Wait Time Analysis for Priority Procedures in Canada

Group Name
Databricks

Group Members
Riya Bennychan(815633)
Karthick Raja Dhanuskodi(8116651)
Melvin Kozhikkadan Jolly(0801117)

Contents

Contents

1. Introduction

The wait times for priority procedures in Canada are a critical aspect of the healthcare system, directly impacting patient outcomes and experiences. As the demand for healthcare services continues to rise, understanding and addressing wait times have become paramount in ensuring timely access to care.

This report aims to analyze and interpret the data tables provided by the Canadian Institute for Health Information (CIHI) regarding wait times for priority procedures across various provinces in Canada. The data encompasses essential procedures such as hip and knee replacement, hip fracture repair, radiation therapy, cataract surgery, and others.

CIHI collects and reports wait time data from multiple sources, including administrative databases and submissions from provincial ministries and agencies. The data provides insights into median and 90th percentile wait times, procedure volumes, and the percentage of patients receiving care within established benchmark time frames.

Over the years, CIHI has been consistently reporting on wait time data for procedures completed between April and September, aiming to improve the timeliness of reporting while ensuring representativeness throughout the year. Moreover, additional data periods have been collected, particularly during the COVID-19 pandemic, to address the evolving healthcare landscape.

Through this report, we seek to delve into the trends, variations, and challenges associated with wait times for priority procedures in Canada. By understanding the underlying factors influencing wait times, stakeholders can collaborate to implement strategies aimed at reducing wait times, enhancing healthcare delivery, and ultimately improving patient outcomes.

As we navigate through the data tables and accompanying notes, we aim to gain comprehensive insights into the dynamics of wait times for priority procedures, contributing to informed decision-making and policy development within the Canadian healthcare system.

2. Methods

In this section, we will outline the key steps and methods employed in our analysis, including exploratory data analysis, preprocessing techniques, machine learning model selection, and evaluation metrics for our dataset.

2.1 Importing Libraries, Packages, and Dataset:

Initially, we imported the necessary libraries and packages required to execute specific sections or all components of the code in our Python notebook (ipynb file). Subsequently, we loaded our dataset and stored it in a data frame to facilitate further processing.

2.2 Exploratory Data Analysis:

We conducted a thorough exploration of the dataset by examining its contents and understanding the variables contained within the data frame. This allowed us to gain insights into the types of columns present and determine which specific columns would be pertinent for predictive modeling. Essential functions such as head, tail, describe, and info were utilized to ascertain the features within the data, including data types, non-null counts, overall counts, means, standard deviations, minimum and maximum values, as well as specific quartiles. Additionally, we investigated the distribution of both numerical and categorical features within the dataset to glean insights visually.

```
data.head(5)
```

	Reporting level	Province/territory	Region	Indicator	Metric	Data year	Unit of measurement	Indicator result
0	Provincial	Alberta	NaN	Bladder Cancer Surgery	50th Percentile	2008	Days	NaN
1	Provincial	Alberta	NaN	Bladder Cancer Surgery	90th Percentile	2008	Days	NaN
2	Provincial	Alberta	NaN	Bladder Cancer Surgery	Volume	2008	Number of cases	NaN
3	Provincial	Alberta	NaN	Breast Cancer Surgery	50th Percentile	2008	Days	NaN
4	Provincial	Alberta	NaN	Breast Cancer Surgery	90th Percentile	2008	Days	NaN

```
data.tail(5)
```

	Reporting level	Province/territory	Region	Indicator	Metric	Data year	Unit of measurement	Indicator result
15759	Provincial	Saskatchewan	NaN	Prostate Cancer Surgery	Volume	2022	Number of cases	93.00
15760	Provincial	Saskatchewan	NaN	Radiation Therapy	% Meeting Benchmark	2022	Proportion	97.91
15761	Provincial	Saskatchewan	NaN	Radiation Therapy	50th Percentile	2022	Days	7.00
15762	Provincial	Saskatchewan	NaN	Radiation Therapy	90th Percentile	2022	Days	20.00
15763	Provincial	Saskatchewan	NaN	Radiation Therapy	Volume	2022	Number of cases	1293.00

