

Group 5

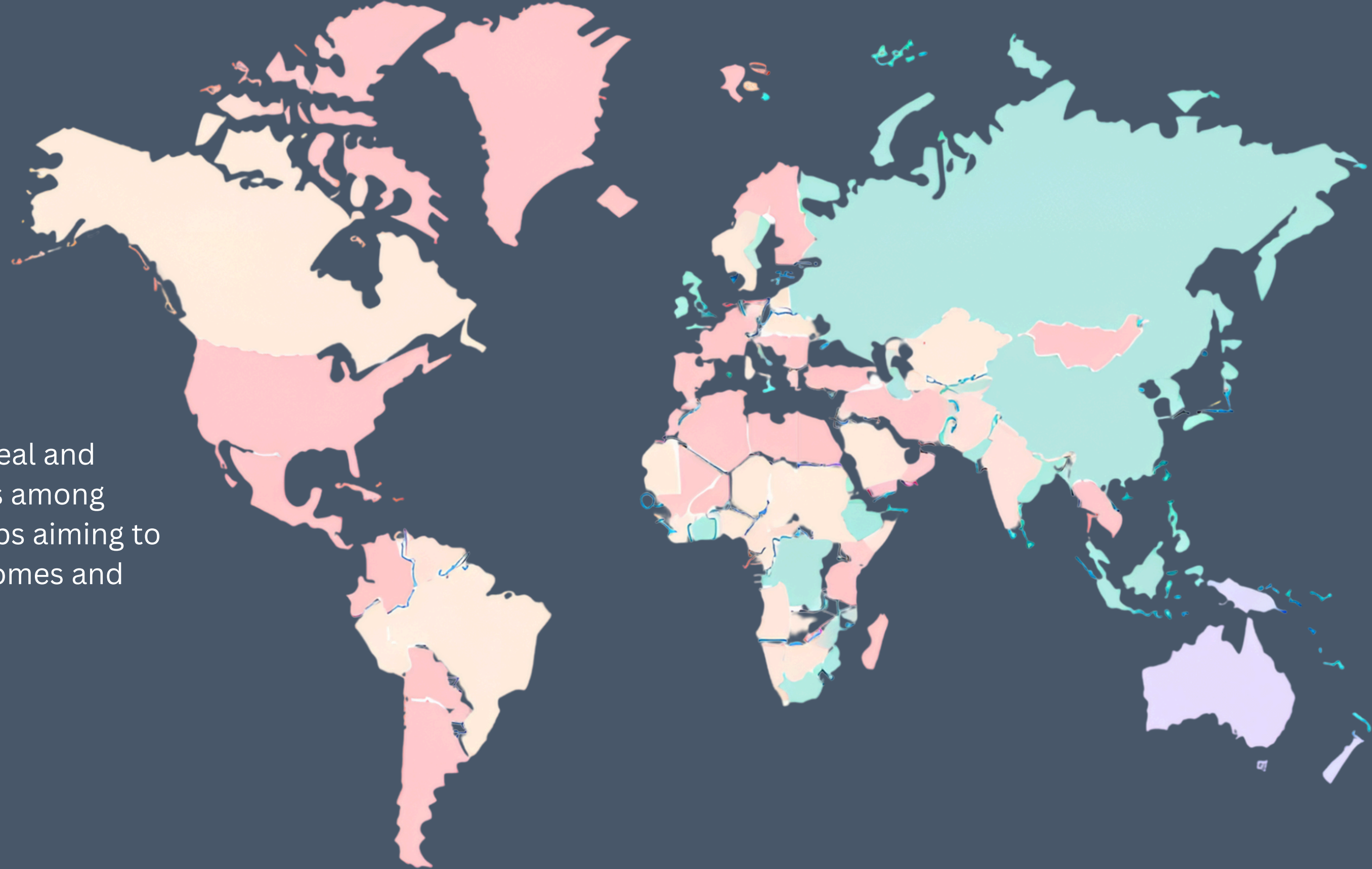
FINAL GROUP PROJECT STORYBOARD

Khadija Arslan, Yu Xue, Sanah Aulakh, Adam Hallal, Alana Yang, Diego Huezo, Christina Jiang

Cancer Rates Comparison Across Race & Geography

The mission of the project is to reveal and address differences in cancer rates among various racial and geographic groups aiming to provide equitable healthcare outcomes and better politics.

[Learn more](#)



About Our Website

This project explores the variation in cancer rates across different racial groups and geographical regions, with a specific focus on understanding the underlying factors contributing to these disparities. By analyzing data from various reliable sources, we aim to highlight significant differences in cancer incidence and mortality rates among different communities. Our goal is to uncover healthcare inequalities and provide insights for targeted interventions and policy changes. Using digital humanities tools, we aim to present this critical information in an accessible and engaging manner, ultimately striving to promote health equity and social justice.

Timeline of Major Events in Cancer Study in the Past Century



1924-1944

-- Discovery of the Pap smear by Dr. George Papanicolaou, leading to early detection of cervical cancer.



1945-1964

--Discovery of the double-helix structure of DNA by James Watson and Francis Crick, providing insight into genetic mutations in cancer.
-- Development of the Papanicolaou test (Pap test) for cervical cancer screening.



1965-1984

--Introduction of combination chemotherapy for the treatment of Hodgkin's lymphoma by Dr. Vincent DeVita.
--President Richard Nixon signs the National Cancer Act, initiating the "War on Cancer."
--Development of monoclonal antibodies by César Milstein and Georges Köhler.
--Approval of interferon-alpha, the first biological therapy for cancer treatment.



1985-2004

--Identification of the HER2/neu gene associated with aggressive breast cancer.
--Approval of paclitaxel (Taxol) for the treatment of ovarian cancer.
--Approval of trastuzumab (Herceptin) for HER2-positive breast cancer.
--Approval of imatinib (Gleevec) for chronic myeloid leukemia.
--Completion of the Human Genome Project.



2005-2024

--Approval of the first HPV vaccine, Gardasil.
--The Cancer Genome Atlas Program is launched.
--Approval of ipilimumab (Yervoy) for metastatic melanoma.
--Approval of CAR-T cell therapy for certain types of leukemia.
--Using the CRISPR-Cas9 gene-editing technology.
--Development of mRNA vaccines for the first.
--Advances in liquid biopsy technology, allowing for non-invasive cancer detection and monitoring.

Timeline of Historical Disparities in Cancer Study in the Past Century



Early 20th Century

- Before that, cancers were attributed to supernatural causes.
- Writers in this period started to address cancer as an unique disease.
- For women, cultural taboos around discussing breast health meant they did not seek medical help early.



Mid-Late 20th Century

- The rise of breast cancer awareness campaigns, such as those led by Susan G. Komen (i.e. Breast Cancer Foundation), began to address some disparities by promoting education, early detection, and fundraising for research.



