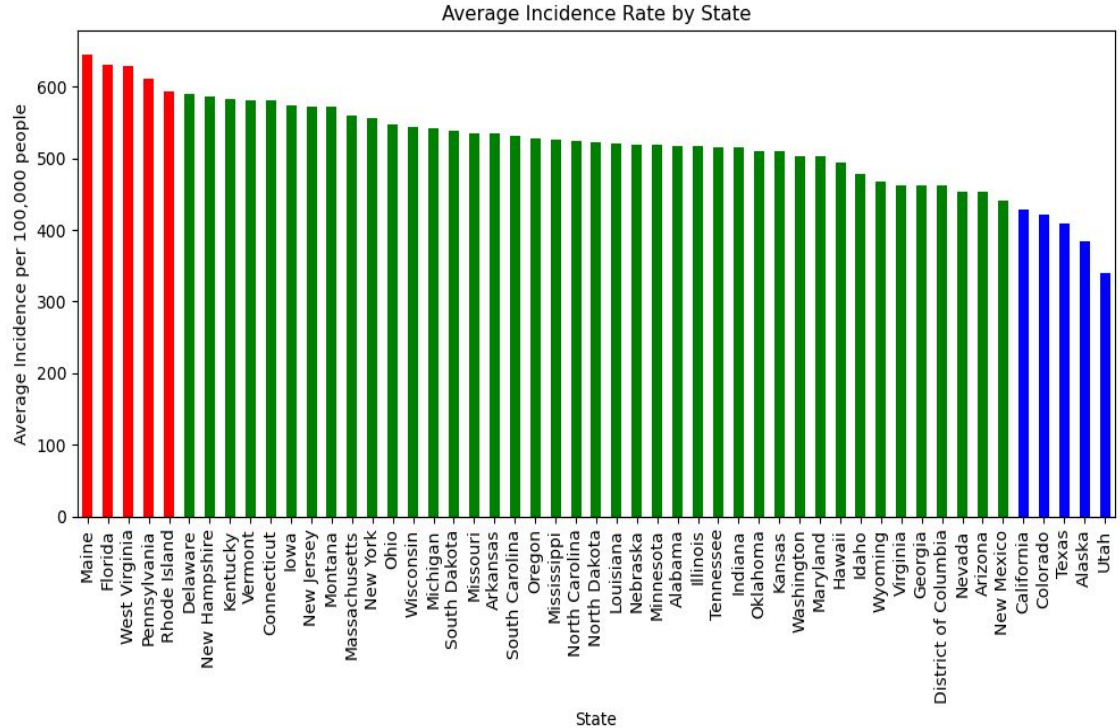
HYPOTHESIS 3 - CONCLUSION

HYPOTHESIS IS TRUE

LOCATION DOES HAVE AN EFFECT ON CANCER RATES BASED ON THE
BAR and LINE PLOT

HYPOTHESIS 4

- The mortality rate for older generation seems to be increasing compare to the younger generation.

HYPOTHESIS

AGE DOES HAVE AN EFFECT ON CANCER RATES-BOTH FOR THE NUMBER OF INCIDENTS AS WELL AS THE MORTALITY RATES.

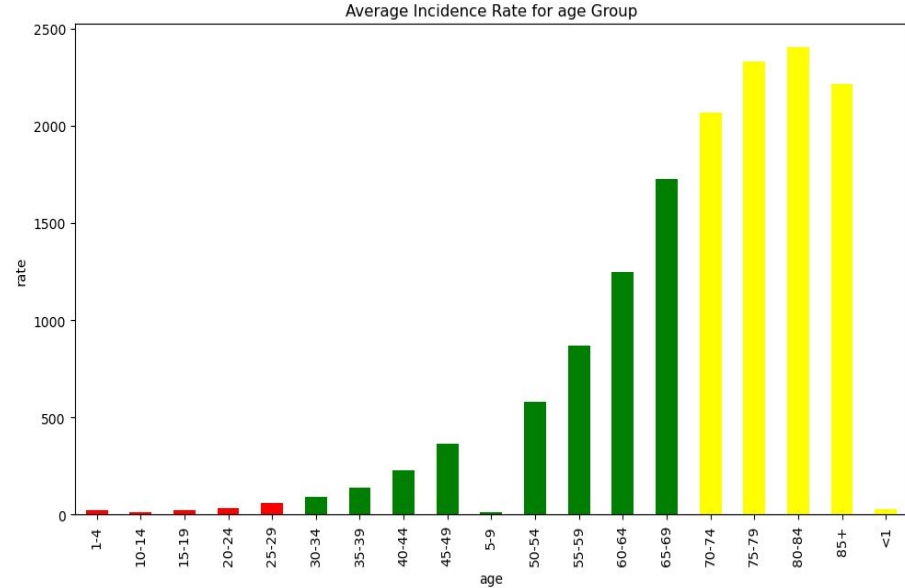| Null-Hypothesis	Alternative Hypothesis
Age does not have an effect on cancer rates.	Age does have an effect on cancer rates.

Visualization:

- Average Cancer rate by age

Observations:

- Cancer rates across cancer types increased as age groups got older.
- As individuals age, the cancer rate increases.

