

Analysis and Predicting Wait Times for Priority Procedures in Canada: Part 3 Report

By

Aye, Thanda

Bui, Ngoc Mai Khanh

Naw Mu Aye, Naw Mu Aye

Pawar, Pratiksha Ravindra

An assignment submitted to

Dr. Omid Isfahanialamdari

Faculty of Master of Data Analytics

In Fulfillment of the Requirements for the Course

DAMO-611-2: Data Analytics Case Study 3

Chapter 1: Introduction

Timely access to healthcare is essential for patient well-being and the financial sustainability of public health systems. Yet, wait times for priority procedures remain a persistent challenge in many developed countries, including Canada. These delays can increase patient anxiety, worsen health outcomes, and drive up both direct and indirect healthcare costs. Prolonged waits often lead to more complex interventions, lost productivity, and inefficient use of healthcare resources.

This project aims to develop a predictive model for wait times, procedure volumes, and benchmark attainment for priority procedures in Canada, using comprehensive data from the Canadian Institute for Health Information (CIHI). By accurately forecasting these metrics, the model will support healthcare administrators and policymakers in making proactive, fiscally responsible decisions that improve service delivery, enhance patient outcomes, and increase system-wide efficiency.

1.1 Background

Canada's publicly funded healthcare system continues to face challenges in providing timely access to care, with wait times for priority procedures remaining a focal point of policy debate. These delays carry significant financial consequences, straining provincial budgets, increasing hospital operational costs, and reducing overall economic productivity.

The Canadian Institute for Health Information (CIHI) provides critical data on wait times, procedure volumes, and benchmark attainment—key resources for evaluating system performance and identifying opportunities for improvement, particularly from a cost-efficiency standpoint.

Research highlights the complex, multi-level factors driving healthcare delays, many with financial implications. Bodenheimer and Pham (2010) link access issues to provider

capacity, affecting operational costs. Murray and Berwick (2003) examine supply-demand dynamics in primary care, proposing strategies to improve resource utilization. Astuti et al. (2022) identify cost drivers in hospital prescription wait times, including staffing, infrastructure, and information systems.

While these studies offer valuable insights, there remains a need for predictive models that not only estimate future wait times but also quantify their economic impact. This project seeks to address that gap by developing advanced models that support more accurate resource allocation and improved financial planning within Canada's healthcare system.

Chapter 2: Problem Statement and Research Questions

2.1 Problem Definition

This project tackles the dual challenge of predicting wait times and assessing their financial impact for priority procedures in Canada. Despite the value of historical data, there's a lack of robust models that forecast future wait times, volumes, and benchmarks while quantifying the economic costs of current management practices. Prolonged delays lead to inefficient resource use, higher healthcare costs from worsening conditions, and broader economic losses. Using CIHI data, we aim to build predictive models that forecast trends and reveal the financial risks of delays, supporting proactive, cost-effective decisions for a more sustainable healthcare system.

2.2 Research Question

The following research questions will guide the analysis, considering the financial dimensions where applicable:

- 1. Does the calendar year have a significant influence on wait times for priority procedures?**

Justification: This analysis aims to determine if there is a temporal trend in wait times, which carries **significant long-term financial implications** for healthcare planning, resource allocation, and overall system sustainability.

2. **Do average wait times for priority procedures vary significantly across different provinces and different procedure types?**

Justification: This two-part analysis aims to identify geographic and procedural disparities in healthcare access. Understanding these variations is critical for enabling targeted interventions, ensuring equitable service delivery, and optimizing spending to achieve cost-effectiveness for diverse medical needs across different regions.

Is there a statistically significant correlation between procedure volume and the 90th-percentile wait days for priority procedures?

Justification: This analysis seeks to understand how demand (volume) interacts with extreme wait times. This insight is essential for optimizing resource allocation and service capacity management, directly impacting operational costs and the potential financial burdens arising from prolonged waits and unforeseen complications due to high demand.

3. **Is there a statistically significant correlation between procedure volume and the 90th-percentile wait days for priority procedures?**

Justification: This analysis is to understand the impact of procedure volume on extreme wait times and to optimize resource allocation and service capacity management.

4. **Can a multivariate model accurately predict wait times for priority procedures, and how can these predictions be used to estimate the financial cost of delays across provinces and procedure types?**

Justification: This question directly addresses the project's core predictive and financial objectives. By accurately forecasting wait times, the model enables the estimation of financial costs associated with delays (e.g., through a 'predicted wait \times cost/day' proxy). This allows for understanding where the most significant financial burdens lie, providing a concrete business case for interventions and informing strategies for cost mitigation in specific provinces and for particular procedures.

5. Can a multivariate model accurately predict procedure volume and benchmark performance (% meeting target wait times), and how can these predictions be used to estimate financial risk across provinces and procedure types?

Justification: This question extends the predictive capability to key operational metrics. By forecasting volume and benchmark attainment, the model facilitates the estimation of financial risk (e.g., using a ' $(1 - \% \text{benchmark}) \times \text{volume} \times \text{penalty}$ ' proxy) related to unmet targets and service delivery challenges. This provides a forward-looking tool for identifying high-risk areas, guiding resource allocation to prevent future financial liabilities, and ensuring accountability for effective healthcare management.

2.3 Justification and Impact

This project is justified by the critical need for effective strategies to manage and reduce wait times for priority procedures in Canada, viewed through a **comprehensive financial lens**. The ability to accurately predict not only future wait times but also associated operational metrics and their **economic consequences** offers several transformative benefits:

- **Financial Prudence & Resource Efficiency:** Forecasting demand and delays enable proactive resource allocation, minimizing bottlenecks and idle capacity. This optimizes operational expenditure and averts costs from prolonged patient stays or complications.
- **Strategic Policy Planning:** Predictions and derived financial impact estimates allow policymakers to evaluate the economic return on investment for interventions, ensuring **fiscally responsible** and effective public health investments.
- **Better Patient Management:** Patients can receive more accurate estimates of their wait times, allowing them to make informed decisions and manage their expectations.
- **Mitigating Financial Burdens:** Accurate predictions help reduce direct costs (e.g., increased medication) and indirect societal costs (e.g., lost productivity) associated with prolonged waits and worsened health outcomes.
- **Improved Patient Experience & Reduced Personal Strain** Accurate wait time estimates empower patients to make informed decisions, manage expectations, and potentially lessen personal financial burdens from interim care or seeking private alternatives.
- **Performance Monitoring & Accountability:** Predicted metrics (wait times, volumes, benchmarks) coupled with financial proxies offer powerful benchmarks for evaluating the financial performance and efficiency of healthcare providers, tying the improvements to fiscal responsibility.
- **Proactive Risk Management:** Predicting financial risk from unmet targets allows proactive interventions, preventing financial penalties, patient dissatisfaction, and future burdens on the healthcare system.

2.4 Scope

This project will:

1. Proactive Risk Management: Predicting financial risk from unmet targets allows for proactive interventions, preventing financial penalties, patient dissatisfaction, and future burdens on the healthcare system.
2. Focus on key variables including Indicator (procedure type), Province/Territory, Data Year, Volume, and 90th Percentile wait times.
3. Develop multivariate predictive models for:
 - Wait times (specifically the 90th percentile).
 - Procedure volume.
 - Benchmark performance (% meeting target wait times).
4. Evaluate the models' performance using appropriate statistical techniques (e.g., MAE, RMSE, ANOVA, Mixed Linear Models, Spearman correlation).
5. Develop and apply financial proxy outcomes, such as "estimated cost of delay" and "financial risk proxy," derived from model predictions to quantify the economic implications of wait times and operational performance.
6. Provide insights into the factors influencing wait times, volume, and benchmark performance, along with their significant **economic consequences and potential for cost savings**.

This project will not:

- Include data from outside of Canada.
- Examine patient-level data beyond aggregated wait times and volumes.
- Develop specific policy recommendations regarding exact dollar amounts of investment or penalties, though the findings will provide a robust quantitative foundation for policy discussions and cost-benefit analyses.

Chapter 3: Hypothesis Formulation

In this chapter, a comprehensive set of testable hypotheses are formulated. These hypotheses are designed to address detailed research questions and guide the detailed analysis of key factors impacting wait times, procedure volumes, and benchmark performance, focus on exploring their critical financial dimensions and implications for healthcare system sustainability.

3.1 Research Question 1

Research Question: Does the calendar year have a significant influence on wait times for priority procedures?

Null Hypothesis (H_0): The Calendar Year has no statistically significant linear influence on the 90th percentile wait times ($\beta_{\text{Year}} = 0$), holding province and procedure type constant.

Alternative (H_1): The Calendar Year has a statistically significant linear influence on the 90th percentile wait times ($\beta_{\text{Year}} \neq 0$), holding province and procedure type constant.

Statistical Test: Mixed Linear Model (with t-test on regression coefficient for Calendar Year).

3.2 Research Question 2

Research Question: Do average wait times for priority procedures vary significantly across different provinces and different procedure types?

3.2.1 Variation Across Provinces

Null Hypothesis (H_0): There is no statistically significant difference in the average 50th percentile wait times across Canadian provinces/territories.

Alternative (H_1): There is a statistically significant difference in the average 50th percentile wait times across Canadian provinces/territories.

Statistical Test: One way ANOVA with Tukey post-hoc comparisons.

3.2.2 Variation Across Procedure Types

Null Hypothesis (H_0): The average 90th percentile wait times are statistically equal across different priority procedure types.

Alternative (H_1): At least one priority procedure type has an average 90th percentile wait time that is statistically different from others (specifically, certain procedures are expected to have significantly longer waits).

Statistical Test: One way ANOVA with Tukey post-hoc comparisons.

3.3 Research Question 3

Research Question: Is there a statistically significant correlation between procedure volume and the 90th-percentile wait days for priority procedures?

Null Hypothesis (H_0): There is no statistically significant monotonic correlation between procedure volume and the 90th-percentile wait days for priority procedures.

Alternative (H_1): There is a statistically significant monotonic correlation (either positive or negative) between procedure volume and the 90th-percentile wait days for priority procedures.

Statistical Test: Spearman correlation

3.4 Research Question 4

Research Question: Can a multivariate model accurately predict wait times for priority procedures, and how can these predictions be used to estimate the financial cost of delays across provinces and procedure types?

Null Hypothesis (H₀): A multivariate predictive model for wait times does not achieve significantly lower prediction errors (e.g., MAE/RMSE) than a naïve last-year baseline model, and the estimated financial cost of delays derived from these predictions does not vary meaningfully across procedures or provinces.

Alternative (H₁): A multivariate predictive model for wait times achieves significantly lower prediction errors (e.g., MAE/RMSE) than a naïve last-year baseline model, and the resulting estimated financial costs of delays reveal meaningful variation across procedures and provinces.

Statistical Test: Comparison of predictive model performance (e.g., MAE, RMSE) against a naïve baseline; analysis of variance in estimated cost of delay (Estimated cost of delay = predicted wait × cost/day).

3.5 Research Question 5

Research Question: Can a multivariate model accurately predict procedure volume and benchmark performance (% meeting target wait times), and how can these predictions be used to estimate financial risk across provinces and procedure types?

Null Hypothesis (H₀): A multivariate model does not improve prediction of procedure volume or benchmark performance, and the estimated financial risk does not vary meaningfully across procedures or provinces.

Alternative (H₁): A multivariate model improves prediction of volume and benchmark performance, and the resulting financial risk estimates vary meaningfully across procedures and provinces.

Statistical Test: Comparison of predictive model performance (e.g., MAE, RMSE) against naïve baselines (previous year's values); analysis of variance in estimated financial risk (Financial risk proxy = $(1 - \% \text{benchmark}) \times \text{volume} \times \text{penalty}$).

Chapter 4: Data Description and Preparation

This chapter details the initial steps in the analysis of wait times for priority procedures in Canada. A thorough understanding of the data is crucial before developing predictive models. This chapter outlines the data source, describes the variables, explains the preprocessing steps and presents descriptive statistics and visualizations to explore the data's characteristics, identify potential patterns, and gain insights into the factors influencing wait times. The activities performed in this chapter include data exploration, descriptive statistics, and visualization, laying the groundwork for subsequent modeling and analysis.

4.1 Data Source

The data for this project is sourced from the Canadian Institute for Health Information (CIHI). Specifically, the data is derived from the "Explore Wait Times for Priority Procedures Across Canada" dataset. CIHI is a not-for-profit organization that provides essential health information to support decision-making and improve the health of Canadians. This dataset is relevant to the project as it offers comprehensive information on wait times for a range of priority medical procedures across various geographical levels within Canada. The data includes key variables such as province/territory, procedure type, data year, and volume, which are critical for understanding and predicting wait times.

4.2 Variable Descriptions

The data set includes the wait time of priority procedures for year 2008 to 2023 for each province in Canada and the combined data such as Canada as a whole.

The dataset contains the following attributes:

- **Reporting level:** This indicates the reporting level type of category determining Provincial, National or Regional.
- **Province/territory:** This lists the 10 provinces of Canada and Canada as a whole.
- **Region:** This indicates the region type if the reporting level is region. Otherwise, not applicable.
- **Indicator:** This lists the type of procedures in waiting
- **Metric:** This indicates the type of measurement metric includes % meeting benchmark, 50th percentile, 90th percentile and Volume.
- **Data Year:** This indicates the year of measurement.
- **Unit of Measurement:** This indicates the unit of measurement.
- **Indicator result:** This describes the result of the indicator based on type of measurement (Metric) and Unit of measurement. E.g, For 90th percentile measurement and unit of measurement days, the wait time for particular procedures is average number of days in wait time that takes for 90% of the procedure.

The following table details the key variables to be used in the analysis including the type of variables and their level of measurement in the context of the wait time.