Present Day

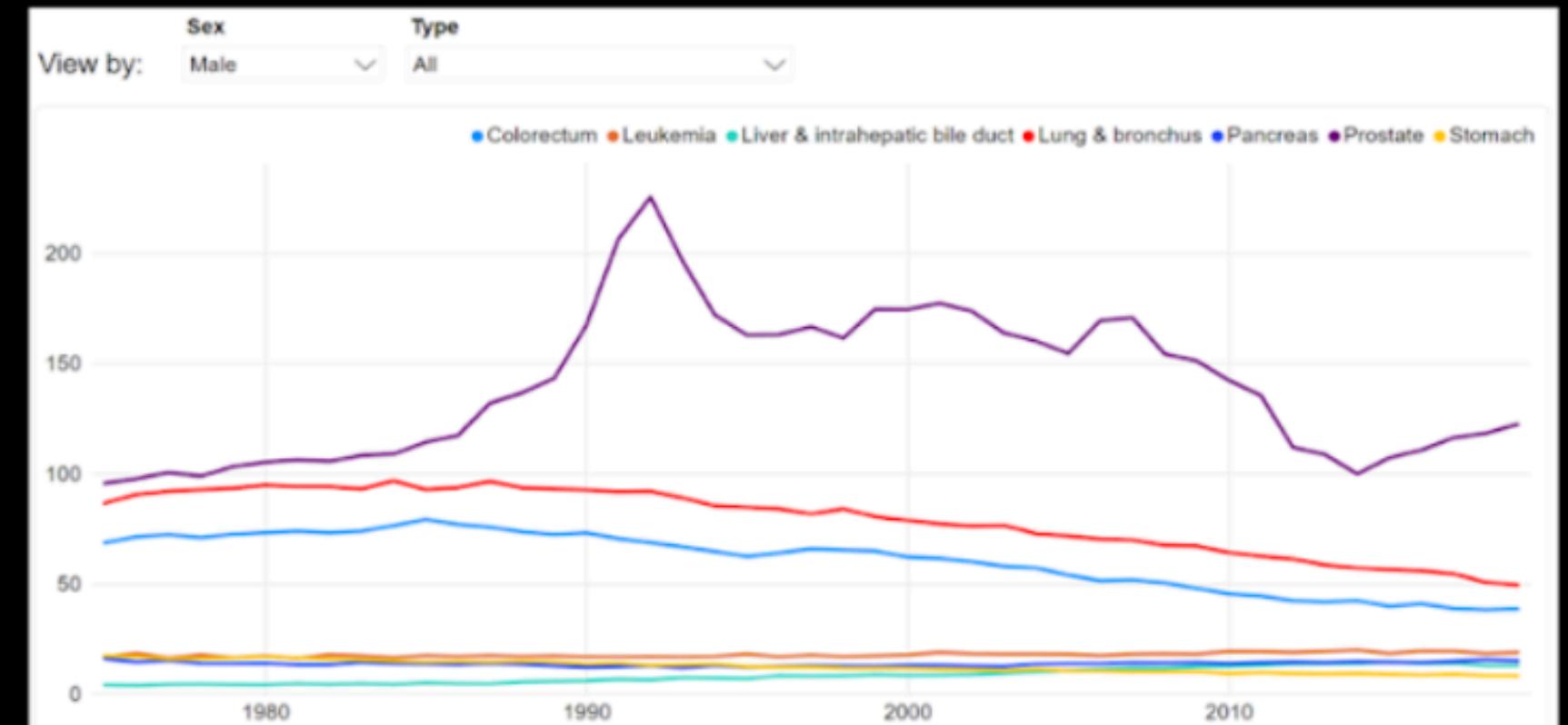
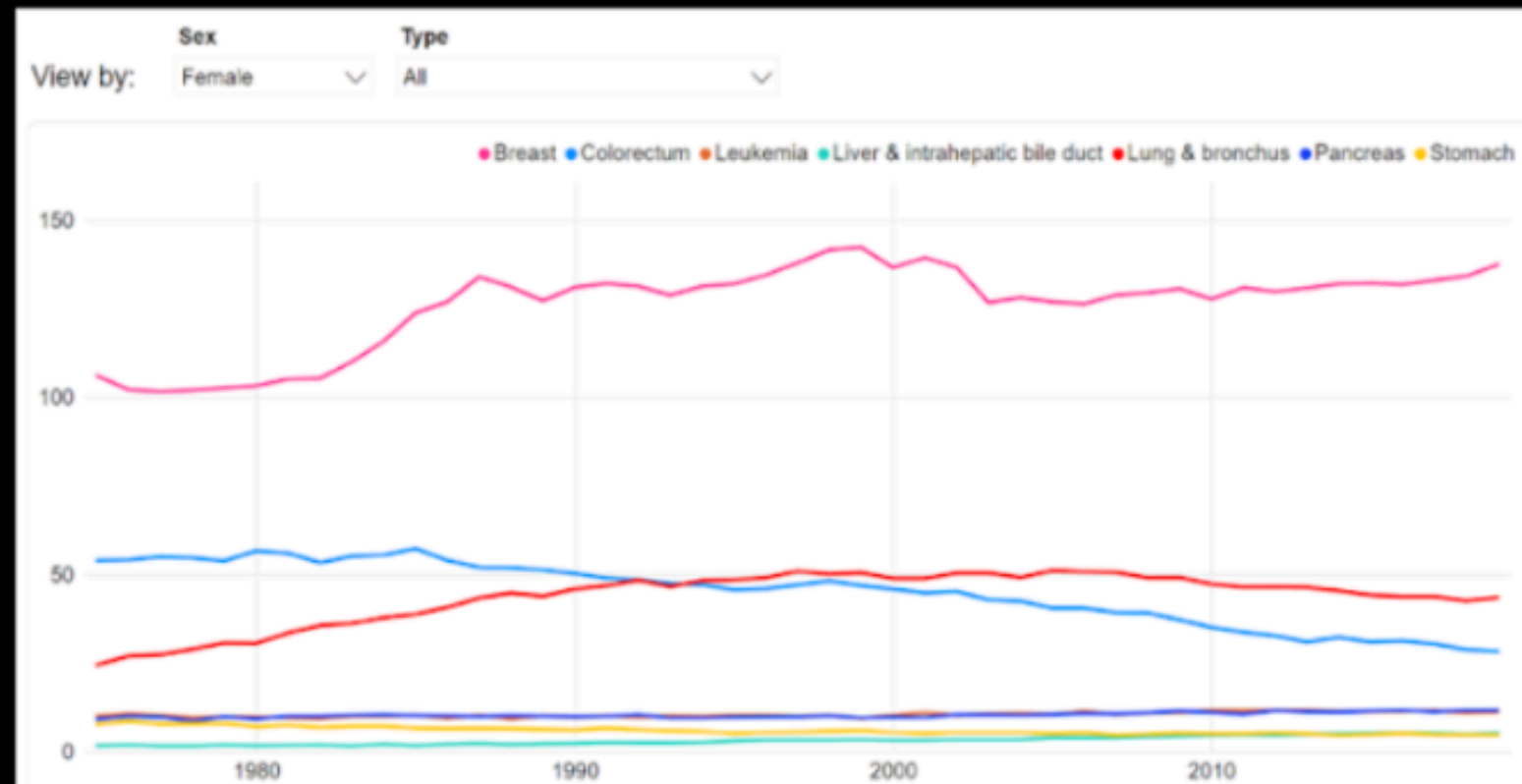
- Rural women often have limited access to specialized cancer care service, leading to delays in diagnosis and treatment.
- Young women, particularly those under 40, face unique challenges including delayed diagnosis due to low suspicion of cancer in this age group and pressure from social relationships and work.
- Financial constraints can lead to delays in seeking care, adherence to treatment, and ability to afford medications.
- Despite great efforts, socioeconomic status, race, and geographic location are still major challenges for promoting medical justice.



Early 21st Century

- Studies showed that people of color had a higher mortality rate from cancer compared to white people, despite having an overall lower incidence rate.
- Factors included later-stage diagnosis, differences in tumor biology, and disparities in access to quality care.
- Cancer awareness grew significantly with campaigns like Breast Cancer Month. Advances in medical research and treatment improved survival rates, contributing to changing public perceptions and reducing stigma.

