



DAB 304 Healthcare Analytics

Project Report

“Cancer Trends/Profiles in Canada”

Submitted By:

Jigar Patel (0794004)

Smit Arora (0795650)

Dhrumi Patel (0793770)

Bhakti Bhatt (0788808)

Vismay Lad (0794141)

Introduction

- According to Ontario Cancer Statistics 2018, cancer accounted for 29.3% of all deaths in Ontario. In Ontario, more than twice as many instances of the disease were reported in 2013 as there were in 1983. According to experts, 1 in 4 Ontarians will die from cancer and 1 in 2 will develop the disease at some point in their lives, with the risk of developing the disease being the same for all racial and ethnic groups. In Ontario, the price of disease care increased from over \$1 billion in 2005 to \$2.6 billion in 2012, or nearly 150%. The most lethal cancers are lung, colorectal, breast, and prostate cancers.
- Nearly half of the newly diagnosed illness cases examined involved malignant growths in the breast, colon, lung, and bronchus (lung), as well as the prostate.
- Lung cellular breakdown is the primary site of disease transmission, accounting for more disease transmission in Canadians than the combined incidence of the other three major malignant growth types (colorectal, breast, and prostate). Despite this major influence, during the past 35 years there has been a large decline in the mortality rate of men from cellular breakdown in the lungs, which has contributed to a decline in the mortality rate of men from all combined tumors.
- Female lung cellular breakdown passing rates have recently begun to fall. Malignant growth passage rates have reduced by 37% for men and 22% for women since their high in 1988, owing to advances in the treatment of lung and other tumors.
- The cancer profile will help with Ontario's ongoing cancer and sickness burden. It will aid in procedure oversight and distinguishing between clinical and exam administration requirements. By comprehending the report, the government may create goals to eradicate specific diseases. For example, the Canadian government has set a target of eliminating cervical disease by 2040.

The motivation of the project is:

- Cancer was the second greatest cause of death in the United States in 2020, after only heart disease.
- There will be 602,350 cancer deaths in 2020, with 284,619 females and 317,731 men.
- At least one third of common cancers are preventable.
- Millions of lives might be saved each year by implementing resource-appropriate preventive, early diagnosis, and treatment techniques.

The analysis will be done using EDA methods in Tableau. The results of the analysis will be presented in the form of tables, graphs, and charts. The findings will be used to provide light on the many profiles depicted on the television show Cancer, as well as how those characteristics affect patients.

Related Work

- This project was an analysis of Cancer's patient profiles and their certain characteristics from cancer. The data is broken down into 7 different Excel files according to Cancer's profiles and will be 602,350 cancer deaths in 2020, with 284,619 females and 317,731 men. The data most importantly reveals whether or not the patients are more likely to tackle the different levels of cancers. The objective of the project was to understand the factors that influence a patient's cancer profiles by different insights and analysis.
- This project's intended audience may include any of the following individuals or groups:
 - Any medical school / institution
 - Any medical student
 - Any analysis student
 - Anyone interested in determining insights into all types of cancer.
- The analysis was carried out in Tableau using multiple fields from the dataset to generate several graphics. The data was cleansed, and the results revealed that characteristics such as gender, age, and the percentage of cancer spots differed by area.
- The successful outcome will be:
 - The cancer profile will aid in addressing Ontario's ongoing cancer disease burden.
 - It will aid in the management of strategies and the identification of priorities for clinical and research services.
 - By understanding the report, government can set their goals to eliminate respective cancer. For example, Canada government has been set target to wipe out cervical cancer by 2040.

Methods

In this section you should outline the steps/approaches taken that lead to the results you will report in the next section. This is where you should include data sources, exploratory data analyses steps, data preprocessing methods, etc. There should be enough detail in here so that anyone reading the report would be able to replicate what you have done.

Following methods and tools will be used to execute this project:

- **Data Collection**

There are several methods for gathering data for Project Cancer Trends/Profiles in Ontario. Utilizing surveys and questionnaires is one strategy. People who have been diagnosed with cancer or who have had cancer in the past can fill out surveys and questionnaires. You can inquire about their cancer experience, diagnosis, and treatment. Profiles of cancer trends in Ontario may then be made using this data.

Utilizing secondary data sources, such as regional cancer registries and databases, is an additional method of data collection. These databases include details on cancer diagnosis, therapies, and results. This information can be used to monitor long-term trends in cancer diagnosis and treatments.