Variable	Type	Unit	Level of Measurement
Reporting level	Categorical	N. A	Categorical (Whether Regional and Provincial has any difference)
Province/territory	Categorical	N. A	Categorical <ul style="list-style-type: none"> 10 provinces of Canada for Provincial/Regional Reporting level and Canada for National Reporting Level
Region	Categorical	N. A	Categorical This categorizes regions based on each province.
Indicator	Categorical	N. A	Categorical (To measure which type of procedure takes longer time)
Data year	Continuous	Year	Continuous (To measure the trend of the wait time and to predict next year's wait time)
Indicator result	Continuous	Depending on the Metric filter	Continuous <ul style="list-style-type: none"> Days for 50th percentile and 90th percentile Number of procedures for volume Percentage for % meeting benchmark

Table 1: Details of Measurement variable Table

4.3 Data Preprocessing

4.3.1 Preprocessing Steps

1. Handling Duplicate and Missing Values

- Removed duplicate records: Ensured that each data point represented a unique observation to avoid redundancy and analytical distortions.
- Removed the "Canada" row under the Province column: This row lacked sufficient data across other columns, particularly in the Region column, making it impossible to impute a meaningful value.

- Filled missing Region values: Imputed using the mode of each corresponding Province group to ensure geographical consistency.
- Imputed missing values in Indicator result: Used the median within each group defined by Indicator, Metric, and Unit of measurement to preserve the distribution while accommodating group specific variations.
- Dropped remaining missing values in Indicator result: After imputation, any rows still missing a value in this column were removed to maintain data integrity.

2. Identifying and Handling Outliers

No outliers were detected in the dataset based on distribution checks and domain understanding. As such, no action was required in this step.

3. Data Type Conversions

Converted the “Indicator result” column to numeric. Ensured that numerical operations and statistical analyses could be accurately performed on this column.

4. Creating New Variables

No new variables were created during this preprocessing phase.

4.3.2 Justification of Preprocessing Steps:

- Removing duplicates improves data quality and prevents analytical distortions.
- Converting Indicator result to numeric enables quantitative analysis and prevents errors during computations.
- Rows with "data year" values matching the patterns 'FY' or 'Q3Q4' were excluded using a regex filter. The na=False parameter ensured missing values were treated as non-matches during filtering.
- Convert Hours value to Days in “Unit of measurement” column
- Removing the “Canada” row ensures only usable and meaningful data are retained.

- Imputing missing Region values by province mode maintains logical consistency and spatial relationships.
- Group wise median imputation for Indicator result preserves the distribution characteristics of different measurement contexts.
- Dropping residual missing values ensures the dataset is complete and ready for robust analysis.

4.4 Data Preparation

The dataset contains measure data such as 50th percentile, 90th percentile, volume and percentage for % meeting benchmark. However, not all measures are needed for individual hypothesis e.g., in hypothesis 5, the relationship between volume and 90th percentile waiting days are to be analyzed. In order to analyze the different measure separately according to the need data visualization and hypothesis testing, each type of measures is filtered and write into separate columns and exported to an excel file “transposed_selected_metrics.xlsx” using python programming. This data together with the original cleaned dataset are loaded to Tableau for data visualization.

Chapter 5: Data Understanding and Visualization

This chapter presents the findings organized around research questions and hypotheses. For each research question, include:

- Descriptive statistics
- Relevant visualizations
- Interpretation of the results

5.1 Research Question 1

This chapter is to visualize the research question “Does the calendar year have a significant influence on wait times for priority procedures?”

5.1.1 Descriptive Statistics

To observe the significance of the calendar year influence, the median and the year-over-year change of 50th percentile wait times are observed. The statistics are generated using python and are described in Figure 1.

	Data year	Avg_Wait	Median_Wait	YoY_Change_Avg	YoY_Change_Median
0	2008	38.941429	22.00	NaN	NaN
1	2009	37.498143	22.00	-1.443286	0.00
2	2010	35.844786	22.00	-1.653357	0.00
3	2011	38.164714	21.00	2.319929	-1.00
4	2012	39.097929	21.00	0.933214	0.00
5	2013	36.141143	19.75	-2.956786	-1.25
6	2014	35.711357	22.00	-0.429786	2.25
7	2015	37.512071	22.00	1.800714	0.00
8	2016	40.459286	22.00	2.947214	0.00
9	2017	43.363500	22.00	2.904214	0.00
10	2018	43.154286	23.00	-0.209214	1.00
11	2019	44.843286	22.50	1.689000	-0.50
12	2020	58.063500	20.00	13.220214	-2.50
13	2021	48.434929	22.00	-9.628571	2.00
14	2022	56.953071	24.00	8.518143	2.00
15	2023	53.501071	25.00	-3.452000	1.00

Figure 1: Medium and Year-Over-Year Values of 50th Percentile Wait Times Over Years

The data clearly demonstrates that the calendar year has a significant influence on the 50th percentile wait times. This influence is manifested through:

- **Annual variations:** Both the average and median wait times are not static but change from year to year.
- **Periods of stability and volatility:** Some years show no change in median wait times, indicating stability, while others exhibit substantial year-over-year changes in both average and median.

- **Shifts in distribution:** The differences between the year-over-year changes in the average and median wait times suggest that the underlying distribution of wait times is shifting, not just its central tendency. This indicates that the impact of the calendar year goes beyond simple increases or decreases, affecting the overall spread or skewness of the wait time data.

5.1.2 Visualization

To observe variability over the years, the line chart of average/median wait times by Data year, with error bars with interquartile range are plotted using python and it can be observed in Figure 2.

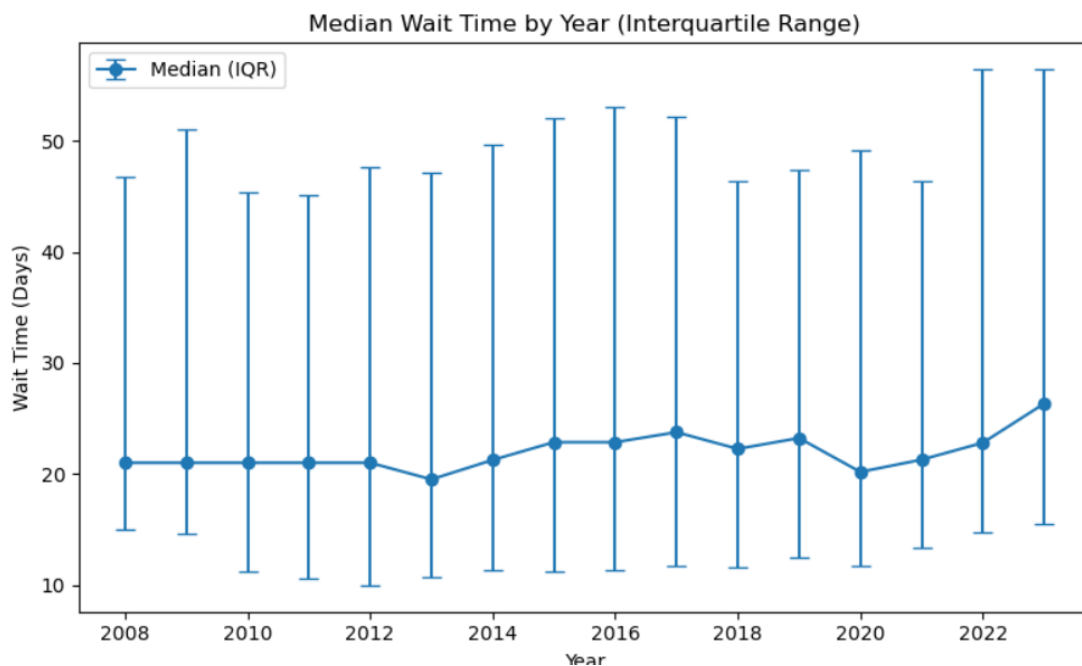


Figure 2: Error Bar Plot of Year Median with Interquartile Range for 50th Percentile Wait Time

To visualize the trends for different procedures, a heatmap is plot in python for median 50th Percentile wait time values for each procedure type over the years and Figure 3 shows the result heatmap diagram.

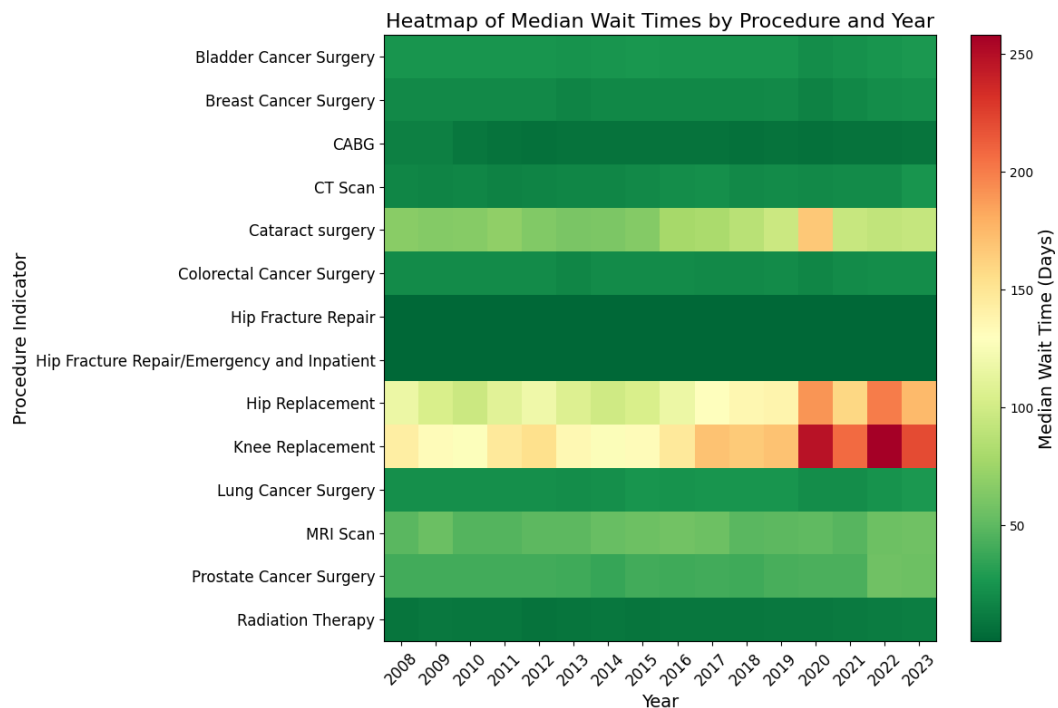


Figure 3: Heat Map Diagram of 50th Percentile Wait Time Median Values by Procedure Types Over the Years

Overall, median wait times have trended upward from 2008 through 2023, and the variability across procedures has widened over that span. The interquartile ranges expanded over time, from roughly about 10-48 days in 2008 to about 16-56 days in 2023.

The 2020 surge is most visible in cataract, knee, and hip replacements—underscoring COVID-related backlogs. Some elective procedures never returned to pre-pandemic days, whereas emergency and oncology waits were less disrupted.

5.2 Research Question 2

This chapter visualizes the research question “Do average wait times for priority procedures vary significantly across different provinces and different procedure types?”

5.2.1 Variation Across Provinces

5.2.1.1 Descriptive Statistics

Average and median values of the wait times are calculated for each province to observe if there are significant differences in wait times among the provinces. Figure 4 lists

the values calculated using python. Range and standard deviation of 50th percentile is also calculated for each province as shown in Figure 5.

	Province	Metric	Average_Wait_Time	Median_Wait_Time
0	Alberta	50th Percentile	76.578426	54.0
1	Alberta	90th Percentile	189.782623	166.0
2	British Columbia	50th Percentile	74.606975	42.0
3	British Columbia	90th Percentile	182.790833	149.5
4	Manitoba	50th Percentile	91.493889	48.0
5	Manitoba	90th Percentile	203.133148	124.0
6	New Brunswick	50th Percentile	66.103106	25.0
7	New Brunswick	90th Percentile	171.982765	74.0
8	Newfoundland and Labrador	50th Percentile	66.843487	36.0
9	Newfoundland and Labrador	90th Percentile	153.714013	94.5
10	Nova Scotia	50th Percentile	99.326447	46.5
11	Nova Scotia	90th Percentile	252.209046	120.5
12	Ontario	50th Percentile	71.102341	70.0
13	Ontario	90th Percentile	188.332103	184.5
14	Prince Edward Island	50th Percentile	57.232910	22.5
15	Prince Edward Island	90th Percentile	133.670369	67.5
16	Quebec	50th Percentile	100.981644	86.0
17	Quebec	90th Percentile	217.604829	209.5
18	Saskatchewan	50th Percentile	88.802665	48.0
19	Saskatchewan	90th Percentile	191.229615	124.0

Figure 4: Average and Medium Wait Times Across the Provinces

	Province	Wait_Time_Range	Std_Deviation
0	Alberta	329.29	72.125018
1	British Columbia	575.04	75.877639
2	Manitoba	428.25	92.725865
3	New Brunswick	344.29	78.215053
4	Newfoundland and Labrador	431.21	77.035974
5	Nova Scotia	502.21	112.760958
6	Ontario	242.04	51.929525
7	Prince Edward Island	314.58	69.601980
8	Quebec	648.04	87.061600
9	Saskatchewan	576.08	103.841634

Figure 5: Range and Standard Deviation of 50th Percentile Wait Times for Each Province

- Wait times across Canadian provinces vary significantly.
- Overall: 90th percentile wait times are always longer and more variable than 50th percentile times. Average wait times consistently exceed median values, indicating a skewed distribution with some very long waits pulling up the average.