Incidence Rates Comparison Between Male and Female



Incidence Rates

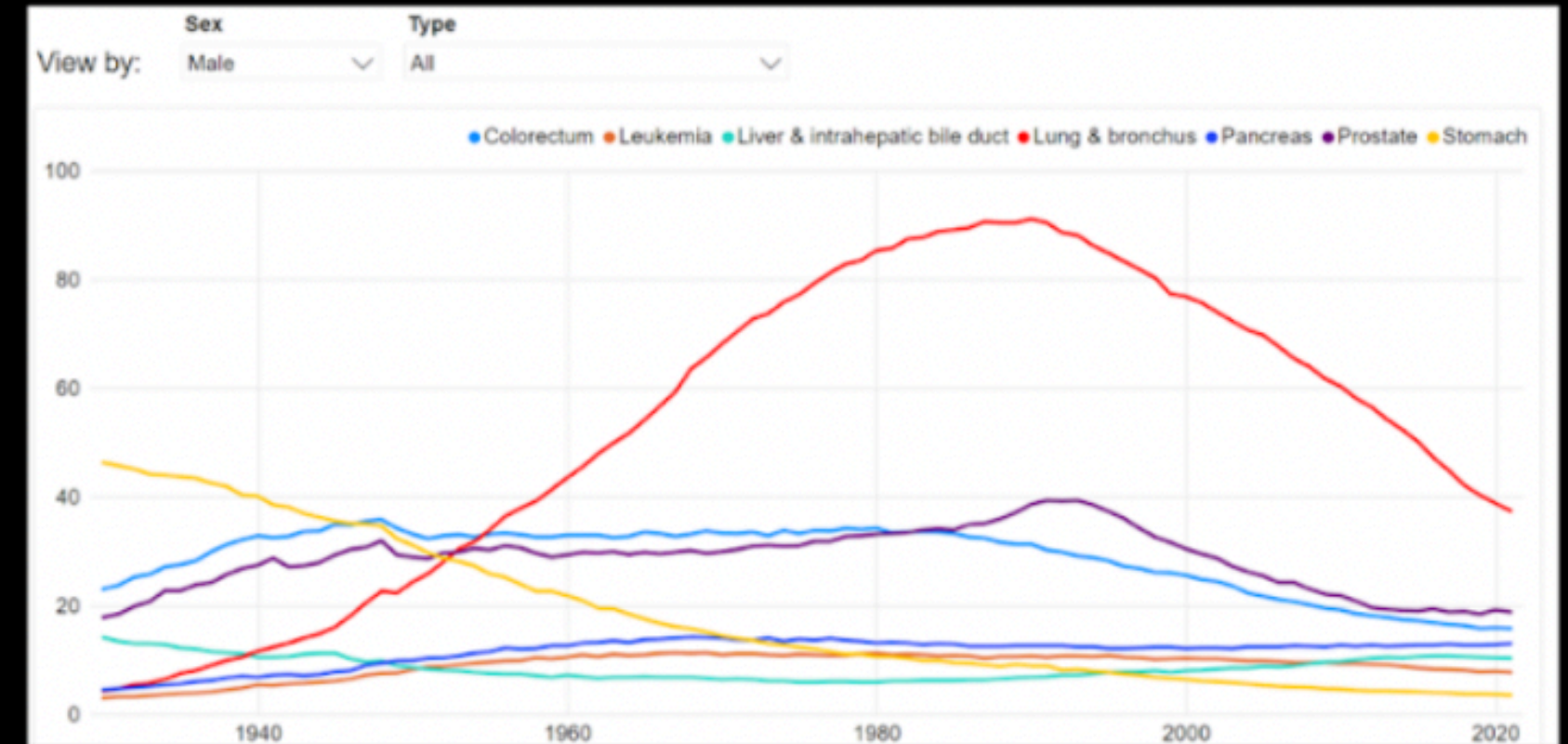
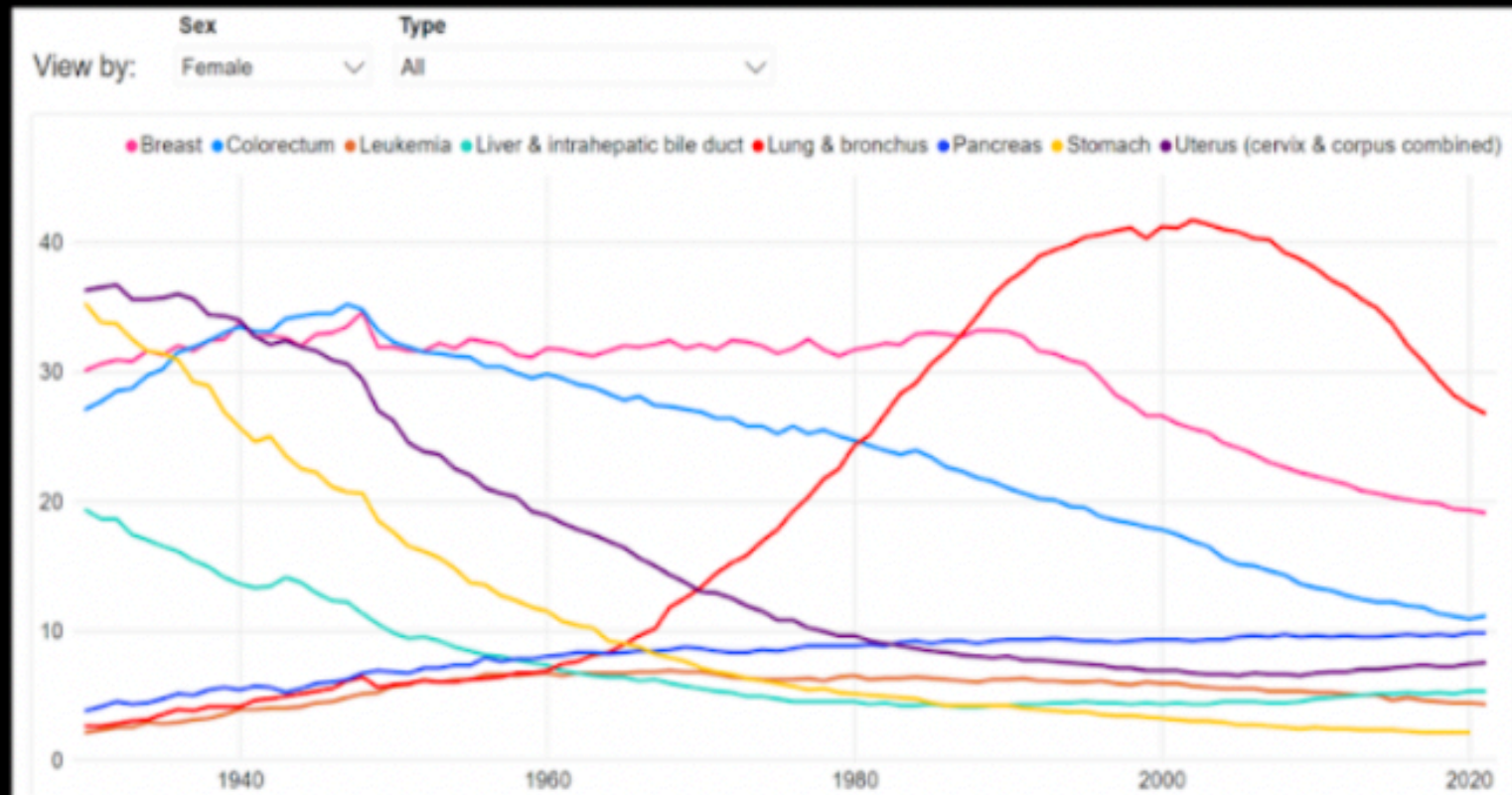
Common Trends:

- Both genders show declining trends in the incidence rates of colorectal and lung & bronchus cancers.
- Increased awareness, preventive measures, and lifestyle changes have contributed to the reduction in incidence rates for these cancers.

Differences:

- **Women:** Breast cancer incidence increased until around 2000, followed by stabilization, likely due to increased awareness and screening.
- **Men:** Prostate cancer incidence showed a significant rise until 2000, followed by a decline, reflecting the impact of widespread screening practices.

Death Rates Comparison Between Male and Female



Death Rates

Common Trends:

- Both genders experienced significant reductions in death rates for major cancers like lung & bronchus and colorectal cancer.
- Advancements in medical technology, improved screening methods, and better treatment options have contributed to the overall decline in cancer death rates for both men and women.

Differences:

- Women:** Breast cancer showed a significant decline in death rates post-1980 due to improved detection and treatment methods.
- Men:** Prostate cancer death rates increased until around 1990, followed by a decline due to better screening and treatment.

DATA CRITIQUE

Data Analysis

To gain deeper insights, we performed several analyses and visualizations on the dataset:

1.Descriptive Statistics:

1. We calculated descriptive statistics for key columns, providing a summary of the central tendencies and variability in cancer rates by age and race.

2.Correlation Analysis:

1. We generated a correlation matrix to identify relationships between different cancer rates and demographic variables.

3.Regression Analysis:

1. We conducted a regression analysis to explore the relationship between total cancer rates and lung cancer rates by race. This helped in understanding how lung cancer rates among different racial groups contribute to the overall cancer rates.

Information Included in the Dataset:

The dataset includes various columns such as:

- **Total.Rate:** Total cancer rate per 100,000 people.
- **Total.Number:** Total number of cancer cases.
- **Total.Population:** Total population considered.
- **Rates by Age:** Cancer rates for different age groups (e.g., < 18, 18-45, 45-64, > 64).
- **Rates by Sex and Age:** Cancer rates segmented by sex and age groups.
- **Rates by Race:** Cancer rates for different racial groups (e.g., White, Black, Asian, Indigenous).
- **Types of Cancer:** Specific cancer types (e.g., Lung, Colorectal, Breast) segmented by age and race.