2.3 Data Preprocessing:

Data preprocessing is a crucial step used to refine raw data into a format that is suitable for analysis and modeling. Raw data often contains inconsistencies and missing values that can lead to errors in subsequent analyses, making preprocessing essential. During this stage, data is cleaned to retain only relevant information, ensuring that it can be effectively utilized by machine learning models.

i. Handling Null Values:

Our data preparation process began with an assessment of null or missing values in each column of the dataset. Utilizing the method .isnull().sum(), we identified null values in two columns: Region (10652) and Indicator result (3386). To address this, we opted to drop the entire Region column and removed rows containing null values for the indicator result.

ii. Handling Unique Values in Metric Column:

Upon examining the Metric column, we noticed inconsistencies in the representation of similar metrics. Variations such as '% Meeting Benchmark', '50th Percentile', '90th Percentile', 'Volume', '50th percentile', '90th percentile', and '% meeting benchmark' were present. These inconsistencies were rectified to ensure uniformity and clarity in the dataset.

iii. Addressing Unique Values in Data Year Column:

The Data year column contained a mix of numerical values representing years (ranging from 2008 to 2022) and alphanumeric values indicating fiscal years or specific quarters ('2019FY', '2019Q3Q4', '2020FY', '2020Q3Q4', '2021FY', '2021Q3Q4'). To maintain consistency, these anomalies were resolved during the data preprocessing phase.

iv. Changed the indicators with hours to Days

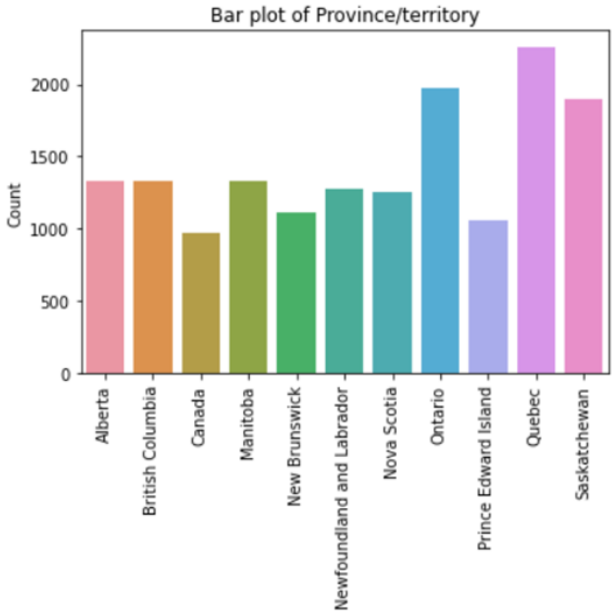
To ensure consistency across the metrics, we've converted the indicators from hours to days. This change allows for a unified measurement approach across all metrics, facilitating easier comparison and analysis.

After implementing the cleaning steps, the dataset now consists of 15,764 rows and 7 columns. These cleaning procedures have effectively addressed null values, standardized the representation of metrics and data years, and ensured data integrity and consistency. With the dataset prepared and streamlined, it is now ready for further analysis and modeling tasks, setting the stage for insightful exploration and meaningful insights.

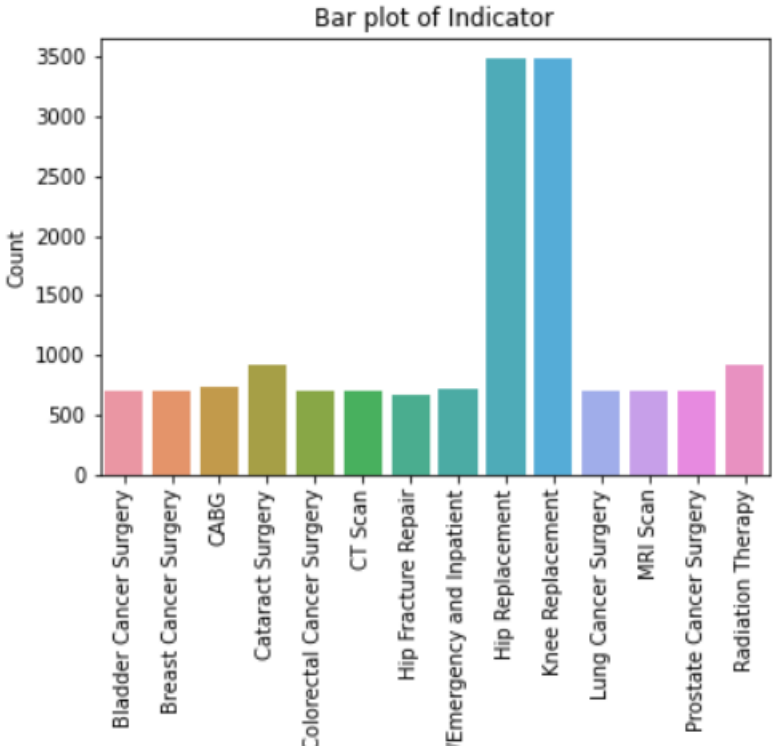
2.4 Exploratory Data Analysis (Visualization):