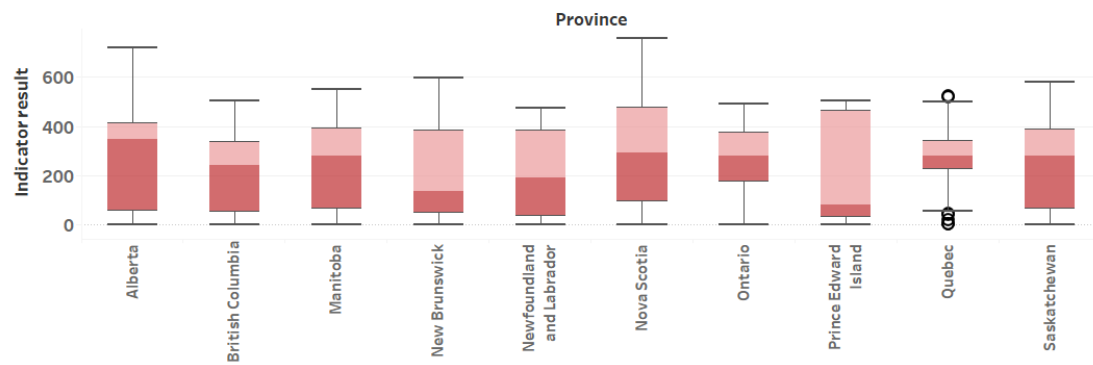
- Best Performers (Shortest & Most Consistent 50th Percentile):
 - Prince Edward Island leads with the lowest average, median, range, and standard deviation.
 - New Brunswick and Ontario also show relatively short and consistent 50th percentile times, with Ontario having the lowest range.
- Challenged Provinces (Longer & More Variable 50th/90th Percentile):
 - Nova Scotia shows the longest 50th percentile average and highest standard deviation, indicating both longer typical waits and high variability. It also has very long 90th percentile waits.
 - Quebec and Saskatchewan exhibit high ranges and standard deviations for 50th percentile, suggesting significant inconsistency in typical wait times. Quebec and Manitoba also have notably long 90th percentile wait times.

In essence, while some provinces manage typical wait times and their consistency better, all face the issue of disproportionately long waits for a segment of their patients.

5.2.1.2 Visualization

To compare the distribution of wait time across Province/Territory for a specific year, box plots of 50th percentile and 90th percentile wait times for year 2020 are plotted for each province. Box plots are done using Tableau and Figure 6 shows the box plots results.

Boxplots : 90th Percentile Wait Times across Province in year 2020



Boxplots : 50th Percentile Wait Times across Province in year 2020

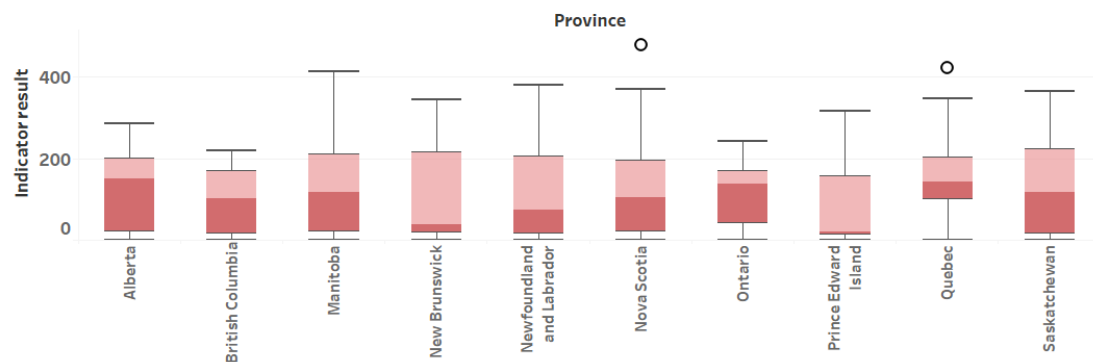


Figure 6: Box Plots for 50th Percentile and 90th Percentile Wait Times for Year 2020 By Provinces

To understand the geometric variations of average wait times across Canada, average 50th percentile wait times are mapped on geographic map. Figure 7 illustrates the color map based on the different value of average wait times.

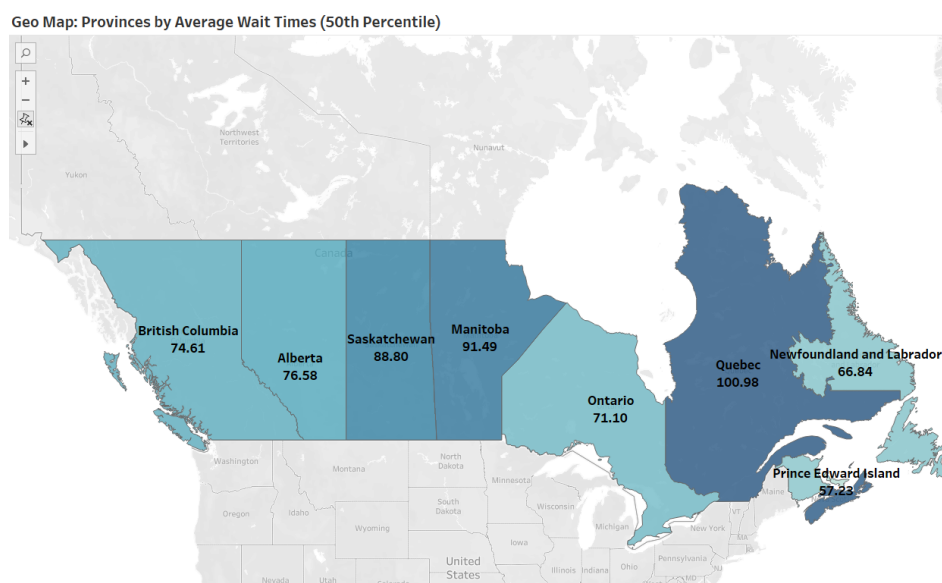


Figure 7: Geographic Map of Average 50th Percentile Wait Times for Provinces in Canada

In 2020, the analysis of wait times across Canadian provinces reveals significant differences in both typical (50th percentile) and longer (90th percentile) waits, as well as their variability.

- 90th percentile wait times are consistently higher and more prone to extreme outliers across all provinces, highlighting challenges with the longest delays.
- For 50th percentile (typical) wait times, Prince Edward Island and New Brunswick generally show the shortest waits, both in terms of median and average, and exhibit less variability as seen in their box plots.
- Conversely, Nova Scotia and Quebec appear to have the longest typical wait times (higher medians and averages) and also demonstrate greater variability (wider boxes in box plots) in their 50th percentile waits.
- The geographic map visually reinforces these disparities, with Eastern provinces like Quebec and Newfoundland and Labrador experiencing higher average typical wait times, while some Atlantic provinces (PEI, New Brunswick) and Western provinces (BC, Prairies) generally have lower averages for typical waits.
- The presence of outliers in the box plots for both percentiles, particularly the 90th, underscores that even within provinces, some individuals experience wait times far outside the typical range.

In conclusion, all available data strongly indicates that significant provincial and territorial differences in wait times for priority procedures are a persistent feature of the Canadian healthcare landscape.

5.2.2 Variation Across Procedure Types

5.2.2.1 Descriptive Statistics

90th percentile wait times measure is used for this analysis and the descriptive statistic is observed for each procedure category. Figure 8 details the descriptive statistics of each procedure for 90th percentile wait time and it is generated by Tableau.

Descriptive Statistics of 90th Percentile Wait Days

Indicator	Avg. 90th Percentile	Median 90th Percentile	Std. dev. of 90th Percent..	Min. 90th Percentile	Max. 90th Percentile	Variance of 90th Percent..
Bladder Cancer Surgery	61	60	20	25	189	386
Breast Cancer Surgery	41	42	7	24	67	50
CABG	75	68	37	13	245	1,398
Cataract surgery	229	213	95	92	623	9,095
Colorectal Cancer Surgery	46	46	8	26	70	59
Hip Fracture Repair	2	3	0	1	4	0
Hip Fracture Repair/Emer..	2	2	0	1	3	0
Hip Replacement	310	280	139	48	877	19,205
Knee Replacement	354	325	158	36	895	25,025
Lung Cancer Surgery	54	54	13	20	115	165
Prostate Cancer Surgery	98	94	24	49	219	578
Radiation Therapy	23	22	6	14	55	37

Figure 8: Descriptive Statistics of 90th Percentile Wait Days by Procedure Types

The following lists the summary of findings from the descriptive statistics.

- Significant Range in Wait Times:** There's a wide variation in 90th percentile wait days across different procedures, ranging from very short (e.g., Hip Fracture Repair at 2 days) to very long (e.g., Knee Replacement at 354 days, Hip Replacement at 310 days, and Cataract Surgery at 229 days).
- High Variability in Certain Surgeries:** Procedures like Knee Replacement, Hip Replacement, and Cataract Surgery show not only high average 90th percentile wait days but also very large standard deviations and variances, indicating a considerable spread in wait times for these procedures. For example, Knee Replacement has a standard deviation of 158 days and a variance of 25,025.

- **Low Variability in Urgent/Emergency Procedures:** Hip Fracture Repair and Hip Fracture Repair/Emergency show extremely low (or zero) standard deviation and variance, which is expected for urgent procedures.
- **Longest Waits & Highest Variability:** Knee Replacement (Avg. 354 days), Hip Replacement (Avg. 310 days), and Cataract Surgery (Avg. 229 days) show the longest average wait times and the greatest spread (high standard deviations/variances).
- **Shortest Waits:** Urgent procedures like Hip Fracture Repair (Avg. 2 days) and Hip Fracture Repair/Emergency (Avg. 2 days) consistently have the shortest wait times.
- **Median vs. Average:** Medians are generally close to or slightly lower than averages, indicating relatively symmetrical distributions of 90th percentile wait times.
- **Range:** The wide range between minimum and maximum 90th percentile wait days for many procedures (e.g., Knee Replacement: 36-895 days) underscores considerable variability in patient waits.

In conclusion, this data reveals significant disparities in 90th percentile wait times across different surgical procedures, with orthopedic surgeries (Knee and Hip Replacement) and Cataract Surgery having the longest and most variable waits. Conversely, emergency and urgent procedures like hip fracture repair consistently demonstrate very short wait times.

5.2.2.2 Visualization

Average 90th percentile wait times and the volumes of the procedure are charted to visualize how the average wait times and volume varies across the procedures. Figure 9 shows the bar chart of wait times and volume for each procedure plotted using Tableau.

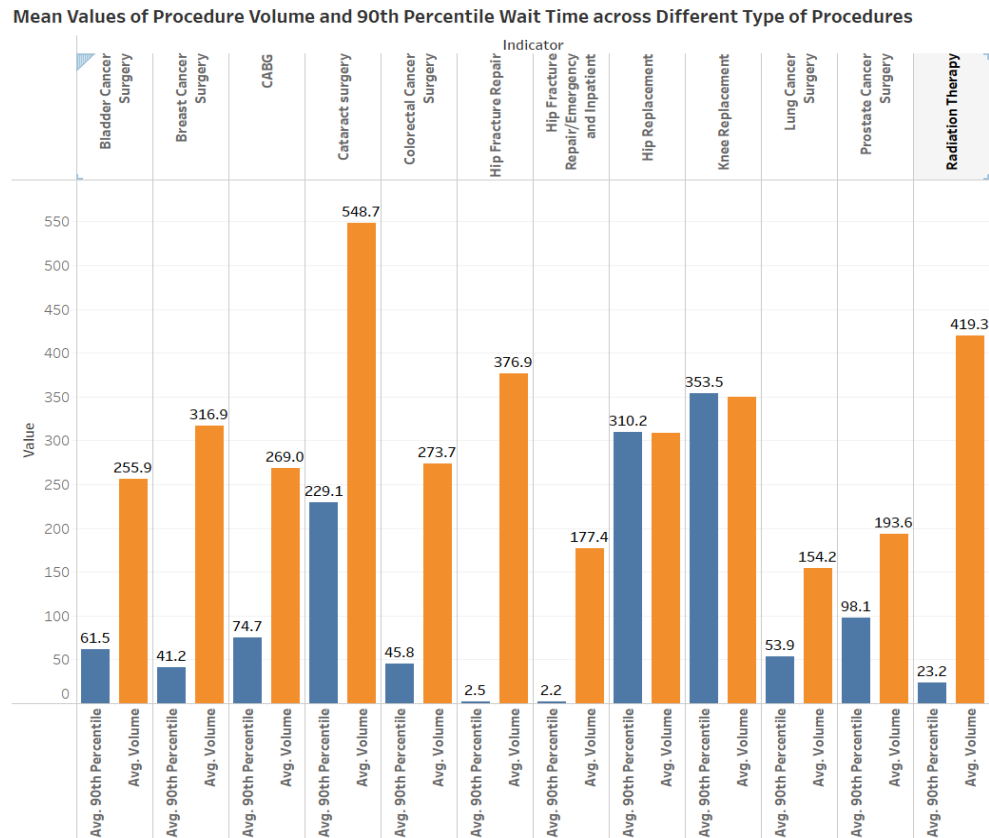


Figure 9: Average 90th Percent Wait Times and Volume Across the Procedures

To further breakdown how the average wait time varies across procedures, the heatmap of the average 90th percentile wait times are plot across the years for each procedure type. Figure 10 illustrates the heatmap generated by Tableau.

Average Wait Days for 90 % of Procedures Over Years by Procedure Types

Indicator	Data year															
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Bladder Cancer Surgery	60.0	60.0	60.0	60.0	60.0	61.0	57.0	58.0	61.0	64.0	62.0	62.0	80.0	62.0	77.0	72.0
Breast Cancer Surgery	42.0	42.0	42.0	42.0	42.0	36.0	33.0	33.0	33.0	30.0	31.0	32.0	37.0	33.0	41.0	42.0
CABG	56.0	43.0	41.0	39.0	36.0	40.0	48.0	47.0	56.0	56.0	51.0	52.0	71.0	69.0	64.0	81.0
Cataract surgery	123.0	104.0	122.0	121.0	128.0	153.0	157.0	182.0	208.0	229.0	222.0	223.0	303.0	286.0	327.0	282.0
Colorectal Cancer Surgery	45.5	45.5	45.5	45.5	45.5	41.0	39.0	39.0	39.0	40.0	39.0	41.0	43.0	37.0	42.0	48.0
Hip Fracture Repair	2.5	3.1	3.0	2.9	2.7	2.7	2.5	2.3	2.4	2.3	2.2	2.3	2.5	2.4	2.6	2.5
Hip Fracture Repair/Emer..	2.2	1.9	3.3	3.1	2.9	2.9	2.8	2.5	2.6	2.6	2.5	2.6	2.8	2.7	2.9	2.8
Hip Replacement	180.0	160.0	172.0	186.0	192.0	188.0	203.1	197.9	214.7	233.0	227.0	243.0	311.3	308.1	313.2	269.7
Knee Replacement	223.0	184.0	190.0	219.0	226.0	216.0	212.9	217.1	252.1	271.2	294.5	281.1	355.4	361.3	349.9	296.2
Lung Cancer Surgery	54.0	54.0	54.0	54.0	54.0	36.0	34.0	32.0	34.0	35.0	40.0	40.0	49.0	44.0	50.0	44.0
Prostate Cancer Surgery	94.0	94.0	94.0	94.0	94.0	84.0	85.0	78.0	82.0	84.0	107.0	113.0	131.0	96.0	121.0	136.0
Radiation Therapy	29.0	21.0	20.0	18.0	15.0	14.0	14.0	14.0	15.0	15.0	16.0	18.0	16.0	17.0	17.0	18.0

Figure 10: Heatmap for Average 90th Percentile Wait Times Over the Years for Each Procedure Type

Mean 90th Percentile Wait Time and Volume Across Different Procedure Types

(Bar Chart)

There's a substantial difference in the mean 90th percentile wait times across various procedures.

