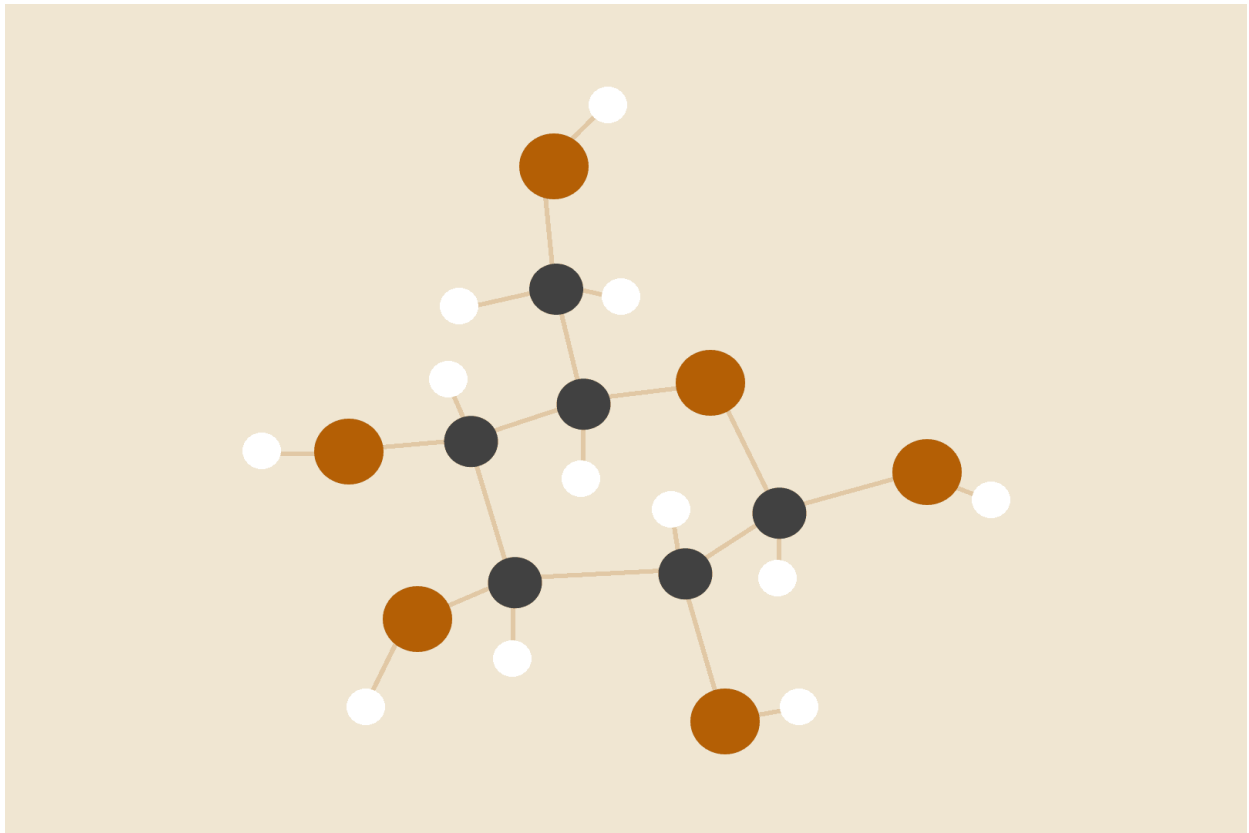


Cancer Isn't Just a Senior Moment: Mapping Common Cancers in Young Adults

DS4630 - Data Visualization



John Chen u1364584, Michael Molenaar u0712166

E-mail addresses: johnlingchen12345@gmail.com michaeldeanmolenaar@gmail.com

UIDs: u1364584, u0712166

PROJECT PROPOSAL

Background and Motivation

Our decision to create a human body map of common cancers in adults is driven by a desire to enhance public awareness and education. Everyone knows (or is) someone who has been affected by cancer. However, a general belief of our population is that cancer is an old person's concern—which is true, but not exclusively true. Specifically, the American Cancer Society annual report shows younger adults to be the only age group with an increase in overall cancer incidence between 1995 and 2020—the rate has risen by 1-2% each year during this period.