i. Distribution of Health Procedures by Province

The plot below illustrates the distribution of overall health process counts across different provinces in Canada. Quebec has the highest number of reported cases, followed by Saskatchewan, with Ontario ranking third. The data includes reporting at both the regional and national levels, which is why Canada appears to have the lowest count of reported cases. Provinces with the least reported cases are New Brunswick and Prince Edward Island.



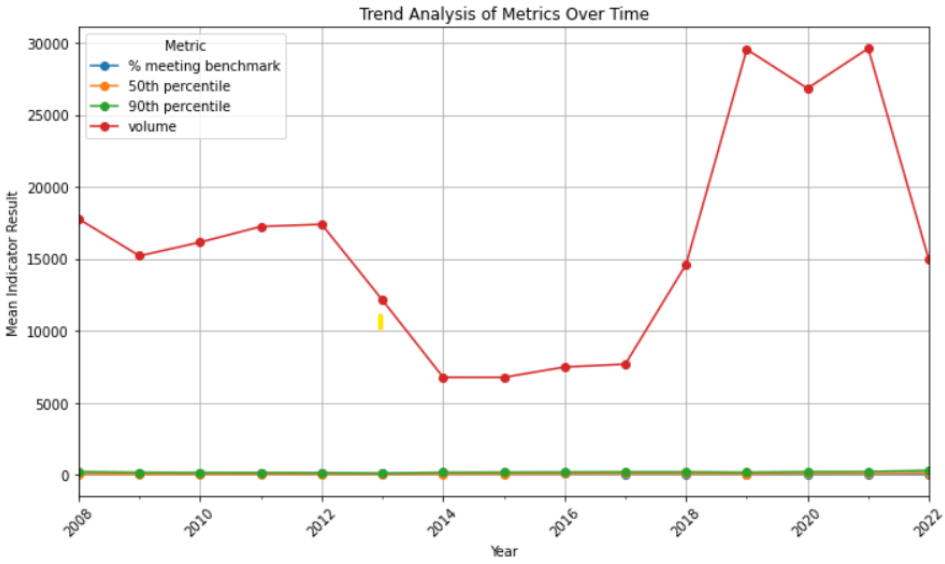
ii. Distribution of Waiting Times for Various Medical Procedures in Canada