Phenomena Illuminated by the Dataset:

•Geographic Variation in Cancer Rates:

- Analysis of cancer rates across different states highlights geographic disparities.

•Racial Disparities in Cancer Rates:

- Segmentation by race reveals significant disparities in cancer incidence among different racial groups.

•Age-Specific Cancer Rates:

- The dataset provides insights into how cancer incidence varies with age.

•Specific Cancer Types:

- Focus on lung, colorectal, and breast cancer rates provides detailed understanding of these common cancers.

Limitations of the Dataset:

•Lack of Socioeconomic Data:

- Missing variables such as income, education, and employment status.

•Absence of Behavioral Data:

- No information on risk factors like smoking, diet, and physical activity.

•Environmental Exposure Data:

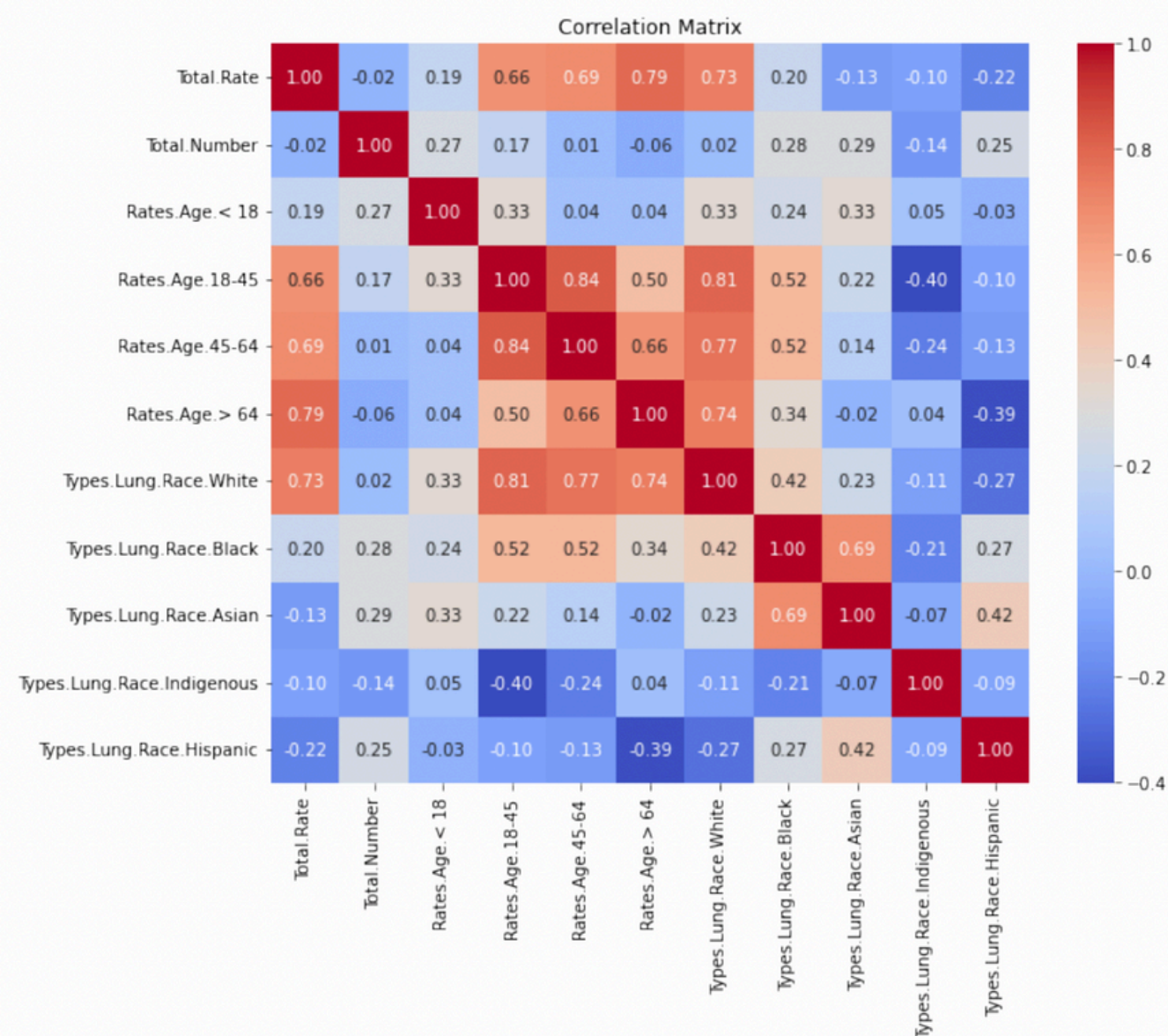
- Lack of data on pollution, occupational hazards, and other environmental factors.

•Temporal Trends:

- The dataset provides a snapshot without longitudinal data to track changes over time.

