Analyzing Reported Changes in Canadian Cannabis Consumption Post Legalization

December 5, 2023

Contents

1. *Project Introduction*

1.1 Abstract

This project will assess the impact of the legalization of cannabis in Canada on cannabis usage patterns in the population. Cannabis was legalized in Canada in October of 2018. Specifically, we are looking to see if there was a surge in cannabis users immediately following legalization and determine if there are any trends regarding cannabis usage following legalization. This analysis will provide insights into the long-term effects of cannabis legalization on usage trends and help inform public policy and healthcare strategies related to cannabis consumption going forward.

1.2 Introduction

This study investigates the implications of cannabis legalization in Canada on the patterns of cannabis usage within the population. With the landmark legalization taking effect in October 2018, our focus is determining whether there was a notable surge in cannabis users immediately following this event. Additionally, the study aims to identify and analyze trends in cannabis usage that emerged post legalization. The outcomes of this analysis should yield valuable insights into the enduring effects of cannabis legalization on usage trends. The results will offer crucial information to guide the development of public policies and healthcare strategies addressing cannabis consumption in the future.

1.3 Data Sources

Data was retrieved from Statistics Canada pertaining to cannabis usage across Canada. The data is collected via a survey, asking participants if they have consumed cannabis in the previous 3 months, meaning the data is self-reported and may not be entirely accurate. The data timeline spans from 2018 Q1 to 2020 Q4 with data missing from the first, second, and third quarter of 2020. This is inconvenient but thanks to the presence of 2020 Q4 data, it should still allow for analysis to determine if there has been a directional trend in cannabis usage. The data is segmented by province, sex, and age allowing for analysis from a variety of perspectives. The data for Canadian territories (Yukon, Northwest Territories, and Nunavut) is sparse so these locations will not be included in the analysis.

1.4   Business Questions

This project will answer several business questions such as:

- Did cannabis usage surge immediately following legalization?

- If there was a surge, was it sustained, or did it gradually decrease over time?

- Were specific demographics disproportionately affected by legalization

- What are the potential long-term implications of cannabis legalization on usage trends?

1.5   Process

The solution strategy for this project involves several key steps. First, data will be processed, ensuring the data is as clean and accurate as possible. Next, a descriptive analysis will be done to provide a high-level view of cannabis usage trends, the findings of which will be documented in this report. Lastly, the analysis will be concluded by summarizing the findings and providing evidence-based recommendations that can be used by policymakers and healthcare professionals to guide future cannabis related policies.

1.5.1   Data Preprocessing

The data is mostly complete, having been preprocessed by statisticians at Stats Canada prior to it being made publicly available. Imputing data for Q1, Q2, and Q3 of 2020 was considered but the data was too sparse to do so in an accurate fashion. Ultimately the data points would have little impact on the overall trend calculation this report intends to show. The many categories that were measured were grouped together in the column "GEO". The column was duplicated, renamed "Category", and the values were adjusted to make the plot legends they are used in more legible. Some example changes are:

- "Canada, 15 to 24" to "15 to 24" (this was done for all age brackets)