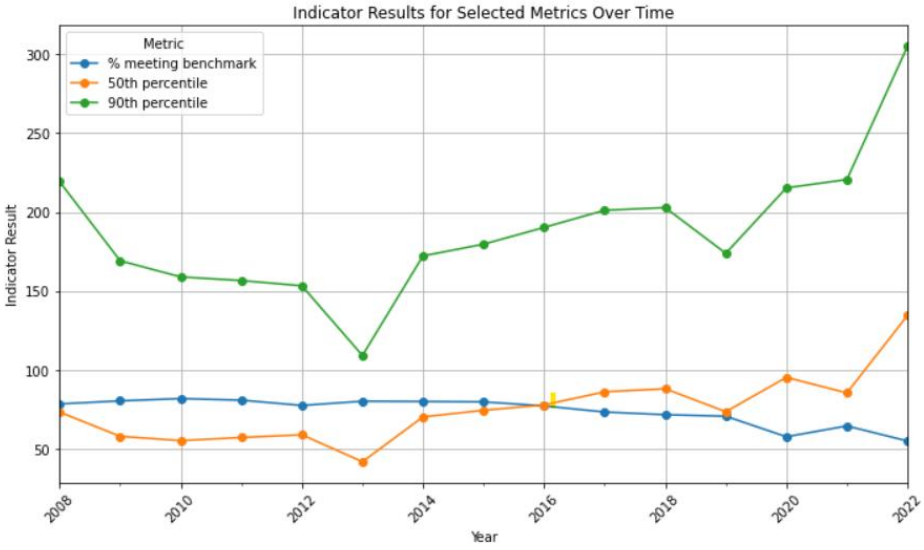
The plot illustrates the distribution of waiting times for different types of medical procedures across Canada. It reveals that the highest number of individuals are awaiting hip replacement and knee replacement surgeries, indicating a significant demand for these procedures. Conversely, the least number of individuals are waiting for hip fracture surgery, suggesting relatively lower demand or quicker access to this specific procedure compared to others.

i. Trend analysis of different metrics over time

The analysis of the volume metric in the provided graph titled "Trend Analysis of Metrics Over Time" reveals interesting trends, particularly in the context of the COVID-19 pandemic. This metric reflects the number of priority procedures conducted in Canada over the years from 2008 to 2022. Notably, there is a noticeable increase in the volume of procedures, especially during the years 2019 to 2021, coinciding with the peak period of the COVID-19 pandemic. While the volume starts at a relatively low level, it demonstrates a gradual upward trend over the years. However, the most striking observation is the sharp increase in volume recorded during the years, indicating a substantial surge in the number of priority procedures performed. This surge likely reflects various factors, including pent-up demand due to postponed procedures during the pandemic, increased healthcare utilization, and efforts to address backlogs in medical services. Overall, the trend underscores the dynamic nature of healthcare delivery and the impact of external factors such as public health crises on healthcare service utilization.

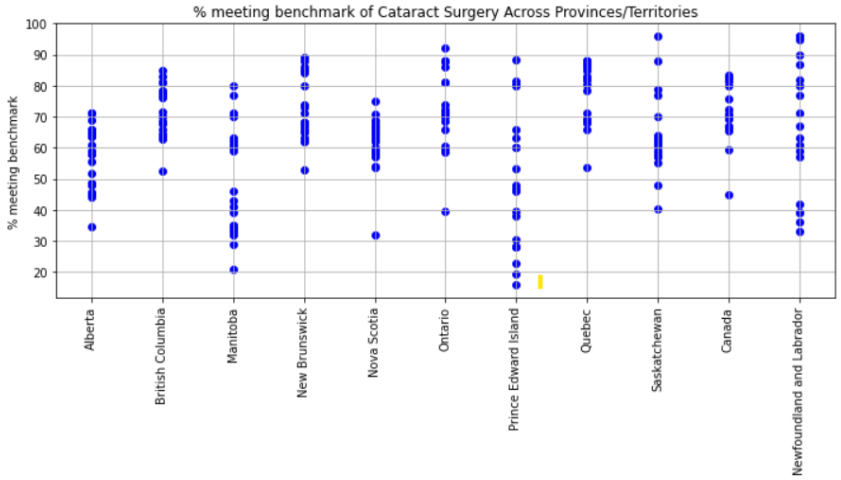


The plot unveils significant insights into the trends of three key metrics: "% meeting benchmark," "50th percentile," and "90th percentile." A notable observation is the decline in the percentage meeting the benchmark from 2020 to 2022, indicative of reduced procedure frequency during the pandemic, leading to increased patient wait times. Moreover, both the "50th Percentile" and "90th Percentile" exhibit fluctuating trends, with decreases observed in 2019. These fluctuations reflect variations in median and 90th percentile values over time, influenced by factors like resource availability and patient demand.