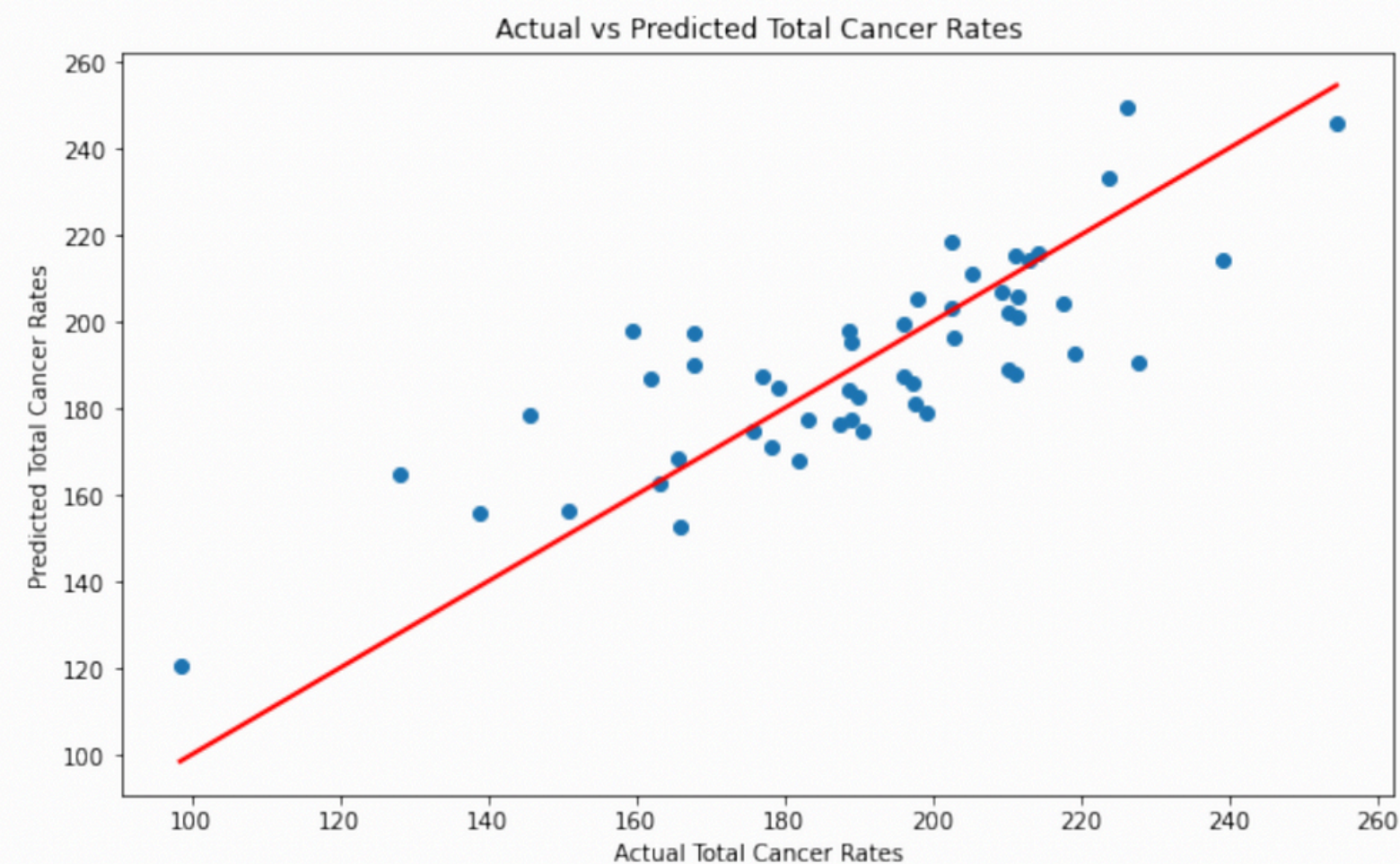
DATA VISUALIZATION

link to dataset: <https://emerald-astrid-91.tiiny.site>



Visualization 1: Correlation Matrix

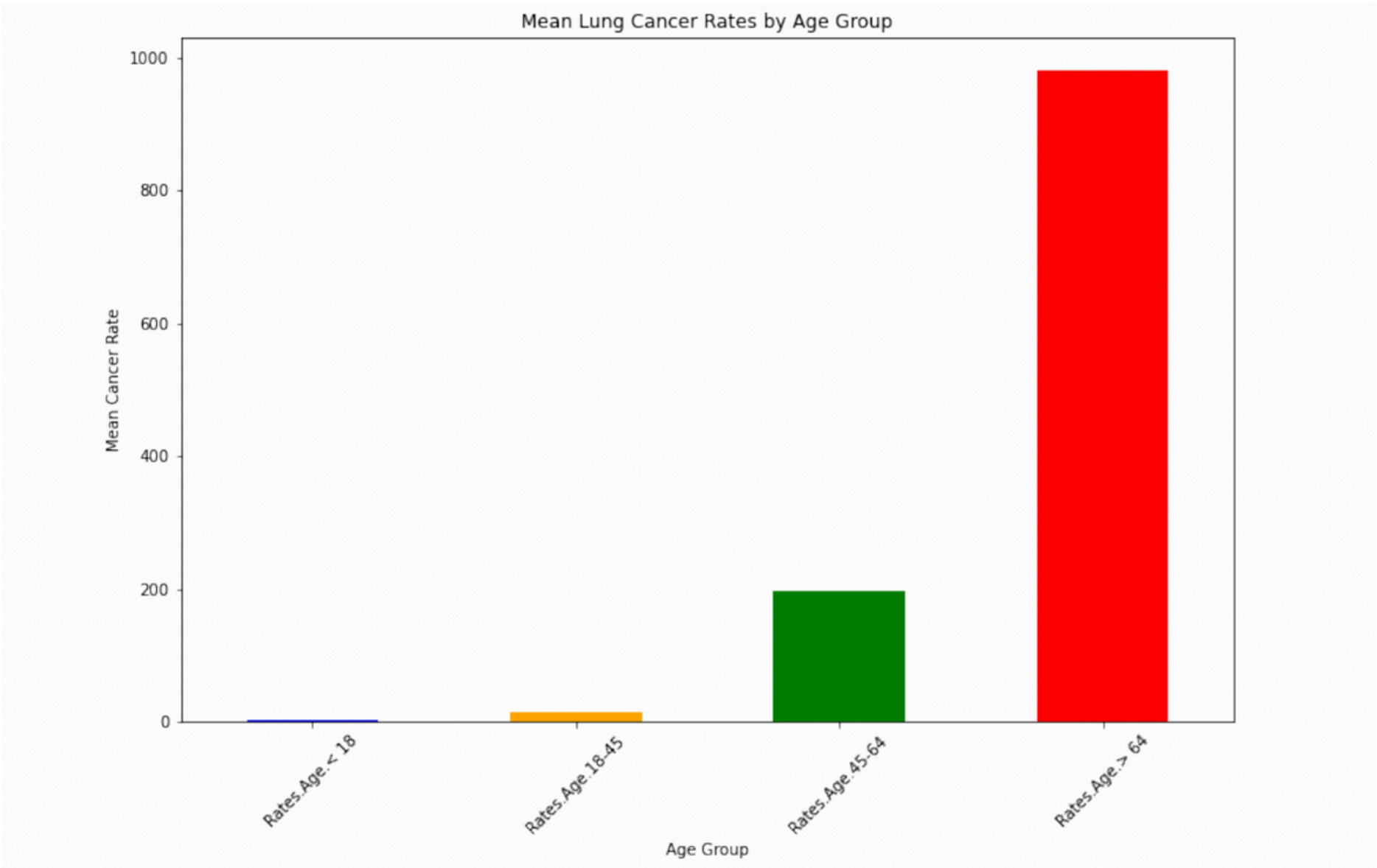
- Objective:** To understand the correlations between different variables in the dataset.
- Description:** A heatmap showing the correlation coefficients between variables such as total cancer rate, cancer rates by age, and lung cancer rates by race.



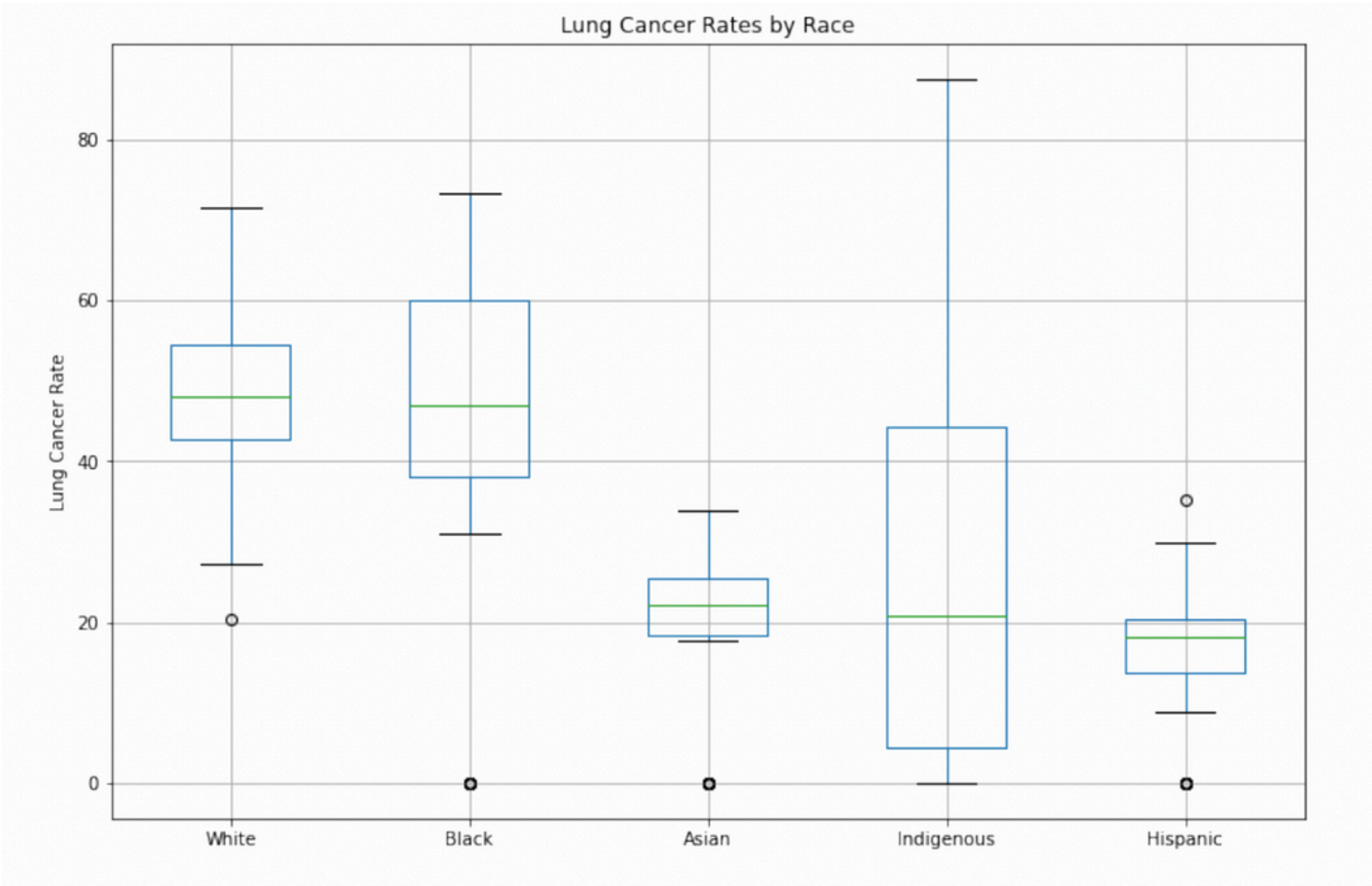
Visualization 2: Actual vs Predicted Total Cancer Rates

- Objective:** To compare the actual and predicted total cancer rates to assess the accuracy of the predictive model.
- Description:** A scatter plot showing actual total cancer rates versus predicted total cancer rates with a line of best fit.