Lastly, information may be gathered through focus groups and interviews with medical professionals and researchers who are studying cancer-related topics. These focus groups and interviews can offer insightful information on current patterns in cancer diagnosis and treatment in Ontario.

- **Data Cleaning**

Data cleansing is a crucial stage in the process of evaluating and comprehending the data for Project Cancer Trends/Profiles in Ontario. It entails locating and fixing mistakes, discrepancies, and missing values in the data as well as making sure the data is in an analysis-ready format.

It is crucial to first review the data for any flaws, contradictions, and missing numbers. This may be achieved by manually going over the data and searching for any patterns or irregularities. Additionally, errors can be found by performing statistical analyses on the data. To find any extreme values in the data, for instance, an outlier test might be applied. Correcting mistakes and inconsistencies is crucial when they have been found. Either manually entering the appropriate numbers or employing data cleaning software to automatically fill in the blanks can be used to do this.

- **Exploratory Data Analysis**

EDA, or exploratory data analysis, is a crucial stage in the data analysis process since it aids in finding patterns and connections in the data. Additionally, it can shed light on the data's underlying structure and aid in the detection of potential bias or outliers. EDA aims to increase knowledge of the data collection and pinpoint areas that require more research.

EDA can assist in identifying patterns in the incidence and mortality rates of several cancer types in Ontario for Project Cancer Trends/Profiles in Ontario. Finding any disparities in the incidence and mortality of various cancer types by gender, age, and other demographic characteristics might also be useful. The initial step in EDA is to produce data visualizations. Charts of many kinds, including bar graphs, line graphs, and scatter plots, may be used to do this. Data patterns and trends, such as variations in the incidence and death rates over time, may be found using these visualizations. Finding connections between various factors is the next stage. Correlation analysis, which may be used to find any connections between various factors like age and cancer type, can be used to do this.

- **Data Visualization**

The visualization for this project can include a variety of charts, graphs, and maps.

Charts:

- Bar Chart – This chart can be used to compare the number of cancer cases in different parts of Ontario over time.
- Pie Chart – This chart can show the proportions of different types of cancer in Ontario over time.
- Stacked Bar Chart – This chart can be used to compare the number of cancer cases in different age groups in Ontario.

Graphs:

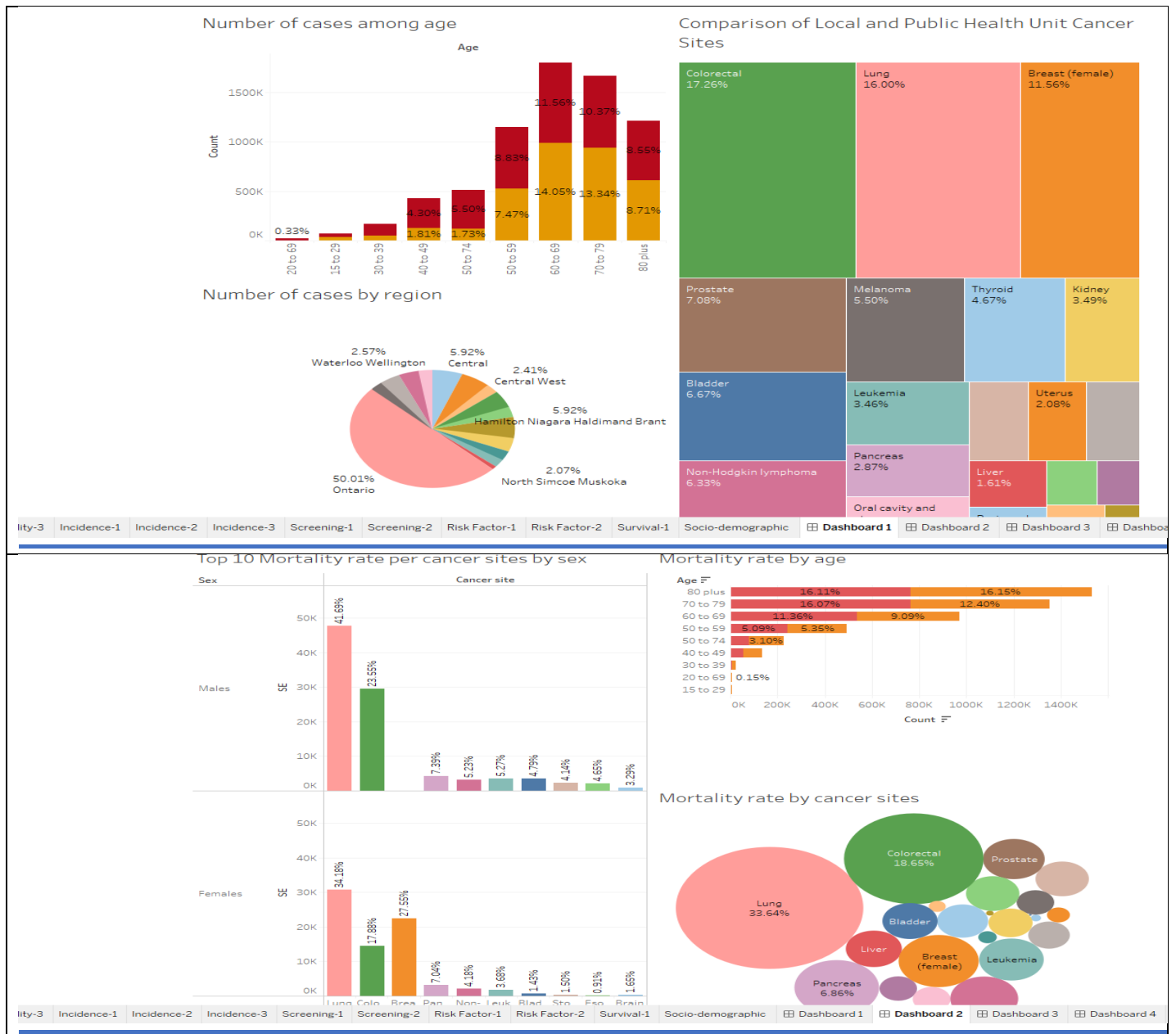
- Tree Plot – This graph can illustrate the relationship between the number of cancer cases and the population size in different parts of Ontario.
- Area Graph – This graph can be used to compare the number of cancer cases in different parts of Ontario over time.

Maps:

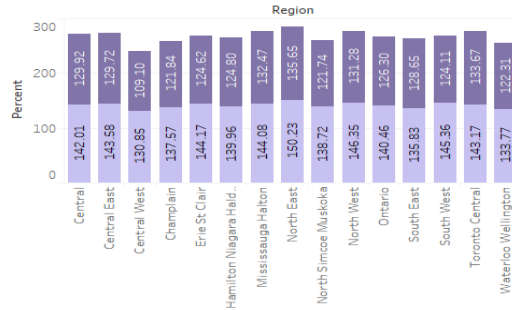
- Heat Map – This map can show the variation in the number of cancer cases in different parts of Ontario.
- Bubble Map – This map can be used to compare the number of cancer cases in different parts of Ontario.

Results

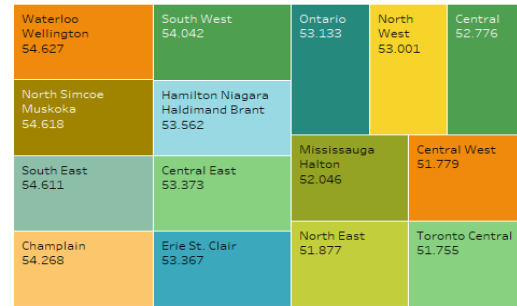
- The Ontario Cancer Profiles project seeks to give a complete picture of cancer in Ontario. My group's contribution to this project would be to create an interactive visual dashboard with extensive information about the province's cancer incidence, trends, and results. This dashboard would feature maps and graphs to display the data, allowing users to easily comprehend the information and act, and our results are shown below in the form of a dashboard.