We believe that a visual representation of where cancers occur within the body can be a powerful tool for health professionals and the general public to better understand the prevalence of different cancer types. We also believe that it is crucial to highlight the specific cancers young people are increasingly being diagnosed with. By offering an anatomical perspective, we can simplify complex medical information, making it more accessible and easier to comprehend for an age group that may be predominantly unaware of these increasing risks.

Furthermore, given the well-documented health disparities across different demographic groups, we aim to highlight the critical information surrounding the heightened risks these disparities present.

Project Objectives

How can a visual representation help people understand and recognize the risks of different cancers?

Questions

1. What are the most common and significant cancers found in young people (Early on-set cancer)?
2. Which cancers are most common in different areas of the human body?
3. How does the prevalence of these cancers vary by demographic group: gender, ethnicity, geographic location, sexual orientation, disability status?
4. What external factors (e.g., lifestyle, environment, or genetics) contribute to the occurrence of different types of cancers?
5. What are the early warning signs or symptoms for these cancers, and where in

the body do they typically manifest?

6. How can a visual representation help people understand and recognize the risks of different cancers?

Benefits

1. Identifying which cancers are most prevalent in young people allows for targeted research, early detection programs, and public health initiatives.
2. Educational campaigns can be developed to raise awareness about symptoms and promote regular checkups in those areas, leading to earlier detection and treatment.
3. Visualizing early warning signs and risks can motivate individuals to seek medical advice earlier or engage in preventive behaviors. Also, by illustrating how certain cancers manifest, this project could help encourage preventative care and screenings.
4. An easier-to-read/follow visual representation could be more accessible to people who might not have a medical or statistical background.
5. The personal connection many people have with cancer can make this project resonate with a wide audience, fostering a deeper understanding and empathy for those affected by cancer.

Data

National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER):

- Cancer data, including incidence, survival rates, and mortality
- U.S. data
- <https://seer.cancer.gov/>

Centers for Disease Control and Prevention (CDC):

- Data on cancer incidence, mortality, and risk factors across different demographic groups
- Global data
- <https://wonder.cdc.gov>

American Cancer Society (ACS):

- Cancer facts and statistics, which include reports on cancer types by age, gender, and other demographic factors
- U.S. data

- <https://cancer.org>

World Health Organization (WHO):

- **Global Cancer Observatory (GLOBOCAN):** reports cancer incidence, mortality, and prevalence
- Global data
- <https://gco.iarc.fr>

Harvard School of Public Health:

- Published studies and datasets on cancer incidence, survival rates, and risk factors across various demographic groups
- Both U.S. and Global data
- <https://www.hsph.harvard.edu>

Data Processing

We expect to do substantial data cleanup for this project, given the potential complexity and variety of cancer datasets available. Data from sources like SEER or similar may contain missing values, inconsistencies in formats (e.g., dates, categorizations), or excess details that aren't relevant to our specific focus on mapping cancers to the human body.

- **Removing Irrelevant Fields** - The datasets may include variables not necessary for our maps, such as geographic information or socioeconomic data, which we'll filter out.
- **Handling Missing Data** - We may encounter missing values in important fields like cancer type, stage, or patient demographics, which will require strategies such as interpolation, imputation, or removal of incomplete records.
- **Standardizing Formats** - Ensuring that all data, such as age ranges, cancer categories, and time intervals, are standardized for consistency across our map.
- **De-duplication** - If the datasets contain repetitive entries or overlapping information, we will need to identify and remove duplicates.
- **Accuracy in Anatomical Illustration** - This is crucial to our overall visualization of cancers in young adults.

Visualization Design

Starting View

Upon opening up the dashboard, the user will see this is the first visualization. Here, the user can click on our Anatomy Explorer (this name is a work in progress) and will be able to explore a vast array of different cancer-located areas.