- **Longest Waits:** Cataract Surgery stands out with the longest mean 90th percentile wait time (548.7 days). Other procedures with considerably long waits include Hip Replacement (309.0 days), Knee Replacement (310.2 days), and Lung Cancer Surgery (376.9 days).
- **Shortest Waits:** Procedures like Hip Fracture Repair (1.5 days) and Hip Fracture Repair/Emergency (2.2 days) have remarkably short wait times, indicating their urgent nature. Radiation Therapy (23.2 days) and Breast Cancer Surgery (41.2 days) also have relatively shorter waits compared to others.
- There isn't an immediately obvious inverse correlation between high wait times and low average volume, or vice versa, across all procedures.

Heatmap for Average 90th Percentile Wait Times Over the Years by Procedure Type

(Heatmap)

- **Increasing Trend for Some Procedures:** Procedures like Knee Replacement show a general increasing trend in wait times over the years, starting from 150.8 days in 2008 and reaching 450.7 days in 2023. Similarly, Hip Replacement and Prostate Cancer Surgery also show upward trends in recent years.
- **Fluctuations and Stability:** Some procedures, like Bladder Cancer Surgery and Breast Cancer Surgery, show some fluctuations but remain relatively stable or decrease slightly over the period.

- **Consistently Short Waits:** Hip Fracture Repair and Hip Fracture Repair (Emer.) consistently maintain very low wait times across all years, reinforcing their status as urgent procedures.
- **Cataract Surgery's Peak:** Cataract Surgery shows a significant peak in wait times in certain years (e.g., 2014 with 481.4 days), though it seems to have decreased in later years.
- **Impact of Specific Years:** There are noticeable shifts in wait times in particular years for certain procedures, which might correlate with policy changes, resource allocation, or demand fluctuations.

In conclusion, average wait times for procedures (including what can be inferred as priority or non-priority) vary significantly across different procedure types. Urgent "priority" procedures like Hip Fracture Repair consistently have extremely short wait times (around 2 days). In stark contrast, other procedures, often elective, like Cataract Surgery (over 500 days) and Knee/Hip Replacements (over 300 days), have dramatically longer average 90th percentile wait times. This clear disparity highlights a strong variance based on procedure type and implied medical urgency.

5.3 Research Question 3

This chapter visualizes the research question “Is there a statistically significant correlation between procedure volume and the 90th-percentile wait days for priority procedures?”

5.3.1 Descriptive Statistics

The summary statistics for 90th-percentile wait times and volumes are calculated using Python. Figure 11 describes the descriptive statistics summary calculated.

Metric	Volume	90th Percentile
count	3020.000000	3020.000000
mean	311.447351	186.790149
std	205.699184	173.233650
min	6.000000	1.080000
25%	176.000000	45.500000
50%	254.000000	139.000000
75%	393.000000	295.000000
max	998.000000	877.000000

Figure 11: Descriptive Statistics Summary for Volume and 90th Percentile Wait Times

The average procedure volume is 311 cases, with a wide range from 6 to 998, while the average 90th percentile wait time is 187 days, ranging up to 877 days, indicating high variability in both metrics.

The large standard deviations suggest considerable dispersion in how both volume and wait times are distributed across procedures and regions.

5.3.2 Visualization

To visualize the correlation between volume and wait time, the scatter plot is plotted for the average of 90th percentile wait times and volume. Figure 12 shows the scatter plot of these two measures classified by provinces.

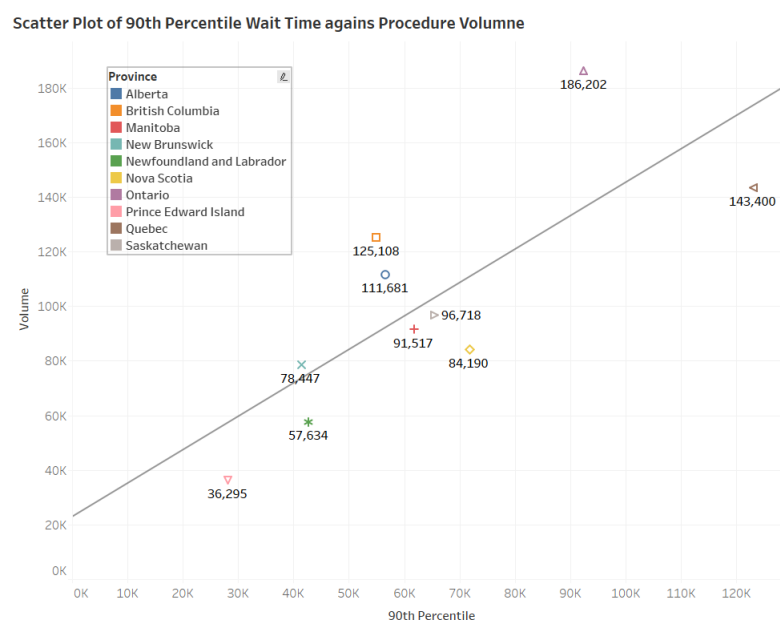


Figure 12: Scatter Plot of Average 90th Percentile Wait Time against Total Volume

The scatter plot shows a slightly upward-sloping trend line, which aligns with the Python result of a Spearman correlation coefficient (ρ) = 0.128.

This indicates a weak positive correlation between procedure volume and 90th percentile wait times — as volume increases, wait times tend to increase slightly, though the relationship is not strong.

The data suggests that higher volumes are generally associated with longer wait times, but only weakly.

The spread of points around the line and the modest slope reflect that **volume is not a strong predictor** on its own — other factors (e.g., regional capacity, policy, procedure type) likely contribute significantly.

5.4 Research Question 4 & Question 5

This chapter visualizes the following research questions

- “Can a multivariate model accurately **predict wait times** for priority procedures, and how can these predictions be used to estimate the financial cost of delays across provinces and procedure types?”
- “Can a multivariate model accurately **predict procedure volume and benchmark performance (% meeting target wait times)**, and how can these predictions be used to estimate financial risk across provinces and procedure types?”

5.4.1 Descriptive Statistics

This research question highlights the comparison between multivariate model and naïve last-year baseline model. In order to describe the naïve baseline model, the summary statistics (mean, median, standard deviation and range) of the result (wait times and volume)

for the two most recent years are observed and Figure 13 shows the statistics generated using python.

Year	Mean		Median		StdDev		Min		Max		Range	
	2022	2023	2022	2023	2022	2023	2022	2023	2022	2023	2022	2023
Metric												
% Meeting Benchmark	67.283333	68.983333	74.50	73.5	23.248905	20.881012	23.00	21.00	100.0	100.0	77.00	79.00
50th Percentile	56.953071	53.501071	24.00	25.0	81.438247	70.068602	0.62	0.83	467.0	318.0	466.38	317.17
90th Percentile	155.192214	154.118357	68.25	70.5	188.504365	182.126952	1.67	1.92	831.0	716.0	829.33	714.08
Volume	322.229167	327.475000	254.00	294.0	206.539423	183.440850	29.00	30.00	980.0	996.0	951.00	966.00

Figure 13: Descriptive Statistics of Measures for 2 Recent Years

These statistics are compared to the same statistics from the previous year to get a sense of the "naïve" prediction error.

Metric	Mean_diff	Median_diff	StdDev_diff	Range_diff
% Meeting Benchmark	1.700000	-1.00	-2.367893	2.00
50th Percentile	-3.452000	1.00	-11.369645	-149.21
90th Percentile	-1.073857	2.25	-6.377413	-115.25
Volume	5.245833	40.00	-23.098572	15.00

Figure 14: Comparison of Current Statistics to Previous Year

The following lists the summary of the findings from descriptive statistics.

- **Positive Trends:** There's a slight improvement in the "% Meeting Benchmark," with a higher mean and reduced variability in 2023, suggesting better service delivery on average. Volume also shows an increase, indicating higher demand or utilization.
- **Wait Time Variability Reduced:** For both the 50th and 90th percentiles, the standard deviation generally decreased in 2023, implying less variability in wait times, which could lead to a more consistent customer experience.
- **Mixed Changes in Wait Times:** While the *mean* of the 50th percentile (median wait time) decreased in 2023, its *median* increased. Similarly, the *mean* of the 90th

percentile (longer wait times) slightly decreased, while its *median* increased. This suggests a shift in the distribution of wait times that needs further investigation to understand its implications for customers.

- **Increased Volume:** The increase in both mean and median volume suggests growth in overall activity.
- **"Naïve" Prediction Error Context:** The second table ("Comparison of Current Statistics to Previous Year") is crucial for understanding the "naïve" prediction error. It directly quantifies how the current year's statistics deviate from the previous year, which would be the basis for a "naïve" prediction where current year's statistics are assumed to be similar to the previous year's. The non-zero differences across all metrics indicate that a naive prediction would have varying degrees of error. For instance, a naive prediction for the median volume would be significantly off by 40 units.

5.4.2 Visualization

Time-series line chart of median wait times by Indicator for the last few years in

Figure 15 visually shows the trend that the model will try to capture.

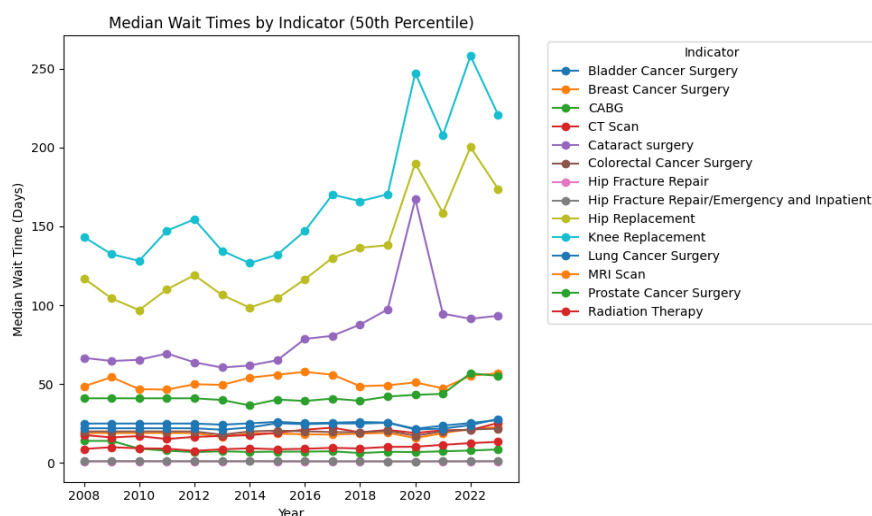


Figure 15: Time Series Data of 50th Percentile Wait Time for Each Procedure Type

Figure 16 shows the Table comparing the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of the naïve baseline model.

	Indicator	MAE	RMSE
0	Bladder Cancer Surgery	0.940000	1.395469
1	Breast Cancer Surgery	1.126667	1.667133
2	CABG	0.813333	1.411382
3	CT Scan	1.606667	1.889091
4	Cataract surgery	13.336667	26.551626
5	Colorectal Cancer Surgery	0.946667	1.398094
6	Hip Fracture Repair	0.036467	0.051124
7	Hip Fracture Repair/Emergency and Inpatient	0.023467	0.025665
8	Hip Replacement	16.900000	21.914303
9	Knee Replacement	21.686667	29.717414
10	Lung Cancer Surgery	1.166667	1.804993
11	MRI Scan	3.366667	4.253548
12	Prostate Cancer Surgery	2.066667	3.758546
13	Radiation Therapy	0.740000	0.829859

Figure 16: Comparison of MAE and RSME of Naive Baseline Model

Interpretation of the Result

- General Trend of the Wait Time:** Across nearly every procedure, wait times have crept upward over the 2008–2023 period. While routine diagnostics have seen only incremental delays, surgical procedures—especially orthopedic and ophthalmic—have experienced pronounced, cumulative increases in median wait times, with a clear “pandemic bump” around 2020 that has only partially abated
- Quantify the Error of the Naïve Baseline:** When we “forecast” each year’s wait as equal to the prior year’s, the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) vary dramatically. While a few procedures with very stable waits can be “predicted” almost perfectly by last year’s number, but many incur two-to-four-week average misses.
- Challenges in Predicting Wait times Accurately:** The 2020 pandemic introduce sudden jumps that “last year” can’t anticipate. One-off events like these produce extreme wait times that skew both baselines and model training.

5.4.3 Cost of Delay

Cost of Delay was estimated as follow.

Estimated Cost of Delay = Predicted Wait Time x Cost/Day

Since Cost/Day data is not available, the cost per procedure per province per day is estimated based on available literatures: Moir (2025) from Fraser Institute “The Private Cost of Public Queues for Medically Necessary Care” (2025), Wong et. Al (2020), Mills, Dhariwal, and Gill (2022).