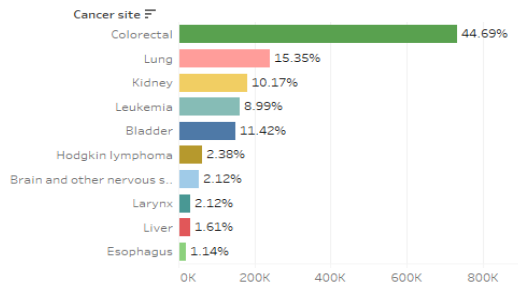
Rate of Cancer Prevalence by Region



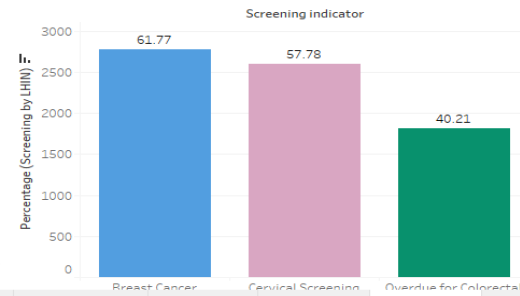
Cancer Screening



Top 10 Cancer sites ratio during year 2018

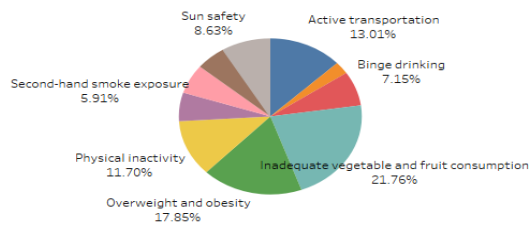


Total rate of screening indicator

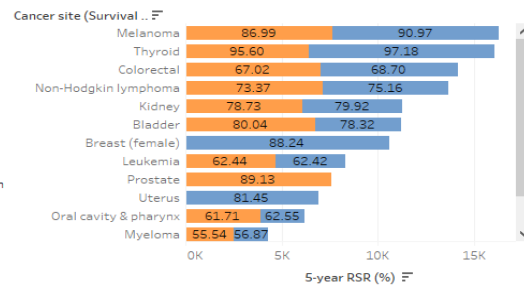


lity-3 Incidence-1 Incidence-2 Incidence-3 Screening-1 Screening-2 Risk Factor-1 Risk Factor-2 Survival-1 Socio-demographic Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4

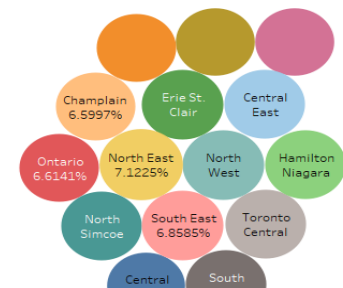
Affected factors for cancer



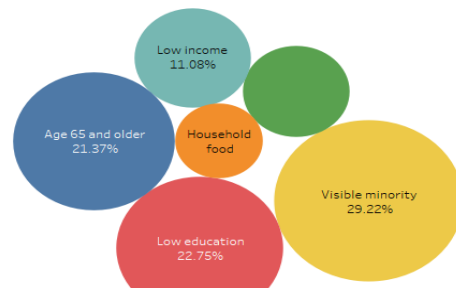
Top 15 Survival rate from cancer sites(Average of 5 year)



Survival rate for both gender



Average ratio of socio-demographic factors



lity-3 Incidence-1 Incidence-2 Incidence-3 Screening-1 Screening-2 Risk Factor-1 Risk Factor-2 Survival-1 Socio-demographic Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4

Conclusion/Insights

“Prevention is better than cure”

- As we all know, there are over 100 different types of cancer that may affect individuals.
- If you only search for a type of cancer, you'll get a long list containing every letter of the alphabet except J, Q, X, and Z.
- Our doctors and researchers are working hard to treat all types of cancer, but we need full-body scans to detect and prevent such diseases for our own safety.

Contributions

- Every member of the group was present during the dataset selection process. Following a careful evaluation of several data sets, we have decided to focus on the datasets listed above that have already been integrated into a single.csv file.
- After discussing thoughts and comments on the issue, each group member creates a project proposal.
- In accordance with the project idea, we divided the datasets among the two participants, who subsequently cleaned the data.
- After a brief break, three people work on the visualization component.
- The presentation was made by two persons, while the report was created by others.
- Each team member contributed the same amount of time to this project.

References

- <https://open.canada.ca/data/en/dataset/b89ab9d1-bddc-4baa-9133-34a446623c5b>
- https://www.who.int/health-topics/cancer#tab=tab_3
- <https://www.kaggle.com/search?q=cancer>
- <https://www.cancercareontario.ca/en/cancer-facts/ontario-cancer-profiles-lets-youexplore-cancer-statistics>
- <https://public.tableau.com/app/discover>
- <https://www.worldcancerday.org/what-cancer>