Illustration

Mapping Common Cancers in Young Adults

Click the Anatomy Explorer to explore different cancer regions.



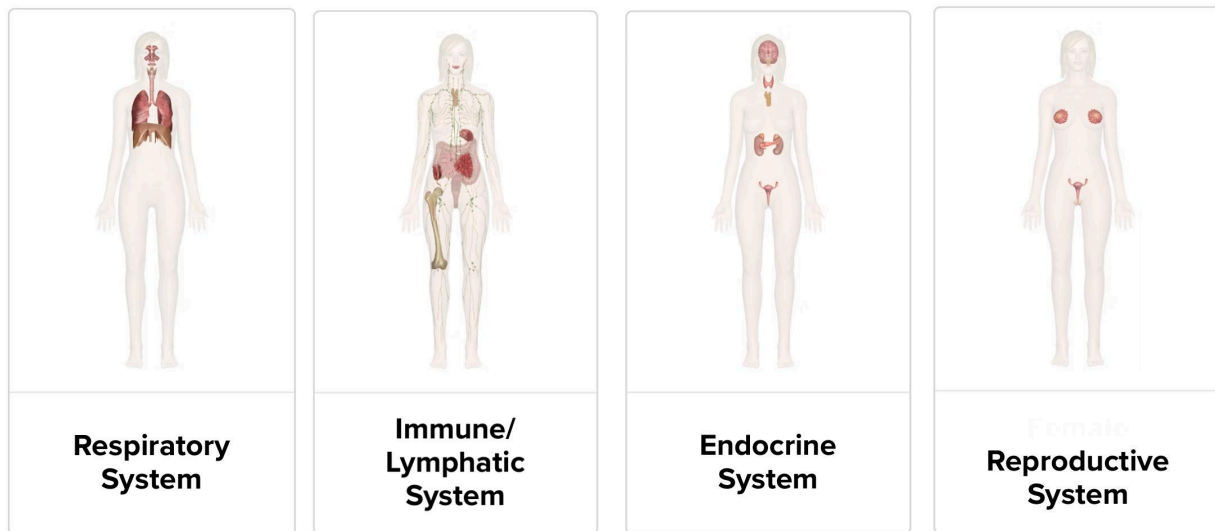
Anatomy Explorer - Systems

After clicking, the user will be able to explore different anatomical systems.

Illustration

Anatomical Systems

Select one system for further insights.



These are all pictures we found online. We hope to have a cleaner view of all three stacked on top of each other, with the ability to choose each system by dragging a mouse over them (highlighting the selected system), and then clicking to enter a more detailed view of the system. This will allow us to have more systems available for users to explore.

Selecting System Example:

Respiratory System

Immune/Lymphatic System

Endocrine System

Reproductive System

Anatomy Explorer - Cancer Locations

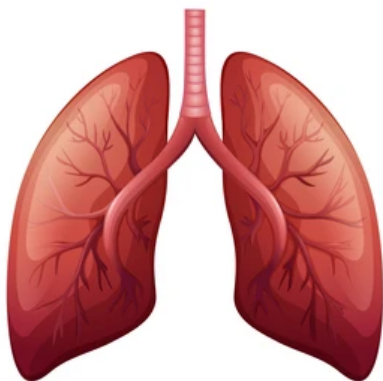
After the user selects the system of interest, a new visualization will appear, only depicting that system and the locations where cancer is most prevalent.

Respiratory System

Select each cancer location to learn more about its risks and causes.



Here, users can individually select cancer locations (above depicted in green). A pop-up window will show a plethora of details regarding the statistics, causes, and more about the individual cancer and its prevalence amongst young adults. We will also be adding a risk rating similar to that of a fire danger rating.