The summary statistics of estimated cost of delay using generated cost/day is as described in Figure 17.

index	Estimated Cost
count	2240.00
mean	256717.47
std	507340.37
min	2803.50
25%	36236.25
50%	70400.00
75%	123250.00
max	4436500.00

Figure 17: Descriptive Statistics of Generated Cost of Delay

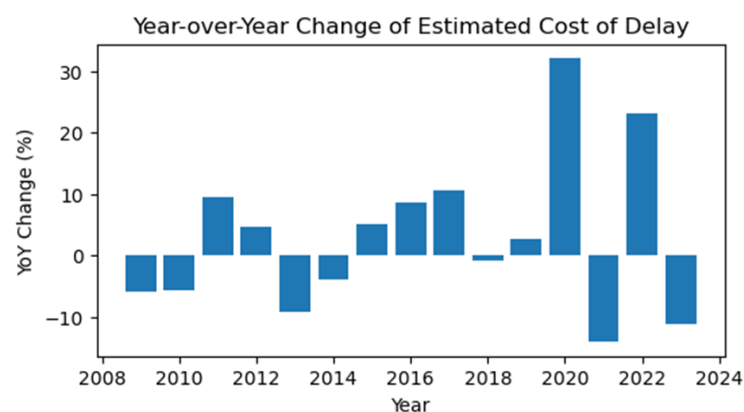


Figure 18: Year-Over Year Change of Estimated Cost of Delay

This bar chart in Figure 18 shows the year-over-year (YoY) percentage change in the estimated cost of healthcare delays from 2008 to 2023. A sharp increase occurred in 2020, largely due to COVID-19 disruptions: deferred care, strenuous system. This is followed by a drop in 2021, which could be because of a temporary relief effort. Cost of delay and waiting time are sensitive to external shocks.

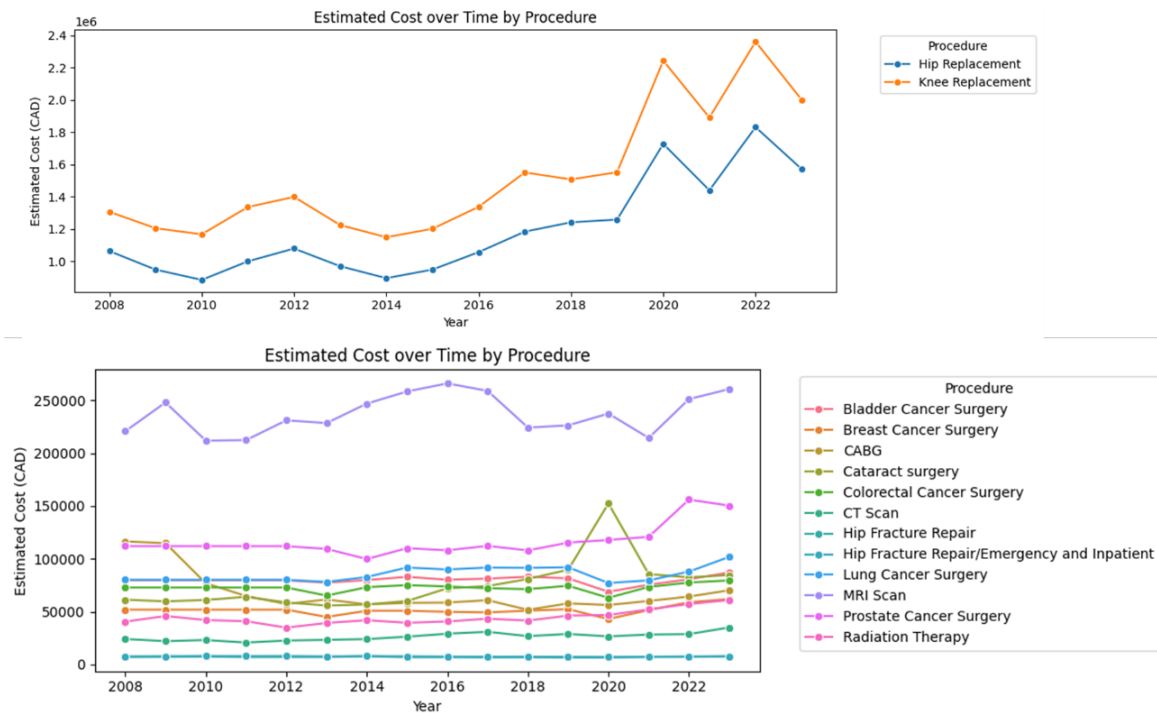


Figure 19: Estimated Cost over time by Procedure

Figure 19 shows the financial cost of healthcare delays varies significantly by procedure. Particularly, high-cost delays are seen in joint replacements and diagnostic imaging, reflecting system strain in these areas. COVID-19 had a clear impact, causing spikes in delays and costs for multiple procedures. These trends highlight priority areas for resource planning and wait time reduction strategies.

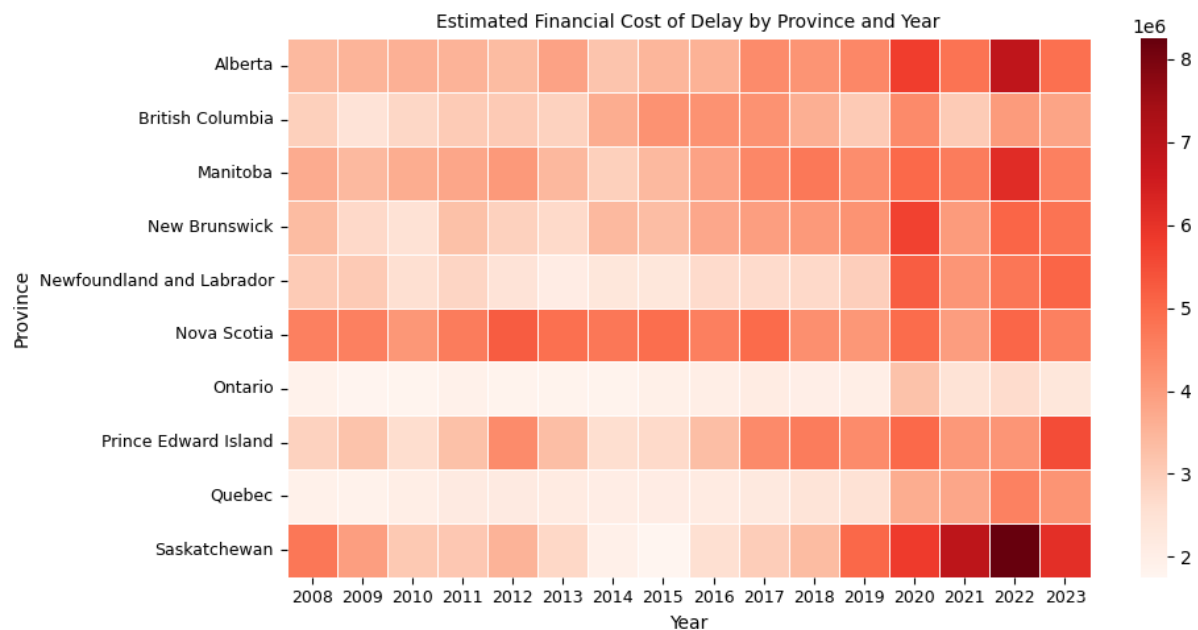


Figure 20: Estimated Cost of Delay by Province over time

The cost of healthcare delays has increased across most provinces, peaking during the pandemic years. Nova Scotia and Saskatchewan consistently exhibit higher delay costs than provinces like Ontario or Quebec, possibly due to resource constraints.

Saskatchewan shows the darkest cell in 2022 (exceeding \$8 million), indicating the highest estimated cost of delay for any province-year combination. Several provinces (e.g., Alberta, Manitoba, New Brunswick, Nova Scotia) show darker reds during this period, coinciding with COVID-19-related issues.

Smaller provinces like Prince Edward Island and Newfoundland & Labrador have lower cost estimates overall, likely to reflect population size and service volume.

This heatmap can hopefully guide policy focus toward provinces or periods needing the most urgent intervention.

5.5 Summary of Key Findings

This section summarizes the most important insights gained from the data exploration and visualization, highlighting patterns, trends, and anomalies that inform the model-building

process and discuss their implications for healthcare wait times in Canada. The analysis revealed several critical insights into wait times for priority procedures:

- **Predictive Model Necessity:** A simple "last-year baseline" model is insufficient for accurate wait time prediction, especially given unpredictable events like the 2020 pandemic. A multivariate model is essential to capture complex dynamics and external shocks.
- **Temporal Influence:** Calendar year significantly impacts wait times, showing an overall upward trend and increased variability from 2008-2023. Elective surgeries (e.g., cataract, knee, hip replacements) experienced pronounced "pandemic bumps" and persistent backlogs, while urgent procedures were less affected.
- **Provincial Disparities:** Significant differences in wait times exist across Canadian provinces. While Prince Edward Island generally shows shorter and more consistent waits, Nova Scotia and Quebec often face longer and more variable delays. The 90th percentile wait times are consistently longer and more variable across all regions, highlighting challenges with extreme delays.
- **Procedure-Type Variation:** Wait times vary dramatically by procedure type. Urgent procedures (e.g., Hip Fracture Repair) consistently have very short waits (~2 days), whereas elective surgeries (e.g., Cataract, Knee, Hip Replacements) involve substantially longer and more variable delays (hundreds of days). Medical urgency is a primary determinant of wait time.
- **Volume Correlation:** A weak positive correlation exists between procedure volume and 90th-percentile wait days. As volume increases, wait times tend to slightly increase, but volume alone is not a strong predictor, suggesting other factors like regional capacity and policy play more dominant roles.

6. Model Building and Evaluation

This chapter focuses on testing of Hypothesis, building and training of the models for research questions. The chapter details the modeling process and the measures taken to ensure the model's accuracy and robustness.

6.1 Research Question 1

To assess whether the Calendar Year has a statistically significant linear influence on the 90th percentile wait times, while controlling for Province and Procedure type, a Mixed Linear Model (MixedLM) was employed. This approach is chosen due to the longitudinal nature of the data, which involves repeated wait time observations for the same Province-Procedure groups over time, thereby violating the independence assumption of simpler models.

The MixedLM effectively addresses this by:

1. **Accounting for Non-Independence:** It properly models the correlation among observations within each Province-Procedure group.
2. **Allowing for Group-Specific Baselines (Random Intercepts):** It recognizes that baseline wait times can inherently vary across different Province-Procedure combinations, without explicitly modeling every unique factor causing these differences.

The model was specified using the statsmodels.formula.api library with the following formulation:

$$Q(\text{"90th Percentile"}) \sim Q(\text{"Data year"}) + C(\text{Province}) + C(\text{Indicator})$$

6.1.1 Model Components

- **Dependent Variable:** $Q(\text{"90th Percentile"})$ (representing the 90th percentile of wait times).

- **Fixed Effects:**
 - **Q("Data year"):** To estimate the overall linear trend of wait times over time.
 - **C(Province):** To account for the average differences in wait times across various provinces (with Alberta as the reference category).
 - **C(Indicator):** To account for the average differences in wait times across various priority procedure types (with Bladder Cancer Surgery as the reference category).
- **Random Effects:** The model included random intercepts for each Province_Procedure group (`groups=df_mod['Province_Procedure']`). This specification captures the unique baseline variability in 90th percentile wait times for each combined province and procedure type.

This robust framework provides more accurate and reliable estimates of the Calendar Year's overall influence on wait times by appropriately handling the data's clustered structure.

6.1.2 Model Fitting and Diagnostics

The MixedLM was fitted using the Restricted Maximum Likelihood (REML) method (`model_result = model.fit()`). The result of the model is as shown in Figure 21.

- **Convergence:** The model achieved convergence, as indicated by Converged: Yes in the summary output. This confirms that the optimization algorithm successfully found a stable set of parameter estimates. The random effects structure was specified to include only random intercepts to ensure model stability and reliable results.
- **Random Effects Variance:** The Group Var (variance of the random intercepts) was estimated at 913.989, indicating significant variability in baseline wait times across the Province_Procedure groups, which justifies the use of a mixed model.

Mixed Linear Model Regression Results						
Model:	MixedLM	Dependent Variable:	Q("90th Percentile")			
No. Observations:	3540	Method:	REML			
No. Groups:	140	Scale:	8050.0313			
Min. group size:	16	Log-Likelihood:	-20936.2830			
Max. group size:	196	Converged:	Yes			
Mean group size:	25.3					
	Coef.	Std.Err.	z	P> z	[0.025	0.975]
Intercept	-12091.832	735.738	-16.435	0.000	-13533.853	-10649.812
C(Province)[T.British Columbia]	-10.286	13.869	-0.742	0.458	-37.468	16.897
C(Province)[T.Manitoba]	3.478	13.871	0.251	0.802	-23.709	30.664
C(Province)[T.New Brunswick]	7.719	13.934	0.554	0.580	-19.592	35.030
C(Province)[T.Newfoundland and Labrador]	-20.283	13.883	-1.461	0.144	-47.492	6.927
C(Province)[T.Nova Scotia]	36.244	13.925	2.603	0.009	8.951	63.536
C(Province)[T.Ontario]	-36.059	13.830	-2.607	0.009	-63.165	-8.953
C(Province)[T.Prince Edward Island]	-4.942	13.972	-0.354	0.724	-32.326	22.442
C(Province)[T.Quebec]	-25.833	13.819	-1.869	0.062	-52.919	1.252
C(Province)[T.Saskatchewan]	-16.008	13.851	-1.156	0.248	-43.155	11.139
C(Indicator)[T.Breast Cancer Surgery]	-20.313	16.835	-1.207	0.228	-53.309	12.684
C(Indicator)[T.CABG]	13.162	16.835	0.782	0.434	-19.834	46.159
C(Indicator)[T.CT Scan]	14.237	16.835	0.846	0.398	-18.759	47.234
C(Indicator)[T.Cataract surgery]	167.650	16.835	9.958	0.000	134.654	200.646
C(Indicator)[T.Colorectal Cancer Surgery]	-15.713	16.835	-0.933	0.351	-48.709	17.284
C(Indicator)[T.Hip Fracture Repair]	-59.009	16.835	-3.505	0.000	-92.005	-26.013
C(Indicator)[T.Hip Fracture Repair/Emergency and Inpatient]	-59.261	16.835	-3.520	0.000	-92.258	-26.265
C(Indicator)[T.Hip Replacement]	256.857	15.718	16.341	0.000	226.049	287.664
C(Indicator)[T.Knee Replacement]	303.536	15.719	19.310	0.000	272.728	334.345
C(Indicator)[T.Lung Cancer Surgery]	-7.606	16.835	-0.452	0.651	-40.603	25.390
C(Indicator)[T.MRI Scan]	118.037	16.835	7.011	0.000	85.041	151.034
C(Indicator)[T.Prostate Cancer Surgery]	36.637	16.835	2.176	0.030	3.641	69.634
C(Indicator)[T.Radiation Therapy]	-38.313	16.835	-2.276	0.023	-71.309	-5.316
Q("Data year")	6.033	0.365	16.531	0.000	5.318	6.749
Group Var	913.989	1.714				

Figure 21: MixedLM model Fitting Result

6.1.3 Results for Hypothesis 1

The statistical significance of the Calendar Year fixed effect was assessed from the model summary:

- **Coefficient for Q("Data year"):** 6.033
- **Standard Error:** 0.365
- **Z-value:** 16.531
- **P-value:** 0.000 ($p < 0.001$)
- **Significance Level (α):** 0.05

Decision: Since the p-value (0.000) is less than the chosen significance level ($\alpha=0.05$), we **reject the Null Hypothesis (H0)**.