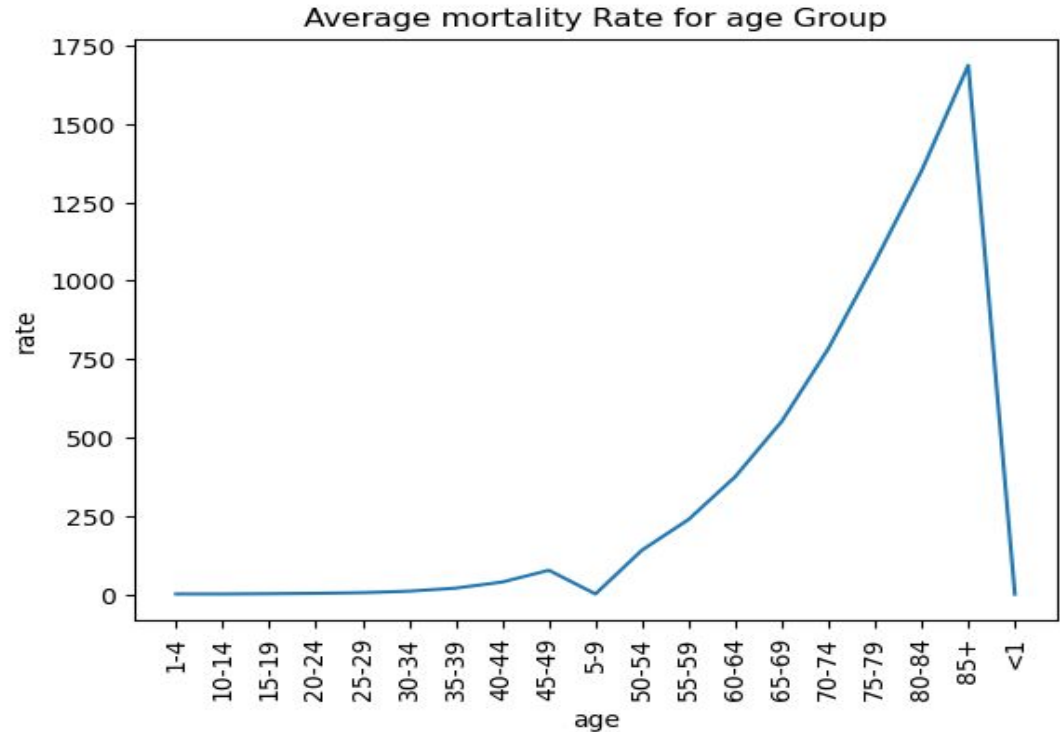
Average Incidence Rate for age Group

CANCER AGE IMPACT

Visualization:

- Average mortality rate for age group

Observations:

- The mortality rate for older generation seems to be increasing compare to the younger generation.

Average mortality Rate for age Group

HYPOTHESIS 4 - CONCLUSION

HYPOTHESIS IS TRUE

AGE DOES HAVE AN EFFECT ON CANCER RATES BASED ON

- Often our body repairs the damage, or damaged cells self-destruct. But sometimes the damage in a cell builds up, and the cell starts to behave differently. This is when it can grow out of control and develop into a cancer.

Research by:

Cancer Research UK

HYPOTHESIS 5

DOES RACE IMPACT CANCER TYPES?

HYPOTHESIS

RACE DOES HAVE AN EFFECT ON CANCER RATES-BOTH FOR THE NUMBER OF INCIDENTS AS WELL AS THE MORTALITY RATES.

| Null-Hypothesis	Alternative Hypothesis
Race does not have an effect on cancer rates.	Race does have an effect on cancer rates.

Visualization:

- Average Cancer rate by race depending on the count/ number of individuals in a given race.

Observations:

- Cancer rates across cancer types varied amongst races non-hispanic white, hispanic of any race, and non hispanic black based off the given data.

- The higher amount of people within a given race the more cases of cancer were found.

Cancer Race Research
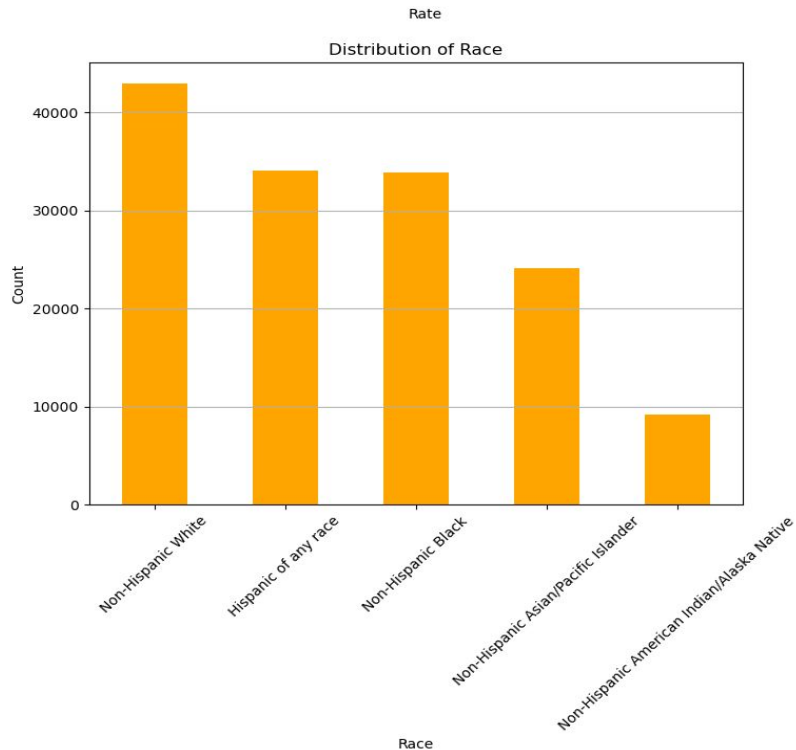
HYPOTHESIS 4 - CONCLUSION

HYPOTHESIS IS TRUE

RACE DOES HAVE AN EFFECT ON CANCER RATES BASED ON

Thank you!