Lung Cancer *Decreasing Risk*

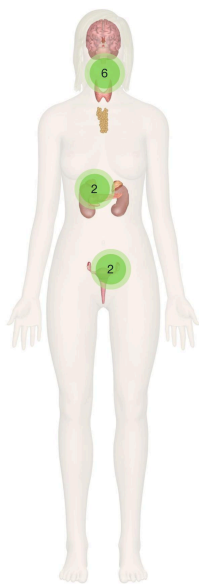
Incident Rate: 2.5 / 100,000 young adults

2.5% annual decrease

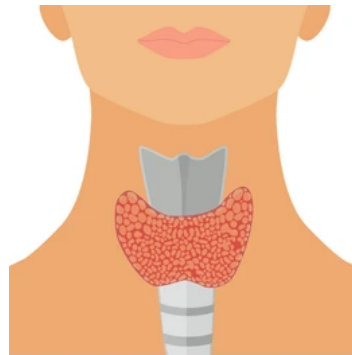
Contributing Factors: Decreased Smoking Rates, Improved Awareness, Genetic Factors



Endocrine System



=>



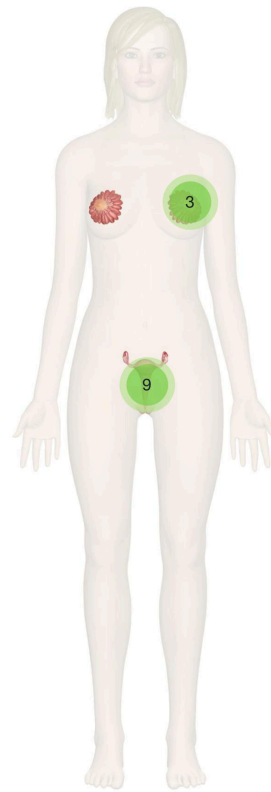
Thyroid Cancer *Increasing Risk*

Incident Rate: 6.7 / 100,000 young adults

3-4% annual increase

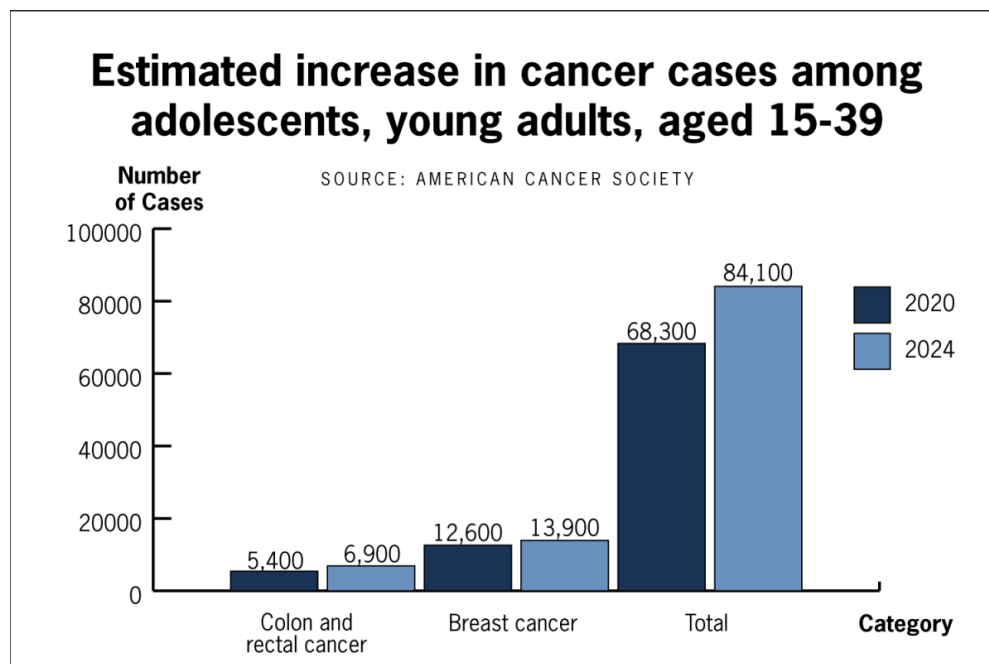
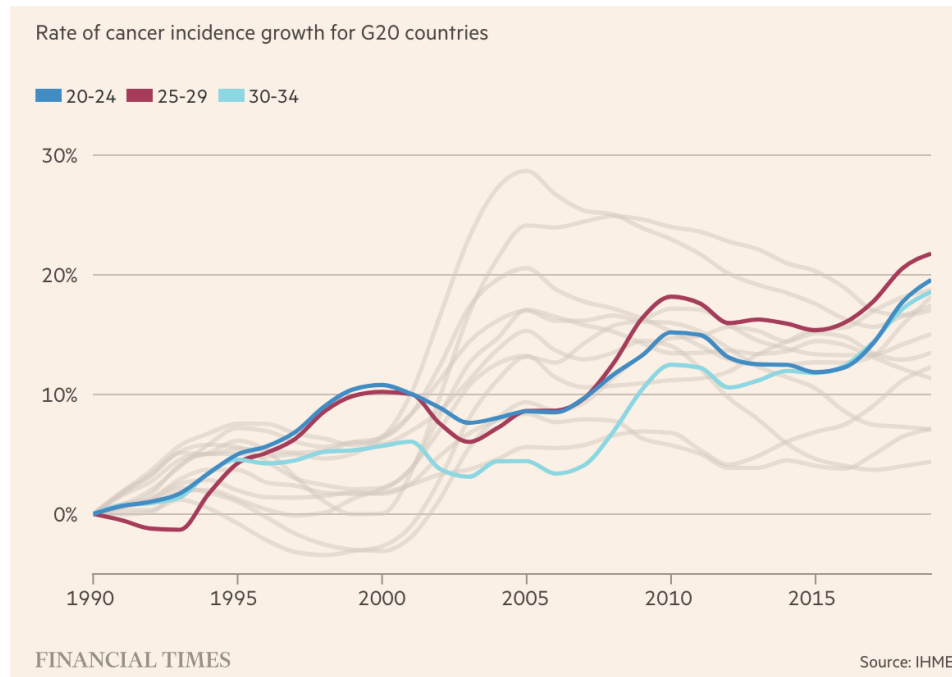
**Contributing Factors: Radiation Exposure,
Improved Detection and Screening**