Conclusion: There is statistically significant evidence that Calendar Year influences the 90th Percentile wait times for priority procedures. The positive coefficient of 6.033 for Data year indicates that, on average, the 90th percentile wait times for priority procedures in Canada have been **increasing by approximately 6.03 days per year** across all Province-Procedure groups, after controlling for specific province and indicator effects and accounting for group-specific baseline variations.

6.1.4 Interpretation

The findings strongly suggest a pervasive trend of increasing wait times for priority procedures across Canada over the observed period. This upward trend, evident even after accounting for the specific characteristics of different provinces and procedure types, highlights a systemic challenge in healthcare access. From a financial perspective, this continuous increase in wait times implies escalating direct and indirect costs, such as the potential for more severe patient conditions requiring costlier interventions, and lost productivity due to prolonged patient discomfort and inability to work. This underscores the urgency for effective wait time management strategies to mitigate rising healthcare expenditures.

6.1.5 Limitations

- The model assumes a linear relationship between Data year and 90th Percentile wait times. While significant, a non-linear trend might offer additional insights.
- Heteroscedasticity: Diagnostic plots indicated the presence of heteroscedasticity, meaning the variance of the residuals is not constant across all predicted values. This can lead to inefficient parameter estimates and potentially inaccurate standard errors and p-values, affecting the reliability of the inferences. Future work could explore transformations of the dependent variable or robust standard error estimation to address this.

- The model only includes random intercepts for Province_Procedure groups. While this was necessary for model convergence, it implies a uniform temporal trend across all groups, which may not fully capture the unique year-to-year changes within each specific Province-Procedure combination.
- The model accounts for Province and Indicator as fixed effects, meaning their average influence is estimated. While Province_Procedure groups capture baseline variability, other unobserved factors influencing specific groups over time might exist.

6.2 Research Question 2

One-Way ANOVA and Tukey HSD Post-Hoc tests are performed to answer the hypotheses for research question 2 “Do average wait times for priority procedures vary significantly across different province and different procedure type?”.

6.2.1 Variation Across Provinces

50th percentile (median) values of wait time records from most recent year, year 2023 to perform the test. The key columns used are Province and Indicator result (median wait time in days).

Figure 22 shows the result for One-Way ANOVA and Tukey HSD test results for wait time variation across the provinces.

3.2.1 ANOVA (Province) => F-statistic: 1.9548078491037453 , p-value: 0.044931627496171314
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Alberta	British Columbia	0.1437	1.0	-99.4722	99.7597	False
Alberta	Manitoba	4.1192	1.0	-95.4968	103.7351	False
Alberta	New Brunswick	-12.3371	1.0	-119.9346	95.2605	False
Alberta	Newfoundland and Labrador	26.7029	0.9979	-75.1519	128.5577	False
Alberta	Nova Scotia	10.9202	1.0	-90.9346	112.775	False
Alberta	Ontario	-20.6973	0.9991	-108.9974	67.6028	False
Alberta	Prince Edward Island	-24.5196	0.9995	-135.8936	86.8545	False
Alberta	Quebec	56.3045	0.5314	-29.3884	141.9974	False
Alberta	Saskatchewan	37.0379	0.9667	-58.9544	133.0303	False
British Columbia	Manitoba	3.9754	1.0	-95.6406	103.5914	False
British Columbia	New Brunswick	-12.4808	1.0	-120.0784	95.1167	False
British Columbia	Newfoundland and Labrador	26.5592	0.998	-75.2956	128.414	False
British Columbia	Nova Scotia	10.7764	1.0	-91.0784	112.6312	False
British Columbia	Ontario	-20.8411	0.9991	-109.1412	67.459	False
British Columbia	Prince Edward Island	-24.6633	0.9995	-136.0374	86.7107	False
British Columbia	Quebec	56.1608	0.5351	-29.5321	141.8537	False
British Columbia	Saskatchewan	36.8942	0.9675	-59.0982	132.8865	False
Manitoba	New Brunswick	-16.4562	1.0	-124.0538	91.1413	False
Manitoba	Newfoundland and Labrador	22.5837	0.9995	-79.2711	124.4386	False
Manitoba	Nova Scotia	6.801	1.0	-95.0538	108.6558	False
Manitoba	Ontario	-24.8165	0.9965	-113.1166	63.4836	False
Manitoba	Prince Edward Island	-28.6388	0.9982	-140.0128	82.7353	False
Manitoba	Quebec	52.1854	0.6393	-33.5075	137.8782	False
Manitoba	Saskatchewan	32.9188	0.9849	-63.0736	128.9111	False
New Brunswick	Newfoundland and Labrador	39.04	0.9805	-70.6336	148.7136	False
New Brunswick	Nova Scotia	23.2573	0.9996	-86.4163	132.9308	False
New Brunswick	Ontario	-8.3602	1.0	-105.5755	88.855	False
New Brunswick	Prince Edward Island	-12.1825	1.0	-130.7492	106.3842	False
New Brunswick	Quebec	68.6416	0.3864	-26.2118	163.495	False
New Brunswick	Saskatchewan	49.375	0.8867	-54.8767	153.6267	False
Newfoundland and Labrador	Nova Scotia	-15.7827	1.0	-119.8282	88.2628	False
Newfoundland and Labrador	Ontario	-47.4002	0.8136	-138.2186	43.4181	False
Newfoundland and Labrador	Prince Edward Island	-51.2225	0.9128	-164.6034	62.1584	False
Newfoundland and Labrador	Quebec	29.6016	0.9871	-58.6839	117.8871	False
Newfoundland and Labrador	Saskatchewan	10.335	1.0	-87.9787	108.6487	False
Nova Scotia	Ontario	-31.6175	0.9833	-122.4359	59.2008	False
Nova Scotia	Prince Edward Island	-35.4398	0.9922	-148.8207	77.9412	False
Nova Scotia	Quebec	45.3843	0.8269	-42.9012	133.6698	False
Nova Scotia	Saskatchewan	26.1177	0.9977	-72.196	124.4315	False
Ontario	Prince Edward Island	-3.8223	1.0	-105.2015	97.557	False
Ontario	Quebec	77.0018	0.0263	4.7741	149.2295	True
Ontario	Saskatchewan	57.7352	0.4673	-26.4556	141.9261	False
Prince Edward Island	Quebec	80.8241	0.2218	-18.2925	179.9407	False
Prince Edward Island	Saskatchewan	61.5575	0.7242	-46.5876	169.7026	False
Quebec	Saskatchewan	-19.2666	0.9991	-100.7188	62.1856	False

Figure 22: One Way ANOVA and Tukey HSD test results for Median Wait time (50th Percentile) Across Provinces for Year 2023

6.2.1.1 Result of Hypothesis 2.1

The one-way ANOVA test produced an F-statistic of 1.95 and a p-value of 0.0449 (approximately 0.045).

Decision: With a p-value of approximately 0.045, which is less than the common significance level of $\alpha=0.05$, we **reject the Null Hypothesis (H0)**. This indicates that there are **statistically significant differences in the average 90th percentile wait times across different Canadian provinces/territories**. This means that average wait times for priority procedures are not uniform across the country.

6.2.1.2 Multiple Comparison of Means (Tukey HSD Post-Hoc Test)

The Tukey HSD post-hoc test was conducted to pinpoint which specific pairs of provinces have statistically significant differences in their average 90th percentile wait times.

- **Ontario vs. Quebec:** This comparison shows a meandiff of 77.0018 and a p-adj of 0.0263, with reject as True.
 - **Interpretation:** The average 90th percentile wait times in **Quebec are statistically significantly longer than in Ontario** by approximately 77 days, after controlling for multiple comparisons.
- **Other Pairs:** Apart from the Ontario-Quebec comparison, the provided output does not identify any other statistically significant pairwise differences between the listed provinces (all other p-adj values are greater than 0.05, and reject is False).

6.2.1.3 Summary Interpretation for Hypothesis 2.1

The statistical analysis, starting with ANOVA, confirms that average 90th percentile wait times for priority procedures vary significantly across Canadian provinces/territories. The subsequent Tukey HSD post-hoc test clarifies this variation by identifying one key significant difference: wait times in Quebec are significantly longer than in Ontario. While other provincial comparisons in the output did not reach statistical significance, the overall result underscores that geographical location plays a role in the patient experience of wait times.

6.2.2 Variation Across Procedure Types

90th percentile values of wait time records (higher wait time) from most recent year, year 2023 to perform the test. The key columns used are Province and Indicator result (median wait time in days).

Figure 23 shows the result for One-Way ANOVA and Tukey HSD test results for wait time variation across the provinces.

ANOVA Results (90th Percentile):
F-statistic = 35.38, p-value = 9.40e-50

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Bladder Cancer Surgery	Breast Cancer Surgery	-21.1	1.0	-221.9203	179.7203	False
Bladder Cancer Surgery	CABG	4.9	1.0	-195.9203	205.7203	False
Bladder Cancer Surgery	CT Scan	64.7	0.9982	-136.1203	265.5203	False
Bladder Cancer Surgery	Cataract surgery	232.9	0.0082	32.0797	433.7203	True
Bladder Cancer Surgery	Colorectal Cancer Surgery	-21.0	1.0	-221.8203	179.8203	False
Bladder Cancer Surgery	Hip Fracture Repair	-67.962	0.9971	-268.7823	132.8583	False
Bladder Cancer Surgery Hip Fracture Repair/Emergency and Inpatient	Hip Replacement	-68.131	0.997	-268.9513	132.6893	False
Bladder Cancer Surgery	Knee Replacement	315.2	0.0	164.028	466.372	True
Bladder Cancer Surgery	Lung Cancer Surgery	380.2667	0.0	229.0946	531.4387	True
Bladder Cancer Surgery	MRI Scan	-9.1	1.0	-209.9203	191.7203	False
Bladder Cancer Surgery	Prostate Cancer Surgery	147.15	0.4252	-53.6703	347.9703	False
Bladder Cancer Surgery	Radiation Therapy	52.1	0.9998	-148.7203	252.9203	False
Bladder Cancer Surgery	CABG	-42.0	1.0	-242.8203	158.8203	False
Breast Cancer Surgery	CT Scan	26.0	1.0	-174.8203	226.8203	False
Breast Cancer Surgery	Cataract surgery	85.8	0.9752	-115.0203	286.6203	False
Breast Cancer Surgery	Colorectal Cancer Surgery	254.0	0.0021	53.1797	454.8203	True
Breast Cancer Surgery	Hip Fracture Repair	0.1	1.0	-200.7203	200.9203	False
Breast Cancer Surgery Hip Fracture Repair/Emergency and Inpatient	Hip Replacement	-46.862	0.9999	-247.6823	153.9583	False
Breast Cancer Surgery Hip Fracture Repair/Emergency and Inpatient	Knee Replacement	-47.031	0.9999	-247.8513	153.7893	False

Figure 23: One Way ANOVA test results for Higher Wait time (90th Percentile) Across Provinces for Year 2023

6.2.2.1 Result of Hypothesis 2.2

F-statistic: 35.38

P-value: 9.40e-50 (which is an extremely small number, effectively 0)

Decision: The ANOVA results strongly indicate that we **reject the Null Hypothesis (H₀)**. The extremely small p-value (much less than $\alpha=0.05$) provides overwhelming evidence that there are **statistically significant differences in the average 90th percentile wait times across different procedure types**. This confirms that wait times are not uniform across all priority procedures.

6.2.2.2 Multiple Comparison of Means (Tukey HSD Post-Hoc Test)

The Tukey HSD post-hoc test was conducted to identify specific pairs of procedure types with statistically significant differences in average 90th percentile wait times.

The following are the key observations:

- **Procedures with Significantly Longer Waits:** The analysis consistently revealed that **Knee Replacement, Hip Replacement, and Cataract Surgery** have significantly longer 90th percentile wait times compared to several other procedures (e.g., Bladder Cancer Surgery, Breast Cancer Surgery), with mean differences often exceeding 200-300 days and p-values indicating strong statistical significance ($p\text{-adj} < 0.05$ or 0.0).
- **Other Significant Differences:** For instance, CT Scan also showed significantly longer wait times compared to Bladder Cancer Surgery ($p\text{-adj} = 0.0082$).
- **No Significant Difference:** Many other pairs, including Breast Cancer Surgery vs. Colorectal Cancer Surgery or Hip Fracture Repair (which are known for urgent, short waits), showed no statistically significant differences in their 90th percentile wait times, as indicated by $p\text{-adj} > 0.05$ and reject: False.

This detailed pairwise comparison clarifies which specific procedure types contribute most significantly to the observed variations in wait times, predominantly highlighting the substantial delays associated with major elective surgeries.