DATA VISUALIZATION

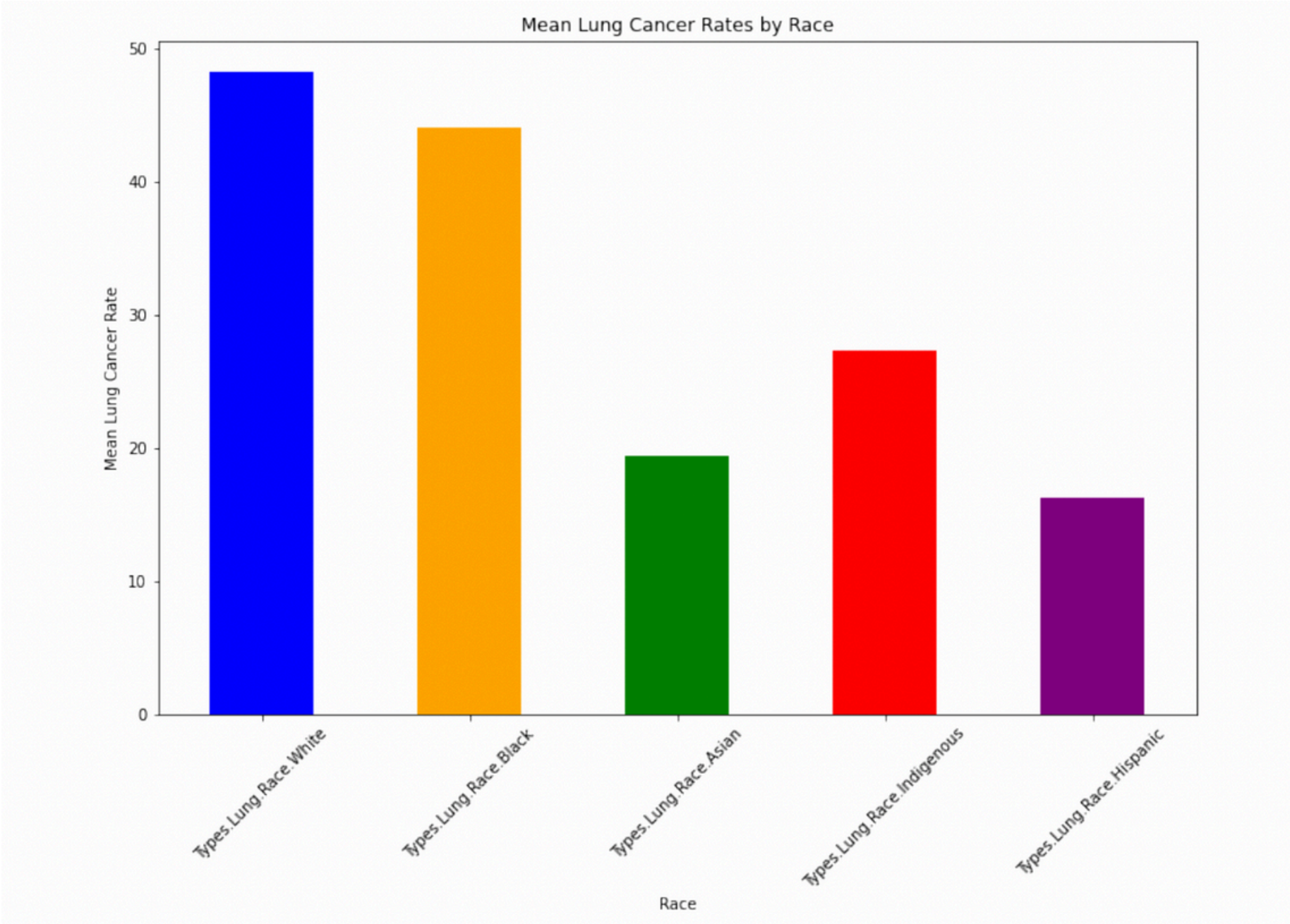


Visualization 3: Mean Lung Cancer Rates by Race
•**Objective:** To compare lung cancer rates among different racial groups.
•**Description:** A bar chart showing mean lung cancer rates for White, Black, Asian, Indigenous, and Hispanic populations.

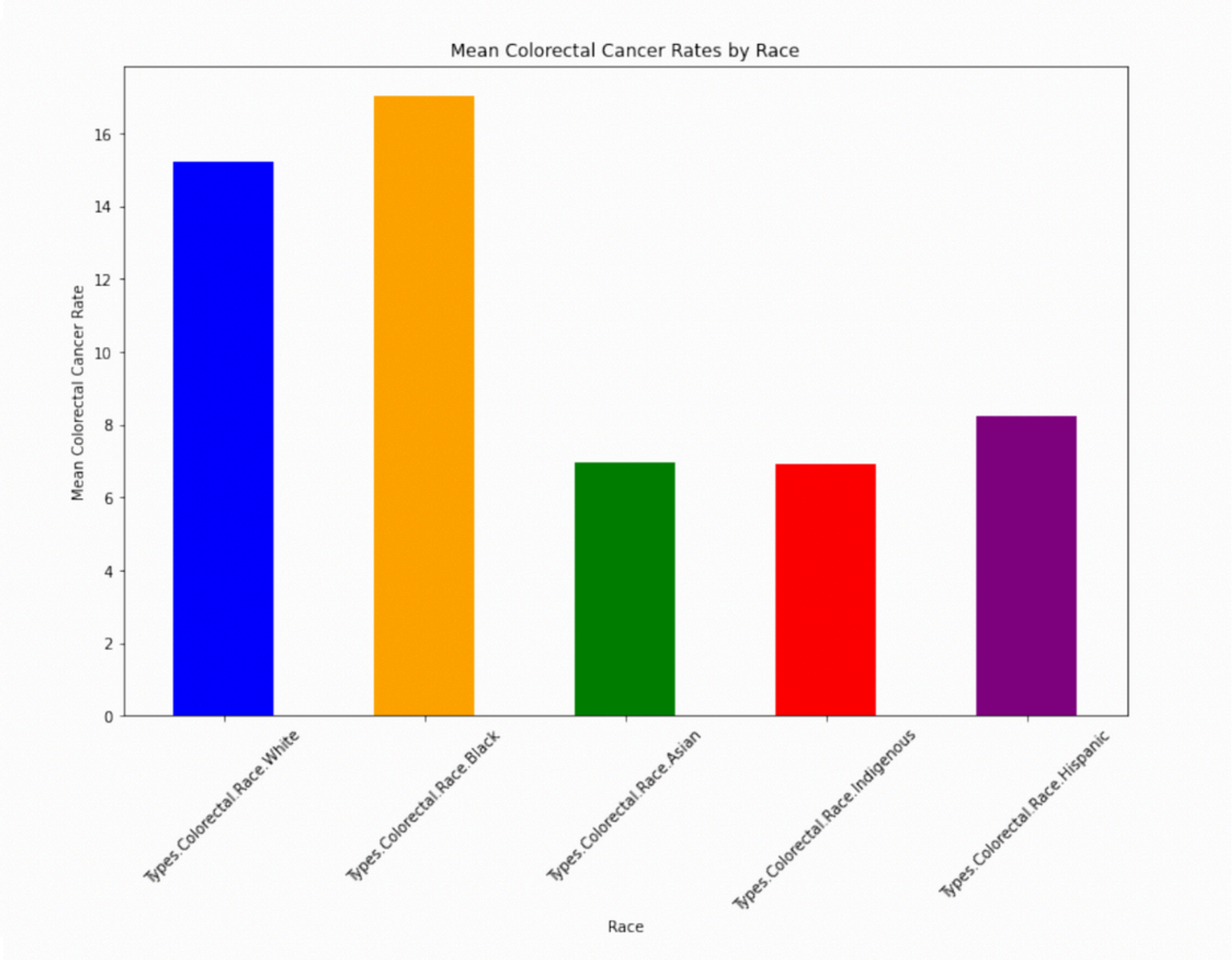


Visualization 4: Mean Breast Cancer Rates by Sex and Age Group
•**Objective:** To illustrate breast cancer rates across different age groups for females.
•**Description:** A bar chart showing mean breast cancer rates for age groups 18-44, 45-64, and > 64.