Reproductive System (this system will have two separate illustrations)



Early On-set Cancer Statistics

Below our comprehensive Anatomy Explorer, we will be providing many different statistics to highlight the magnitude of this dangerous increase we are seeing in early on-set cancers. Included will be many charts, similar to the ones depicted below.



Resources

At the very end of our dashboard, we will have resources for people who may be nervous or believe they are at more risk of developing cancer. We want to also provide links to additional information, as well as briefly list our sources.

Must-Have Features

Including Common Adult Cancers - The project should focus on common adult cancers with accurate statistics. Leaving out major cancers (e.g., lung, breast, prostate) or misrepresenting their prevalence would limit the educational impact and relevance of the project.

Clear Indication of Risk and Early Symptoms - Providing data on early warning signs and risk factors for each cancer is crucial for supporting early detection efforts. Without this, the map would lack practical application for users interested in prevention and self-awareness.

Health Disparities in Specific Demographic Groups - Showing visual comparisons of cancer prevalence by gender, ethnicity, etc. would provide deeper insights.

Key External Factors Contributing to Cancer Increases - The map should highlight how external factors (such as lifestyle, environment, and genetics) contribute to cancer occurrence. Without this feature, the project would miss an important opportunity to educate users about modifiable risk factors.

Clarity and Simplicity of Visualization - The human body map should convey information without overwhelming or confusing the user. A cluttered or overly complex visualization would fail to meet the goal of making cancer information more accessible to the general public.

Optional Features

Interactive Elements - An interactive map where users can click on different regions of the body to learn more about specific cancers, risk factors, and symptoms would add a dynamic element, but the core information could still be conveyed in a static format if necessary.

Mobile-Friendly Design - Optimizing the map for mobile devices would improve accessibility but isn't critical if the map is functional and clear on larger screens.

3D Visualization: A 3D rendering of the human body to showcase cancer locations would be an advanced feature, but a 2D map can effectively communicate the necessary information.