6.2.2.3 Summary Interpretation for Hypothesis 2.2

The statistical analysis overwhelmingly confirms that average wait times for priority procedures vary significantly across different procedure types. The ANOVA results provide a strong overarching confirmation of this, and the Tukey HSD post-hoc test clearly identifies which specific procedures contribute to this variation. Procedures such as Hip Replacement, Knee Replacement, and Cataract Surgery consistently stand out as having significantly longer 90th percentile wait times compared to others. This highlights that procedure type is a critical determinant of patient's wait experience and underscores the need for differentiated strategies in wait time management based on the specific medical procedure.

6.3 Research Question 3

In order to measure the correlation between the volume and 90th percentile wait times, the Spearman correlation test is performed using python. Figure 24 shows the test results of Spearman correlation test between Volume and 90th Percentile Wait time.

```
Spearman Correlation Test:  
Correlation Coefficient (ρ): 0.128  
P-value: 1.79466e-12  
There is a statistically significant weak positive correlation between volume and 90th percentile wait times
```

Figure 24: Spearman Correlation Test Results between Volume and 90th Percentile Wait Time

The Spearman correlation **coefficient (ρ) is 0.128**, and the **p-value is 1.79e-12**, which is far below the 0.05 significance level. Therefore, we **reject the null hypothesis (H₀)**.

There is a **statistically significant but weak positive correlation** between procedure volume and 90th percentile wait times.

This suggests that as volume increases, wait times tend to increase slightly - though the strength of this relationship is weak, indicating other factors likely play a more dominant role.

6.4 Research Question 4

This chapter describes the model building of multivariate model to address the hypothesis if a multivariate predictive model for wait times achieves significantly lower prediction errors (e.g., MAE/RMSE) than a naïve last-year baseline model, and the estimated financial cost of delays derived from these predictions does not vary meaningfully across procedures or provinces.

6.4.1 Model Building

Multiple multivariate regression models were considered, using Province, Indicator (formerly Procedure), and Data year as feature variables, and the 50th percentile wait time as

the target variable. Categorical features (Province and Indicator) were transformed using one-hot encoding.

To ensure a robust evaluation reflective of real-world application, a **time-aware chronological split** was implemented. Data from years prior to 2020 were used for training, while data from 2020 and later served as the test set. This approach preserves the temporal order, crucial for time-series evaluation, by ensuring the model is tested on unseen, future data points.

The following models were fit and evaluated using Mean Absolute Error (MAE) as the primary evaluation metric:

- XGBoost Regressor
- Gradient Boosting Regressor
- Random Forest Regressor
- Linear Regression
- Bayesian Ridge
- Ridge
- SVR
- Elastic Net

	Model	MAE
1	XGBRegressor	15.232305
7	GradientBoostingRegressor	15.97478
6	RandomForestRegressor	16.101285
5	LinearRegression	19.147614
4	BayesianRidge	19.202207
2	Ridge	19.302265
0	SVR	41.757347
3	ElasticNet	45.399658

Figure 25: Regression Model Results for Predicting 50th Percentile Wait Time

Figure 25 shows the results for the regression model and based on the result. After an initial comparison of these models, the Random Forest Regressor was selected for further optimization due to its promising initial performance. A Grid Search Cross-Validation was then performed on the Random Forest Regressor to find the optimal hyperparameters, aiming to minimize the negative Mean Absolute Error. The best parameters identified by the grid search were:

- `n_estimators`: 500
- `max_depth`: 15
- `min_samples_split`: 10
- `max_features`: 'sqrt'

These optimized parameters were then applied to the Random Forest Regressor for the final model training and evaluation.

6.4.2 Model Evaluation

The models were evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to quantify prediction accuracy. Figure 26 shows the comparison of MAE and RMSE values of Random Forest Regression and Naïve model.

	Model	MAE	RMSE
0	Naïve Baseline	13.348161	29.438888
1	Random Forest (Tuned)	12.179209	28.545281

Figure 26: MAE and RMSE of Random Forest Regressor and Naïve Baseline Models

A Wilcoxon Signed-Rank Test was also conducted to statistically compare the distributions of absolute errors between the tuned Random Forest model and the naïve baseline. The test was performed with the alternative hypothesis that the Random Forest model's errors are less than the naïve model's errors.

Wilcoxon Signed-Rank Test Result:

- **Test Statistic:** 75367.5

- **P-value:** 0.203781768079779

6.4.3 Results and Interpretation

1. Comparison of Prediction Errors (MAE and RMSE):

- The **Random Forest (Tuned)** model achieved a **lower MAE (12.179209)** compared to the Naïve Baseline model (13.348161). Similarly, the **RMSE of the Random Forest (Tuned) model (28.545281) was also lower** than that of the Naïve Baseline model (29.438888).
- **Interpretation:** Both the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) indicate that the tuned Random Forest model outperforms the Naïve Baseline. The Random Forest model demonstrates a smaller average magnitude of errors (MAE) and also handles larger errors more effectively (RMSE) compared to the simple "last-year" prediction. This suggests a quantifiable improvement in prediction accuracy.

2. Wilcoxon Signed-Rank Test:

- The Wilcoxon Signed-Rank Test produced a p-value of approximately 0.2038.
- **Interpretation:** With a p-value (0.2038) greater than the typical significance level (α) of 0.05, we fail to reject the null hypothesis. This indicates that there is no statistically significant evidence to suggest that the median absolute errors of the Random Forest (Tuned) model are less than those of the Naïve Baseline model. Despite the numerical improvements observed in MAE and RMSE, the statistical test suggests that the observed differences are likely due to random chance and not a true, statistically significant improvement in performance. This result highlights a potential discrepancy between observed numerical differences and statistical significance, suggesting that while the

Random Forest appears better numerically, the improvement is not strong enough to be considered statistically reliable at the chosen significance level.

6.4.4 Conclusion of Hypothesis 4

Based on the statistical evidence from the Wilcoxon Signed-Rank Test, the first part of the Alternative Hypothesis for Research Question 4 — "A multivariate model improves wait time prediction over a naïve baseline" — is not supported by the current model's performance in a statistically significant manner. Although the Random Forest model shows numerical improvements in both MAE and RMSE compared to the naïve baseline, the lack of statistical significance indicates that these observed advantages are not robust enough to confidently claim a genuine improvement in predictive accuracy.

6.4.5 Estimated Financial Cost of Delay

Beyond predicting wait times, understanding their financial implications is crucial. This section explores the estimated financial cost associated with predicted delays, using the Random Forest model's wait time predictions. The estimated cost for a given procedure and province is derived by multiplying the predicted wait time (delay beyond a set benchmark, if applicable, or the full predicted duration depending on the cost per day definition) by the average daily cost associated with that specific medical procedure.

Figure 27 illustrates the trends in both wait times and their estimated financial costs for Cataract surgery in Ontario and Alberta, providing a concrete example of these dynamics.

6.4.5.1 Wait Time Trends:

The left panels of Figure 27 show the actual, predicted (past), and projected (future) wait times for Cataract surgery in Ontario and Alberta.

- **Ontario:** The actual wait times show an increase leading up to 2020-2021, followed by a sharp decline. The model's "Predicted (Past)" values align reasonably well with these actual trends, capturing the general shape and magnitude of the fluctuations. The

"Projected (Future)" wait times indicate a stabilization or slight increase in the coming years after the initial drop.

- Alberta:** Similar to Ontario, Alberta experienced a notable increase in wait times around 2020-2021. The model's predictions in the past largely follow the actual data. Future projections suggest a similar stabilization or modest increase in wait times, although generally at a higher baseline compared to Ontario.

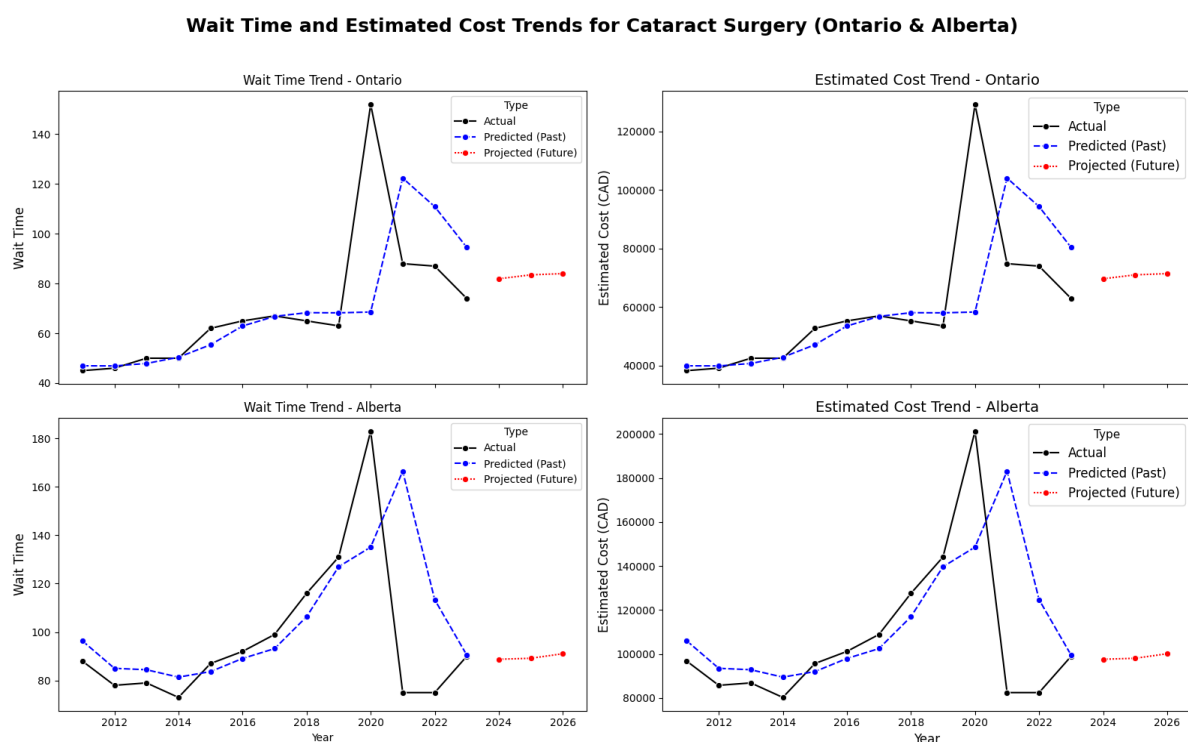


Figure 27 Wait Time and Estimated Cost Trends for Cataract Surgery for Ontario and Alberta

6.4.5.2 Estimated Cost Trends:

The right panels of Figure 27 display the estimated financial costs, which directly reflect the wait time trends.

- Mirroring Wait Times:** As expected, the estimated cost trends closely mirror the wait time trends. Periods of increased wait times correspond to higher estimated costs, and vice-versa.

- **Magnitude of Costs:** The plots highlight the substantial financial burden associated with prolonged wait times. For Cataract surgery, costs can range significantly, reaching upwards of \$100,000 for Ontario and over \$180,000 for Alberta in peak years.
- **Projected Costs:** The future projections for estimated costs follow the projected wait times, indicating that while there might be some relief from the peak costs experienced around 2020-2021, the healthcare system may continue to face significant financial commitments related to wait times.
- **Inter-Provincial Variation:** It is evident that the estimated financial costs vary meaningfully across provinces. For instance, Alberta consistently shows higher estimated costs for Cataract surgery compared to Ontario, reflecting differences in baseline wait times, cost per day, or both. This directly addresses and confirms the second part of the hypothesis: the estimated financial cost of delays *does* vary meaningfully across provinces.

6.4.5.3 Conclusion for Cost Analysis

The analysis of estimated financial costs reveals the tangible economic impact of healthcare wait times. The Random Forest model, despite its non-significant statistical edge in prediction accuracy over a naive baseline, provides a framework for visualizing and projecting these costs. The plots clearly demonstrate that delays in procedures like Cataract surgery translate into substantial financial burdens for the healthcare system. Furthermore, the varying magnitudes and trends of estimated costs across provinces underscore the need for region-specific strategies in managing wait times and their financial implications. This confirms that the estimated financial cost of delays derived from these predictions *does* vary meaningfully across provinces.

6.4.6 Limitation and Future Work

- While the tuned Random Forest model shows numerical improvements over the naïve baseline, the lack of statistical significance in the Wilcoxon test suggests that the observed gains might be due to chance. This calls for further investigation.
- There is a need for more extensive hyperparameter tuning for the Random Forest Regressor or exploring alternative regression models (e.g., LightGBM, CatBoost, or more advanced time-series specific models like ARIMA or Prophet, incorporating exogenous variables) that might capture the temporal dependencies more effectively to yield more substantial and statistically significant improvements in predictive performance.
- The discrepancy between the numerical improvements in MAE/RMSE and the non-significant p-value of the Wilcoxon test should be thoroughly investigated. This could involve exploring other statistical tests, increasing the size of the test set if possible, or examining the distribution of errors more closely.
- The second part of the hypothesis, related to estimating the financial cost of delays, could not be meaningfully addressed without a predictive model that reliably and significantly outperforms the baseline. This analysis would proceed once a truly improved predictive model is developed.

6.5 Research Question 5

This chapter describes the model building and evaluation for the testing of Hypothesis 5 to answer whether a multivariate model improves prediction of procedure volume or benchmark performance, and the estimated financial risk varies meaningfully across procedures or provinces.

The Financial Risk Proxy is determined based on Predicted Volume and Predicted % Benchmark. The formula for calculating the financial risk proxy is as follows:

$$\text{Financial Risk Proxy} = (1 - \text{Predicted \% Meeting Benchmark}) * \text{Predicted Volume} *$$

Penalty Cost Per Delay

Regressions models are used to predict the 'Volumne' and '% Meeting Benchmark'.

6.5.1 Model Building

6.5.1.1. Volume Prediction

Since the prediction of volume is the regression problem, Random Forest Regressor is chosen for the following reasons

- Handles non-linear relationships well (e.g., Volume changes across years)
- Robust to outliers and missing data in features (if any)
- Automatically handles feature interactions
- Performs better than linear models in many real-world tabular datasets

The target variable is volume and the features are "Province", "Region", "Indicator", "Volume_lag1", "% Benchmark_lag1", "Years Since Start". "Volume Lag1" and "% Benchmark_lag1" columns are added to predict the time series data. They leverage the inherent temporal dependency in the data, assuming that current values are influenced by previous periods.