DATA VISUALIZATION

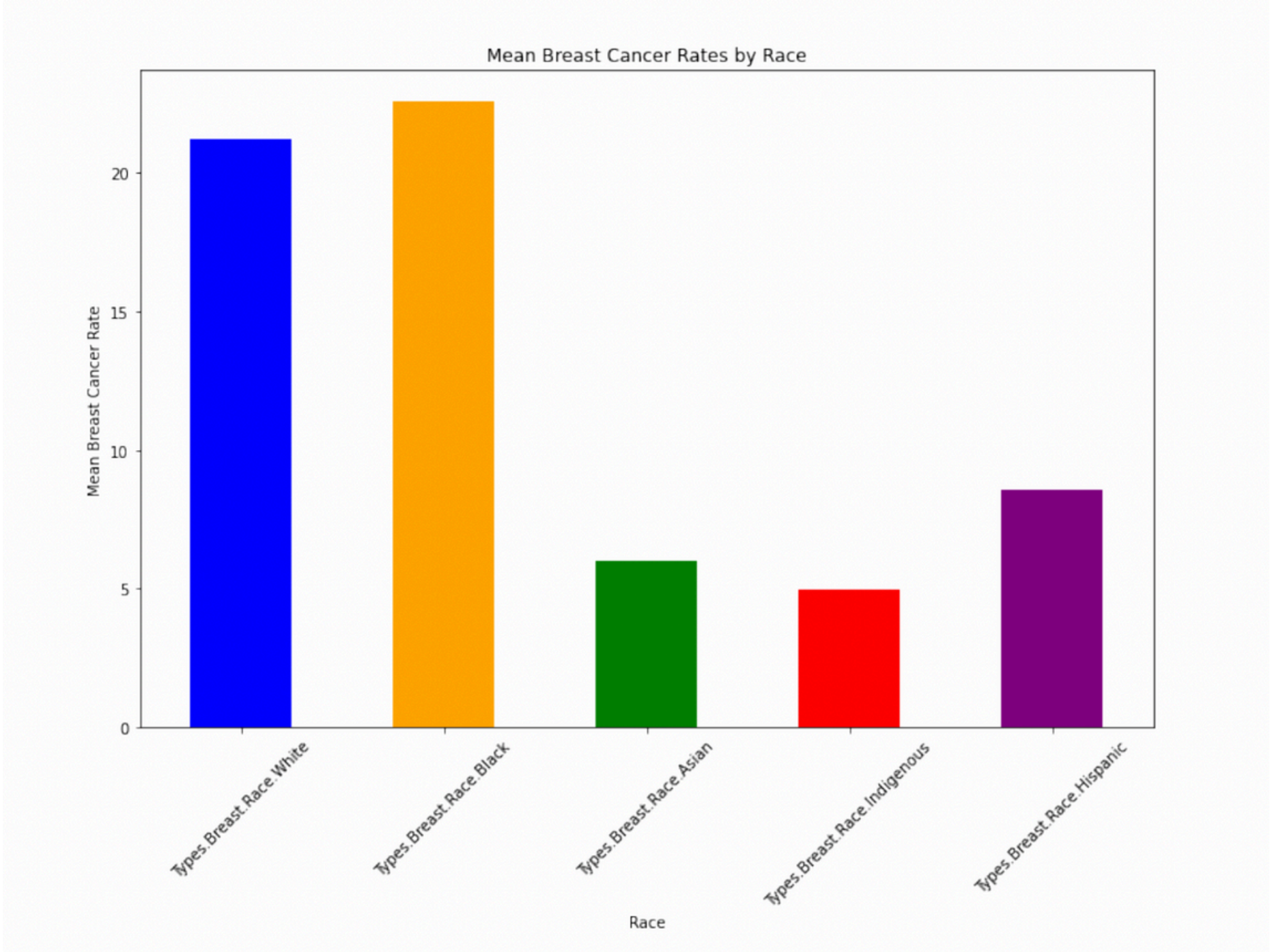


Visualization 5: Mean Lung Cancer Rates by Age Group
•**Objective:** To illustrate how lung cancer rates vary across different age groups.
•**Description:** A bar chart showing mean lung cancer rates for age groups < 18, 18-45, 45-64, > 64.



Visualization 6 : Mean Colorectal Cancer Rates by Race
•**Objective:** To compare colorectal cancer rates among different racial groups.
•**Description:** A bar chart showing mean colorectal cancer rates for White, Black, Asian, Indigenous, and Hispanic populations.

DATA VISUALIZATION



Visualization 7 : Mean Breast Cancer Rates by Race

- Objective:** To compare breast cancer rates among different racial groups.
- Description:** A bar chart showing mean breast cancer rates for White, Black, Asian, Indigenous, and Hispanic populations.

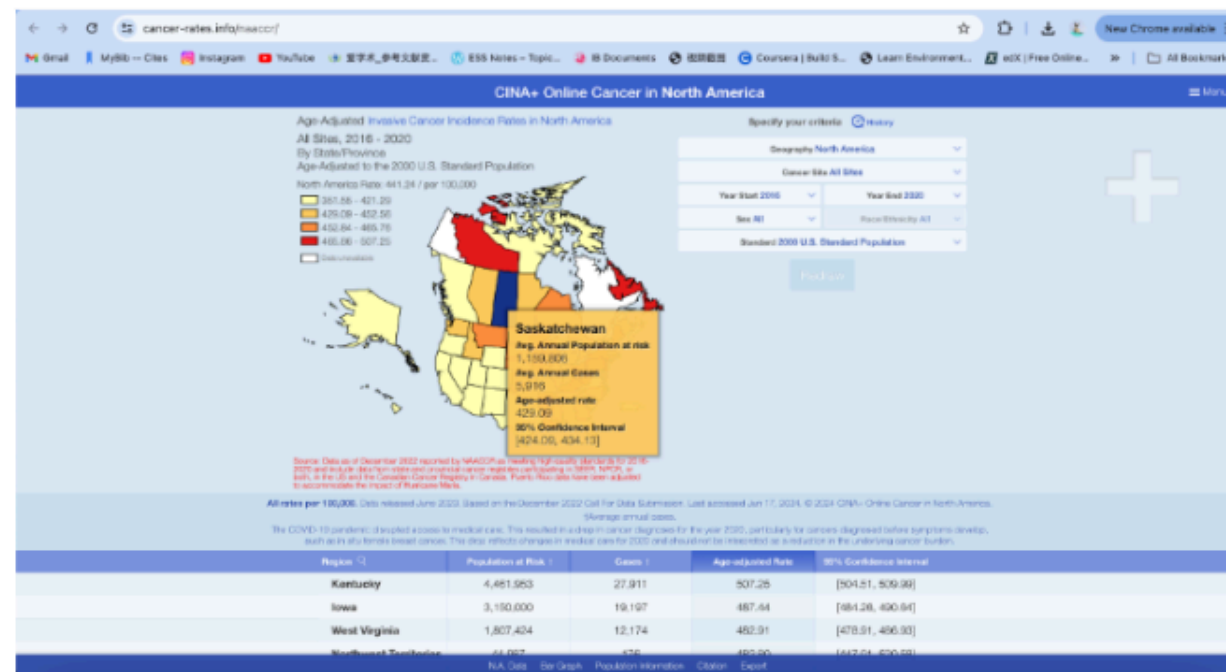
DATA VISUALIZATION PLAN

1. Interactive Map:

Visualize the total cancer incidence rates by state.