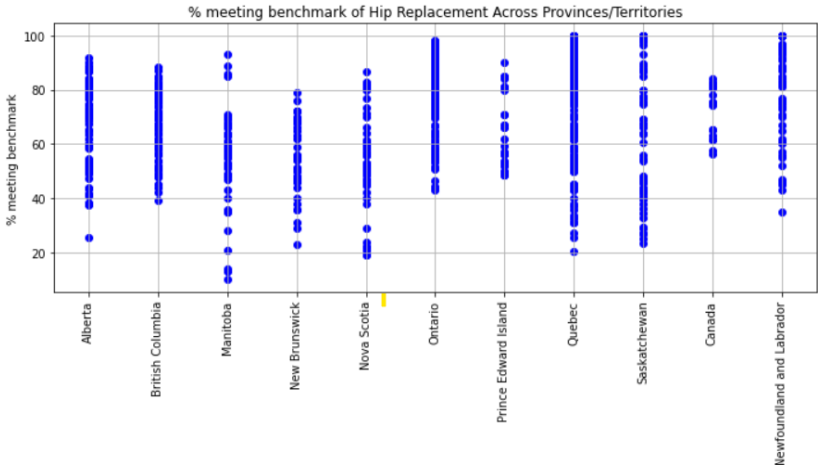
3. Results

3.1 percentage meeting benchmark of Cataract Surgery



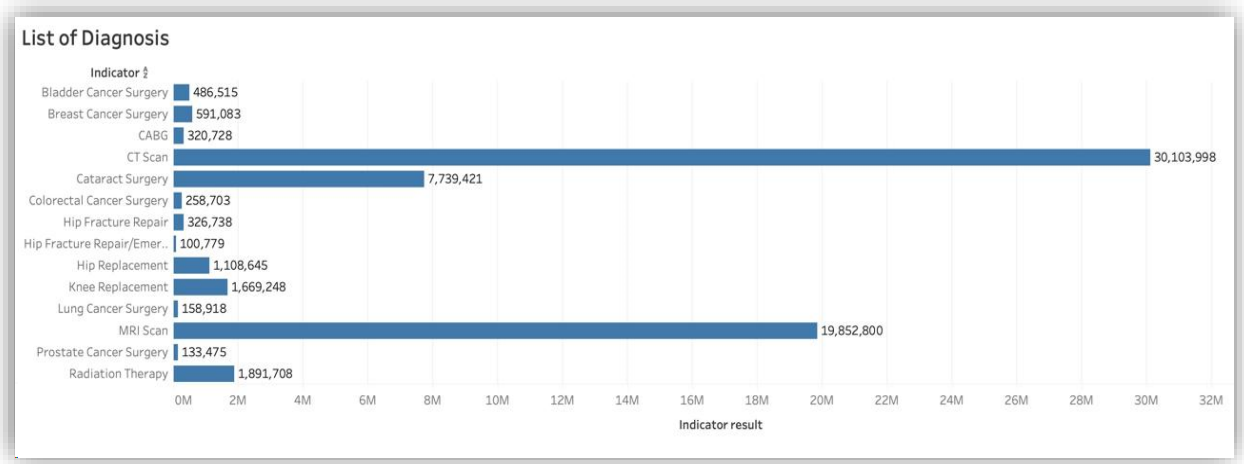
The scatter plot titled "% Meeting Benchmark of Cataract Surgery Across Provinces/Territories" provides insight into the percentage of cataract surgeries meeting benchmarks in various provinces and territories of Canada from 2008 to 2022. Notably, British Columbia, Quebec, and New Brunswick consistently maintain benchmarks above 50%, indicating relatively high performance in meeting targets. Conversely, Manitoba and Prince Edward Island exhibit lower performance rates, consistently falling below the 20% benchmark threshold. Ontario and New Brunswick display a pattern of consistently striving to meet benchmarks, often achieving values closer to the higher end. This variability among provinces and territories underscores the diverse performance trends observed across different regions of Canada over the specified time period.