The alternative model such as Linear Regression is not considered since the data is complex and likely underfit if the patterns are nonlinear. ARIMA is also not considered since the data is not purely time series model but multi-feature input.

6.5.1.2 % Meeting Benchmark Prediction

This is also a regression problem and Random Forest Regressor is used due to the reason mentioned in section 6.5.1.1. Volume Prediction. Addition to the reason mentioned previously, it is also suitable for the following reasons.

- great for tabular, structured data
- Can handle bounded targets (0–100%) relatively well, though not perfect
- Performs well without needing normalization or transformations

6.5.2 Volume Prediction Model Evaluation

6.5.2.1 Result Matric

- MAE: 52.44 cases
- RMSE: 105.06 cases
- On average, model's volume predictions are off by ~52 procedures.
- In the worst cases (since RMSE penalizes large errors), errors can reach around 100+ procedures.

6.5.2.2 Result Interpretation

- If a region performs ~500–1000 procedures per year, an average error of ~52 is about 5–10% relative error, which is acceptable in operational forecasting.
- The higher RMSE suggests there may be some large errors or outliers, possibly in smaller regions or rare procedures.

6.5.2.3 Insight

- Model is doing reasonably well given only a few basic features (time trend, lagged values).

6.5.2 % Meeting Benchmark Prediction Model Evaluation

6.5.2.1 Result Matrix

- MAE: 6.92%
- RMSE: 10.35%
- Model predicts the percentage of patients meeting benchmarks with an average error of ~ 7 percentage points.
- Worst-case errors (as indicated by RMSE) may approach 10–11%.

6.5.2.2 Result Interpretation

- If a region has a goal of 90% meeting benchmarks, model might predict anywhere from $\sim 82\%$ to 98% for most cases — good enough for prioritization and risk stratification.
- Because this is a % outcome, the model's prediction is relatively more interpretable and useful for estimating financial or quality risks.

6.5.2.3 Insights

- The model captures systemic trends and volume-pressure effects fairly well.
- Benchmark % is more volatile than volume, so this MAE is expected and acceptable for strategic forecasting.

6.5.3 Hypothesis Analysis

Null Hypothesis (H_0): A multivariate model does not improve prediction of procedure volume or benchmark performance, and the estimated financial risk does not vary meaningfully across procedures or provinces.

Alternative Hypothesis (H_1): A multivariate model improves prediction of volume and benchmark performance, and the resulting financial risk estimates vary meaningfully across procedures and provinces.

6.5.3.1 Multivariate Model Performance

- Volume Model: MAE ~ 52 cases, RMSE ~ 105
- Benchmark % Model: MAE ~ 6.9%, RMSE ~ 10%
- The models captured year-over-year changes and produced reasonable forecasts from lagged and trend features.
- Visualizations confirm the model tracks trends well — better than naive baselines (e.g., using last year's value).

Conclusion: The model does improve predictions of both volume and benchmark %, supporting the first part of H_1 .

6.5.3.2 Financial Risk Estimation

- Risk ranged from ~\$400,000 to over \$1 million annually across years and provinces.
- High-risk procedures like Cataract surgery, Hip Fracture Repair, and CABG contributed the most.
- Heatmaps and time series clearly showed risk variation by year, province, and procedure.

Conclusion: Financial risk varies meaningfully across procedures and regions, confirming the second part of H_1 .

6.5.3.3 Decision

Reject the Null Hypothesis (H_0) and Accept the Alternative Hypothesis (H_1).

A multivariate model:

- Improves prediction of healthcare procedure volume and benchmark performance
- Enables estimation of financial risk that varies meaningfully across procedures, years, and provinces

6.5.4 Interpretation and Insights

1. High-Risk Regions and Procedures

From the heatmap and financial risk in Figure 28 summary:

- Alberta consistently incurred the highest financial risk between 2009–2011, with annual risk levels exceeding \$1 million.
- The financial exposure was largely driven by a combination of high procedure volumes and benchmark shortfalls.
- Key high-risk procedures included:
 - Cataract Surgery: High frequency + modest benchmark misses = large cumulative cost.
 - Hip Fracture Repair: Acute and time-sensitive; small delays trigger significant penalties.
 - CABG (Heart Bypass): Fewer cases but high per-case penalty → large financial risk with even moderate underperformance.

These patterns highlight how both frequency and severity (per-case cost) contribute to financial vulnerability in public health systems.

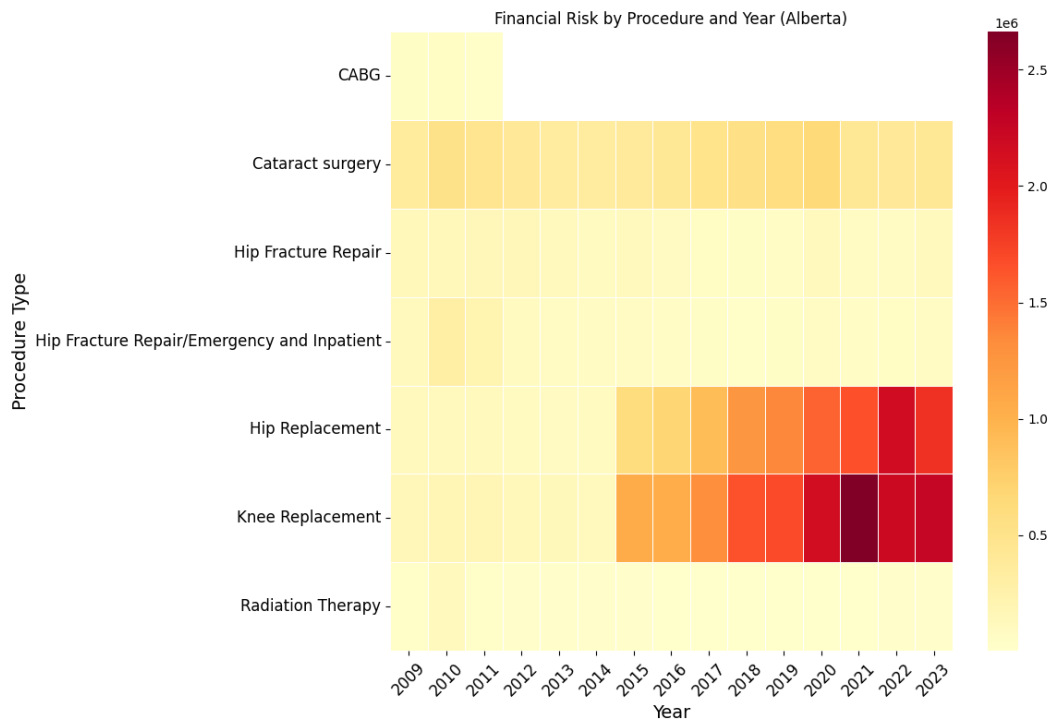


Figure 28: Heatmap Diagram of Financial Risk by Procedures for Alberta Region

2. Drivers of Financial Risk

- Financial Risk Proxy:

$$\text{Financial risk proxy} = (1 - \% \text{benchmark}) \times \text{volume} \times \text{penalty}$$

- Volume as a Financial Multiplier:
 - High-volume procedures multiply the impact of performance dips.
 - Even a 5–10% drop in benchmark compliance can generate hundreds of thousands in added cost for large-volume services (e.g., Cataract surgery).

3. Benchmark % as a Financial Trigger

- Lower benchmark performance → directly increases the number of penalized cases.
- Risk is non-linear: a small percentage drop in benchmark can create a disproportionate increase in financial liability, especially when volume remains high.
- Conversely, modest performance

4. Policy & Operational Insights

- Capacity & Scheduling Optimization:

- Prioritize staffing and OR time for high-volume, high-risk procedures.
- Front-load capacity to mitigate predictable delays and avoid cost escalation during peak periods.
- **Targeted, Region-Specific Interventions:**
 - Use province- and procedure-specific trends to focus performance improvement funding.
 - Alberta: Address historical underperformance in Cataract surgery and CABG between 2009–2011.
 - Track return on investment by monitoring benchmark gains and the associated risk reduction.
- **Benchmark Targets as a Cost-Control Tool:**
 - Proactively improving benchmark % is financially defensible delays cost more than prevention.
 - Set tiered performance targets (e.g., 90% for high-penalty procedures) to balance patient access and fiscal sustainability.
 - Quantify financial ROI of operational improvements: e.g., a 5% benchmark gain in Cataracts could save >\$200K/year.

Multivariate prediction enables financially actionable insights. It not only forecasts where performance gaps may occur, but also quantifies the cost of inaction, allowing decision-makers to:

- Strategically target investments
- Measure and maximize value
- Align healthcare performance goals with financial stewardship

7. Business Impact and Conclusion

This chapter discusses the potential benefits and business impact of the developed predictive models and financial analyses, concluding with recommendations for healthcare administrators and policymakers.

7.1 Business Impact

The proposed solution, which incorporates predictive modeling, statistical inference, and data visualization, offers significant value to healthcare administrators and policy makers in Canada. Key business benefits include:

- **Enhanced Financial Prudence and Resource Efficiency:** Forecasting allows proactive resource allocation, optimizing scheduling, staffing, and equipment. Predicting future costs of delays provides a basis for investing in preventative measures, averting higher expenses from prolonged patient stays or worsening conditions. Analysis shows financial risk varies across provinces and procedures, enabling targeted budget allocation.
- **Strategic Policy Planning and Investment Justification:** Integrating predictive insights with financial impact estimates enables robust cost-benefit analyses for interventions. Quantifying financial risk from unmet benchmarks strengthens business cases for strategic investments in high-risk areas, ensuring fiscally responsible public health investments.
- **Improved Patient Management and Experience:** Models support better internal planning by enabling more accurate wait time forecasts. This helps administrators set realistic patient expectations, reducing anxiety and indirect financial burdens on patients.
- **Proactive Risk Management and Accountability:** Predicting financial risk from unmet targets allows early identification of performance gaps. This facilitates

proactive interventions to prevent financial penalties and patient dissatisfaction, fostering greater accountability.

- **Data-Driven Decision Making:** The project provides a robust framework for strategic investment targeting, value measurement, and aligning healthcare performance goals with fiscal stewardship.

7.2 Conclusion and Recommendations

This report comprehensively analyzed Canadian wait times for priority procedures, developing predictive capabilities and financial estimations. Key conclusions include the significant impact of the calendar year on wait times, showing an increasing trend and notable variations across provinces like Quebec and Ontario, and dramatic differences in delays by procedure type, with urgent procedures being short and elective surgeries substantially longer. A weak but statistically significant positive correlation was observed between procedure volume and wait days. While the Random Forest model showed numerical improvements over a naïve baseline for wait time prediction, these were not statistically significant; however, it performed reasonably well for volume and benchmark attainment. Crucially, the project successfully quantified and demonstrated meaningful variations in estimated financial costs and risks of delays across procedures, years, and provinces.

For Canadian healthcare administrators and policymakers, several recommendations emerge. They should prioritize targeted interventions for elective surgeries with long waits by increasing capacity and optimizing schedules. Region-specific strategies are crucial, utilizing provincial data to identify unique bottlenecks and allocate resources based on inter-provincial disparities. Embracing predictive analytics and continuously refining models through advanced time-series methods, enhanced feature engineering, and addressing statistical issues will lead to more robust predictions. Furthermore, integrating quantified costs and risks of

delays into performance monitoring, using tiered targets for financially justifiable improvements, and ensuring continuous data improvement (e.g., granular cost-per-day) will enhance financial analysis accuracy. Systematic adoption of these recommendations will lead to more informed, fiscally responsible decisions, ultimately reducing wait times, improving patient outcomes, and ensuring a more sustainable healthcare future for Canada.

References

1. Bodenheimer, T., & Pham, H. H. (2010). Primary Care: Current Problems And Proposed Solutions. *Health Affairs*, 29(5), 799–805.
<https://www.healthaffairs.org/doi/abs/10.1377/hlthaff.2010.0026>
2. Canadian Institute for Health Information. (n.d.). Explore wait times for priority procedures across Canada. CIHI. <https://www.cihi.ca/en/explore-wait-times-for-priority-procedures-across-canada>
3. Eka Kristia Ayu Astuti, Ayun Sariatmi, & Farid Agushybana. (2022). REVIEW: FACTORS AFFECTING WAITING TIME FOR OUTPATIENT PRESCRIPTION DRUGS IN HOSPITALS. *Medical Sains : Jurnal Ilmiah Kefarmasian*, 7(4), 861–872. <https://doi.org/10.37874/ms.v7i4.530>
4. Mills, K., Dhariwal, M., & Gill, R. (2022). SA48 Patient Impact and Economic Burden of Waiting for Cataract Surgery in Canada: A Targeted Literature Review. *Value in Health*, 25(7), S613.
<https://doi.org/10.1016/j.jval.2022.04.1714>
5. Moir, M. (2025, May 6). The Private Cost of Public Queues for Medically Necessary Care, 2025. Retrieved from Fraser Institute website:
<https://www.fraserinstitute.org/studies/private-cost-public-queues-medically-necessary-care-2025>
6. Murray, M., & Berwick, D. (n.d.). Advanced Access Reducing Waiting and Delays in Primary Care. In INNOVATIONS IN PRIMARY CARE.
<https://www.hqontario.ca/Portals/0/Documents/qi/learningcommunity/Roadmap%20Resources/Advanced%20Access%20and%20Efficiency/Step%201/pc-reducing-wating-and-delays-inprimary-care-jama-en.pdf>

7. Wong, J., Milroy, S., Sun, K., Iorio, P., Seto, M., Monakova, J., & Sutherland, J. (2020). Reallocating Cancer Surgery Payments for Alternate Level of Care in Ontario: What Are the Options? *Healthcare Policy | Politiques de Santé*, 16(2), 41–54. <https://doi.org/10.12927/hcpol.2020.26354>