Project Schedule (Subject to Improvisation)

Week 1: Project Planning & Initial Research

- **Tasks:**
 - Define the scope of the project, focusing on common adult cancers and external risk factors.
 - Research sources of cancer data (e.g., SEER, WHO, CDC).
 - Create a rough outline of the human body map layout.
- **Responsibility:**
 - John: Research data sources and project scope.
 - Michael: Draft the map layout and interface concept.

Week 2: Data Collection

- **Tasks:**
 - Download and explore datasets related to cancer types, body regions, prevalence, and risk factors.
 - Begin cleaning the data (removing irrelevant information, handling missing data).
- **Responsibility:**
 - John: Download datasets and start data cleaning.
 - Michael: Investigate how cancer types will be visually represented on the map.

Week 3: Data Cleanup & Preparation

- **Tasks:**
 - Finalize data cleaning: handle missing data, ensure consistency in units and formats.
 - Organize datasets by body region and cancer type.
- **Responsibility:**
 - John: Data cleanup and preparation.

- Michael: Validate that the cleaned data aligns with the map concept.

Week 4: Initial Data Analysis

- **Tasks:**
 - Perform initial analysis to identify trends in cancer types by body region.
 - Identify relevant external factors (e.g., smoking, genetics) contributing to cancer.
- **Responsibility:**
 - John: Conduct initial analysis of cancer prevalence.
 - Michael: Correlate cancer types with external risk factors.

Week 5: Body Map Design & Tool Selection

- **Tasks:**
 - Choose the software and tools (e.g., Python, Tableau, Plotly) to visualize the human body map.
 - Start developing the visual layout for the map.
- **Responsibility:**
 - John: Research visualization tools.
 - Partner: Create a draft version of the map design.

Week 6: Prototype Development

- **Tasks:**
 - Begin creating a prototype of the map with basic cancer type and body region visuals.
 - Incorporate a few external factors (like smoking) to test how they can be visually represented.
- **Responsibility:**
 - John: Build the first version of the body map.
 - Michael: Add basic interactivity or visual enhancements.

Week 7: Feedback & Refinement

- **Tasks:**
 - Present the prototype to a small audience (peers, mentors) for feedback.
 - Collect feedback and suggestions for improvements in visualization and

data representation.

- **Responsibility:**
 - John: Compile feedback.
 - Michael: Adjust visual elements based on suggestions.

Week 8: Full Data Integration

- **Tasks:**
 - Integrate the complete dataset into the map, including all body regions and cancer types.
 - Incorporate external risk factors, ensuring they are visually distinct and informative.
- **Responsibility:**
 - John: Finalize data integration.
 - Michael: Refine how external factors are visually represented.

Week 9: Final Visualization Enhancements

- **Tasks:**
 - Polish the map's visual design, ensuring clarity and usability.
 - Test the map on different devices and formats to ensure accessibility.
- **Responsibility:**
 - John: Refine map design for clarity.
 - Michael: Conduct usability testing.

Week 10: Documentation & Explanation

- **Tasks:**
 - Prepare documentation that explains the map's purpose, data sources, and features.
 - Include descriptions of cancer types, body regions, and how external factors influence cancer risks.
- **Responsibility:**
 - John: Write the project documentation.
 - Michael: Create a user guide for navigating the map.

Week 11: Final Testing & Debugging

- **Tasks:**
 - Test the map thoroughly, ensuring all data is accurately represented and interactive features work smoothly.
 - Debug any issues related to data display, interactivity, or responsiveness.
- **Responsibility:**
 - John: Final testing and data accuracy check.
 - Michael: Debug and ensure the map is functional.

Week 12: Final Presentation Preparation

- **Tasks:**
 - Finalize the map and all supporting materials.
 - Prepare for the final presentation, including slides, demos, and speaking points.
- **Responsibility:**
 - John: Final map checks and presentation material preparation.
 - Michael: Coordinate the demo and prepare for presenting the results.