Use color gradients to represent the incidence rates, with interactive features to display detailed data when hovering over each state.

Example gained online: <https://www.cancer-rates.info/naaccr/>



2. Bar Charts:

Compare cancer incidence rates among different racial groups.

Use different colors for each racial group to highlight disparities.

3. Heat Maps:

Show the intensity of specific cancer types (e.g., breast, colorectal, lung) across different regions.

Use varying shades to represent different levels of incidence.

WEBSITE LAYOUT AND DESIGN

Since we haven't finalized the content and layout of the web page, I'll just list the pages that I think we need for our website and the content that will be filled with great wisdom after each page. If you are willing to use my plan to make a website after the design can follow this idea.

- Intro Page: Briefly explain what we are doing in this project, including the hypothesis of the project, the purpose of the research, and a brief description of the research results.
Describe the database we use and why we chose it (I wonder if I need a separate page)
- Member page: bio introduction of each member of the group and the specific part of the participation
- Results display and analysis: Adopt the form of two-layer web page, with subheadings on the main menu: gender difference analysis, racial difference analysis, regional difference analysis
Each small page explains how to analyze and how to process the data, along with the visualization of the results of the analysis, showing the different incidence rates of different cancers in different regions and different populations under the influence of these two different conditions, and the reasons why.
Put up an interactive map showing cancer rates by race and location.
- Annotated Bibliography Page: Explain the significance of the three readings for our project, and attach all the data, knowledge, and relevant reference sources and explanations.
- Contact page: Contact form and social media links for readers to get access and contact.

ANNOTATED BIBLIOGRAPHY

1. Delgado, Richard, and Jean Stefancic. *Critical Race Theory: An Introduction*. NYU Press, 2017.

Abstract: This paper explores Critical race theory (CRT), which explores how race and racism intersect with social structures. This article emphasizes the social construction of race, the importance of minority voices and narratives, and the systemic nature of racism.

Evaluation: This source is critical to understanding the systemic inequalities that drive cancer disparities. We used CRT to analyze quantitative data on cancer incidence in different ethnic groups. Stereotypes and biases are always present and inevitable, so comparing the differences between different races in different types of cancer can help us show the different situations that different races face when it comes to health pathology. This analysis helps us understand how systemic racism and social determinants influence health disparities. It guides our interpretation, ensuring that we place the data within the broader social and historical structure of race and inequality.

Reflection: This reading will provide a theoretical framework for analyzing how race and racism affect cancer disparities. This framework helped us find a very interesting research question and hypothesis that supports both the humanities and data links, guiding us to investigate not only the "what" and "how" of cancer differences, but also the underlying "why." It may be that different ethnic groups have different incomes, social status, and depending on their medical conditions and healthy living environments, all of which may affect the incidence of cancer.

ANNOTATED BIBLIOGRAPHY

2. Manovich, Lev. "What is Visualization?" 2011.

Abstract: This paper introduces Manovich's definition of information visualization, and discusses the historical background, principle and application of information visualization. He emphasized the importance of visualization in understanding complex data and highlighted the images, colors, spatial structures, etc. required for different types of visual representations.

Evaluation: This source is critical to understanding how to effectively present cancer identification data. This will help our website visually and effectively show the impact of different social and historical environmental factors on different types of cancer, explore the indirect impact of humans on human health, and effectively present it with appropriate graphics. Manovich's work is valuable in the field of digital humanities and has provided valuable insights into the creation of influential visual DH projects. Visualization is essential to show geographic and demographic patterns of cancer incidence, to concretize abstract data, and to provide interpretability to a wider audience.

Reflection: This reading helps to help us achieve better website visualizations, helps to decide which type of visualizations (e.g., bar charts, heat maps) are best for different types of data, ensures that the information is accessible, and engages the audience. This approach not only makes the information easy to understand, but also interactive, allowing users to engage with the data in a meaningful way.

ANNOTATED BIBLIOGRAPHY

3. Cowles, David L. "Reader Response Criticism." *Literary Criticism: An Introduction to Theory and Practice*, 2001.

Abstract: This paper explains reader response criticism, focusing on the role of readers in text interpretation. He discusses how personal interpretations based on personal experiences have changed and stresses the importance of engaging the audience. The theory is linked to the digital humanistic environment by designing the project interface to make it interactive and user responsive. It encourages users to interpret visual data based on their own context and experience, making user interaction an integral part of narrative and analysis.

Evaluation: This resource is useful for designing a project's website and visualizations to engage users, while allowing us to think about who our project's audience is and what kind of digital humanities information we need for our research station. Cowles' work is solid and provides a foundation for understanding how to create interactive and meaningful content. This approach tells us how to build engagement on a website. It ensures that our content is not only informative, but also resonating and can provoke individual reflection and response, which is essential for educational and advocacy purposes.

Reflection: The theory will guide the narrative structure of the website, ensuring that it engages users in a meaningful way. It will help create interactive elements, allow users to explore data based on their interests and background, and provide channels to connect with creators to enhance the overall impact of the project and help the project improve its content. This way of communicating with our users has provided us with great inspiration, and we also want to complement the message of our digital humanities website more comprehensively by creating areas of interaction with our readers.

NARRATIVE: RESEARCH QUESTIONS AND HYPOTHESIS

Q: Do race and gender have an effect on a person's chance of being diagnosed with cancer?

H: We predict that race and gender do have an effect on a person's chance of being diagnosed with cancer.

Q: How can a person's background impact various aspects of life and ultimately lead to differing outcomes in health?

H: There are so many different ways that background can potentially affect health. For one, access to healthcare / insurance or how a person is treated by healthcare professionals can be impacted by their race, gender, geographic location, and socioeconomic status (not to mention, a lot of these influence each other). Not having access to healthy food is another factor that can be detrimental to someone's health. To look more into these hypotheses, my group can look into disparities in health insurance, differential treatment by healthcare professionals due to race or gender, and research where food deserts are.

NARRATIVE: MAP OF METHODOLOGY

**not entirely sure what to write here besides the theories we would use, but I'll also write about how we found sources ?

Critical race theory: focus on studying racism not on the individual level, but systemically

- View food deserts / access to healthcare through this lens
 - Who is making the decisions to not open affordable grocery stores in black and brown neighborhoods? Or if it's not a decision, what is preventing it from happening?
 - How has healthcare treated people differently based on identity? Where does this stem from?

Intersectionality theory: belief that multiple facets of identity contribute to an individual's privilege / oppression

- Prompts us to ask more questions about combinations of identities in our dataset
 - How does being a woman or a man change someone's likelihood of getting cancer? What if we add race on top of that? Which combination of gender and race is most prone to having cancer? Why?

NARRATIVE: EXPANSION OF DIGITAL HUMANITIES

Our project contributes to the expansion of the Digital Humanities by combining quantitative techniques in Data Science with intersectionality. By sourcing data from accredited organizations such as the National Center for Biotechnology Information (NCBI), our mission lies in breaking barriers to medical knowledge – especially for those not necessarily in the field. Our exploration of cancer-related disparities among race and geographical location enriches the Digital Humanities through:

1. An ample breadth of data collection, incorporating diverse perspectives & backgrounds
2. Several data visualizations, all illustrating different stories under the same dataset

This approach provides a nuanced understanding of how social, cultural, and environmental factors are capable of influencing health outcomes. Uncovering these trends comes with the responsibility of educating present, and future, generations through an intersectional lens.

NARRATIVE: EXPANSION OF DIGITAL HUMANITIES

**How might
differences in
geographical access
to healthcare
influence cancer
detection?**

**Does this see itself
reflected in our
findings?**

**What does this
entail for future
research in public
health and policy?**

Such questions incentivize dialogue in the Digital Humanities, catering to diverse learning styles and setting the stage for intersectional frameworks (i.e. Feminism and Critical Race Theory). As an example, the Association of Medical Colleges (AAMC) uncovers similar findings within the field – denoting how clinical trials are currently seeking to diversify their participant pools to correct medical biases of the past.

Ultimately, we hope our project serves as an invitation for other scholars interested in “democratizing” biological research. Science can act as a catalyst for societal change, and we believe our findings are capable of raising awareness in the accessibility to cancer prevention, detection, and treatment.