3.2 Percentage meeting benchmark of Hip replacement surgery

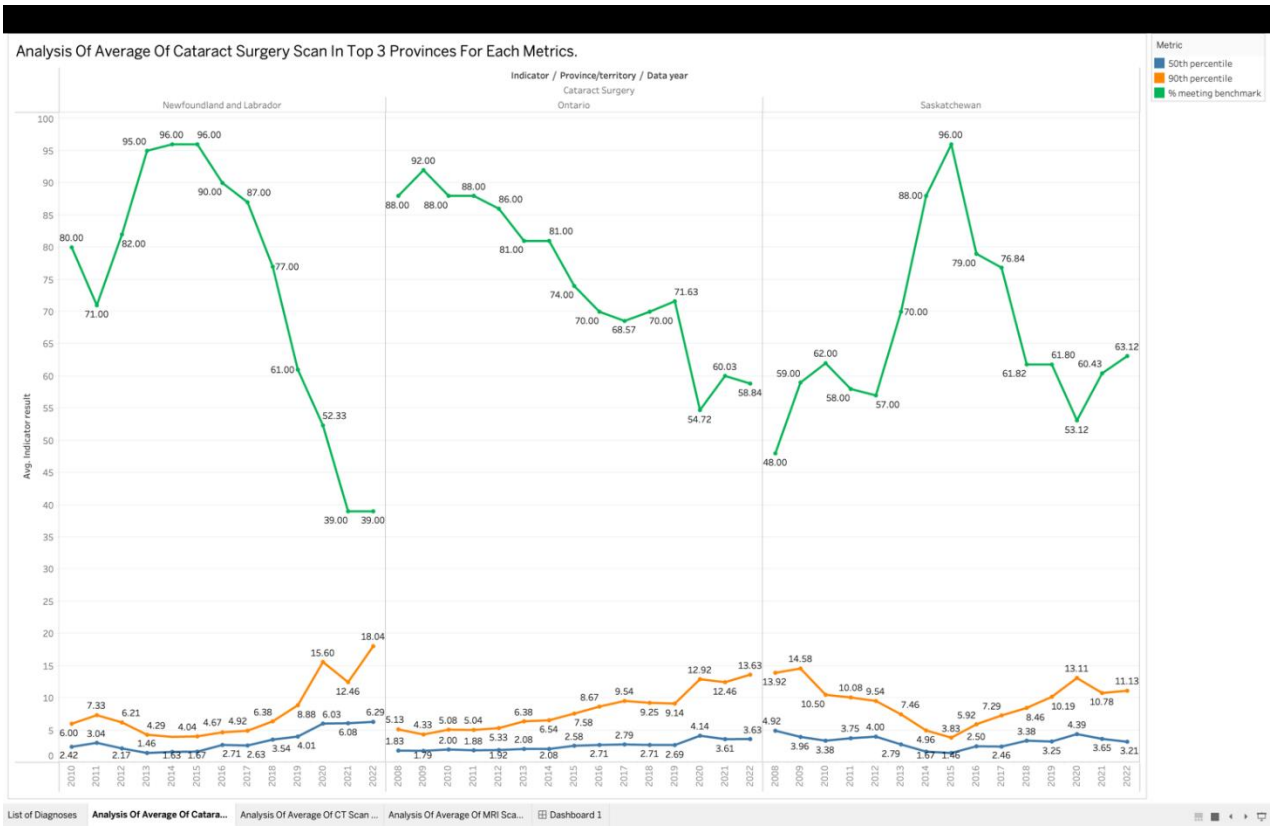


The data analysis highlights distinct trends in meeting benchmark percentages for hip replacement surgery across different provinces in Canada. Ontario consistently demonstrates strong performance, with benchmark completion rates never falling below 40% and often reaching 100% in various years, indicating a consistent and high level of achievement in completing surgeries within the established benchmarks. In contrast, provinces like Quebec, Saskatchewan, and Newfoundland and Labrador exhibit fluctuating trends, with completion rates varying from as low as 20% in some years to achieving 100% in others. Manitoba, on the other hand, consistently shows lower performance, with completion rates occasionally dropping below 10% in certain years, suggesting challenges in meeting benchmark targets. In contrast, British Columbia consistently strives to meet benchmarks, maintaining completion rates above 40% throughout the analyzed period, reflecting consistent efforts and achievements in completing hip replacement surgeries within established benchmarks.

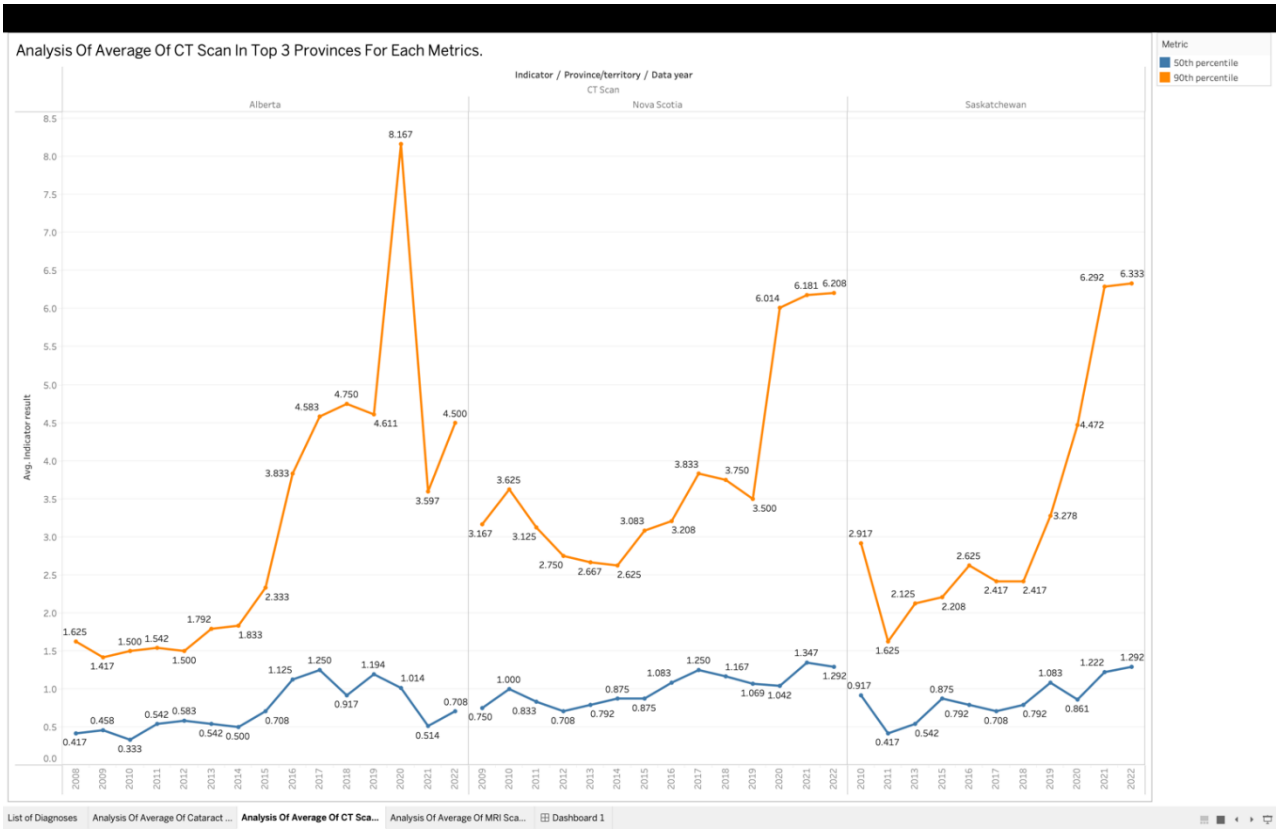
3.3 Tableau Dashboard



This chart depicts the wait time procedures for a list of indicators, showcasing the volume of surgeries, scans, therapies pending or completed each year. It provides a comprehensive overview of the total volume of indicator results for each indicator present in the dataset.



In analyzing the top three indicators performed across nine different provinces in Canada, we focused on the top three provinces. The green line in the plot represents the percentage meeting benchmark for each diagnosis, with Newfoundland and Labrador and Saskatchewan averaging 96.00% in 2015. However, in 2020, amid the COVID-19 pandemic, this percentage decreased significantly by around 52.00% in these provinces. Concurrently, the 50th percentile, indicating the point where half of the patients are treated on time, saw an increase from 1.67 in 2015 to 6.08 in 2020. This rise suggests a notable increase in average wait times between these years.



So, from the list of indicators across total of nine different provinces of Canada, we analysed Top 3 indicators performed from top 3 provinces in Canada. In this plot the orange line indicates 90th percentile that is out of 100 patients 90 patients have been treated on time. Here we took top 3 provinces performing CT scan in Canada during 2018 in Alberta the average has been peaked to 8.16. whereas in 2020 it faces a sudden decrease in the average of around 3.5 and the 50th percentile created an impact because of the 90th percentile in 2020 so it clearly defines the wait times have been increased.

4. Discussion

The results obtained from our analysis provide valuable insights into wait times for priority procedures in Canada and their implications for healthcare delivery. Overall, we achieved our objective of exploring trends, variations, and challenges associated with wait times across different provinces and territories. Our results shed light on the performance of healthcare systems in meeting benchmarks for various procedures and highlight the impact of external factors such as the COVID-19 pandemic on wait times.

In comparing our results with existing literature, we find alignment with previous reports regarding the challenges faced in meeting benchmark targets, particularly during periods of increased demand or resource constraints. Our findings corroborate the notion that wait times for priority procedures can vary significantly across regions and are influenced by factors such as population density, healthcare infrastructure, and funding allocation.

However, we also encountered some failures and challenges along the way. One challenge was handling inconsistencies and missing values in the dataset, particularly in the Metric and Data Year columns. This required careful data preprocessing and cleaning to ensure the integrity and accuracy of our analysis. Additionally, interpreting the impact of the COVID-19 pandemic on wait times posed challenges due to the dynamic and evolving nature of the healthcare landscape during the pandemic.

To overcome these challenges, we employed rigorous data cleaning and preprocessing techniques, including removing null values, standardizing metric representations, and converting indicators to a consistent measurement unit. We also conducted thorough exploratory data analysis and visualization to gain insights into the trends and patterns present in the data. Collaborative efforts among team members facilitated problem-solving and decision-making throughout the analysis process.

5. Conclusion

In conclusion, our analysis provides valuable insights into wait times for priority procedures in Canada, highlighting the variability and challenges faced across different regions. Moving forward, there are several avenues for improvement and future directions. Firstly, continued monitoring and reporting of wait time data are essential to track performance trends and identify areas for improvement within the healthcare system. Additionally, targeted interventions and resource allocation strategies may be necessary to address disparities in wait times and improve access to care for all Canadians. Furthermore, leveraging advanced analytics techniques such as machine learning and predictive modeling could enhance our ability to forecast and manage wait times more effectively. Overall, our work underscores the importance of ongoing research and collaboration in addressing wait time challenges and improving healthcare delivery in Canada.

6. Contributions

Name	Contribution
Riya Bennychan	• Conducted thorough data exploration to understand dataset structure, identify patterns, and check for missing values and outliers. • Implemented data cleaning procedures, including addressing inconsistencies, removing duplicates, and handling missing values, to ensure dataset quality for analysis. • Attempted to create predictive models, including a regression model, to predict or estimate the value of the dependent variable based on independent variables in the dataset.
Melvin K Jolly	• Utilized data exploration techniques to gain insights into dataset characteristics, distributions, and potential relationships between variables. • Created visually appealing presentations to effectively communicate data insights and findings to stakeholders, team members, or clients. • Experimented with modeling techniques such as ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal AutoRegressive Integrated Moving Average) to build predictive models for time-dependent data.
Karthik Raja	• Explored dataset through statistical analysis and visualization to uncover key trends, patterns, and outliers. • Leveraged Tableau to design interactive visualizations and dashboards that present data insights in an accessible and engaging manner. • Contributed to data exploration and visualization efforts to support understanding of dataset characteristics and findings.

7. References

Dataset link: https://www.cihi.ca/sites/default/files/document/wait-times-priority-procedures-in-canada-2023-data-tables-en.xlsx

Cover image : https://blog.nexa.com.au/blog/1-challenges-communicating-wait-times

8. Appendices

- **Databricks_final_report.pdf**: Contains final report.
- **Databricks_python.ipynb** : The provided code includes all the necessary instructions for our project, covering tasks such as importing libraries, preprocessing data, conducting exploratory data analysis, implementing machine learning algorithms.
- **Databricks_dashboard.twbx**: Tableau file includes our dashboard.