- "Canada, males" to "males" (this was done for all sexes)
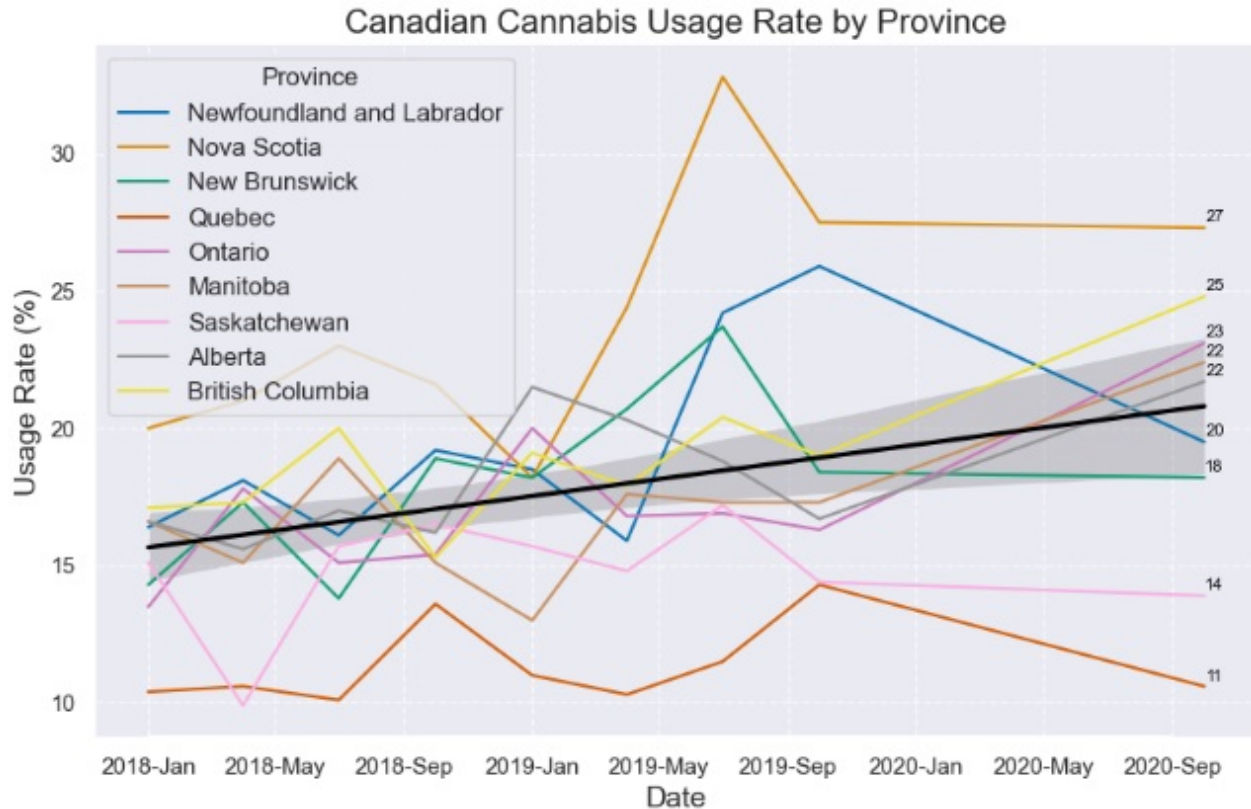
1.5.2   Data Visualization and Analysis

Figure 1.1: Provincial Usage Rates

Geography was the first dimension of the data that was plotted. Looking at the plot above, we can see that there is a positive trend in the data which indicates that as time goes on, the proportion of Canadian citizens who consume cannabis is expected to grow. Drilling down to the provincial level, there are several patterns that emerge immediately following legalization in October, 2018. Quebec, New Brunswick, Newfoundland and Labrador saw a pronounced increase in the proportion of population who consumed cannabis. Nova Scotia, British Columbia, and Manitoba reported a decrease in the number of cannabis users. Ontario, Alberta, and Saskatchewan reported little change. This wide variety of trends suggests that cannabis legalization did not effect all provinces equally. It also confirms that there was not a significant surge in users across Canada immediately following the legalization of cannabis.

Looking further ahead into 2019, all provinces with the exception of Alberta saw a significant increase in the proportion of cannabis users. One possible reason for this could be that it took time for cannabis dispensaries to become licensed. Once the method of distribution was established, the public rushed to engage with the novelty.

The data for 2020 is quite sparse with only one data point however it will be enough to

show the trend for the year. The one year anniversary of legalization saw the emergence of several trends. Ontario, Alberta, Manitoba, and British Columbia saw further growth throughout 2020. Nova Scotia, New Brunswick, and Saskatchewan plateaued, while Quebec, Newfoundland and Labrador saw a decrease in the proportion of cannabis users in their populations.

Although the provinces' populations reacted to the legalization of cannabis in various ways at various times, every province has seen an increase in the proportion of cannabis users since its legalization.
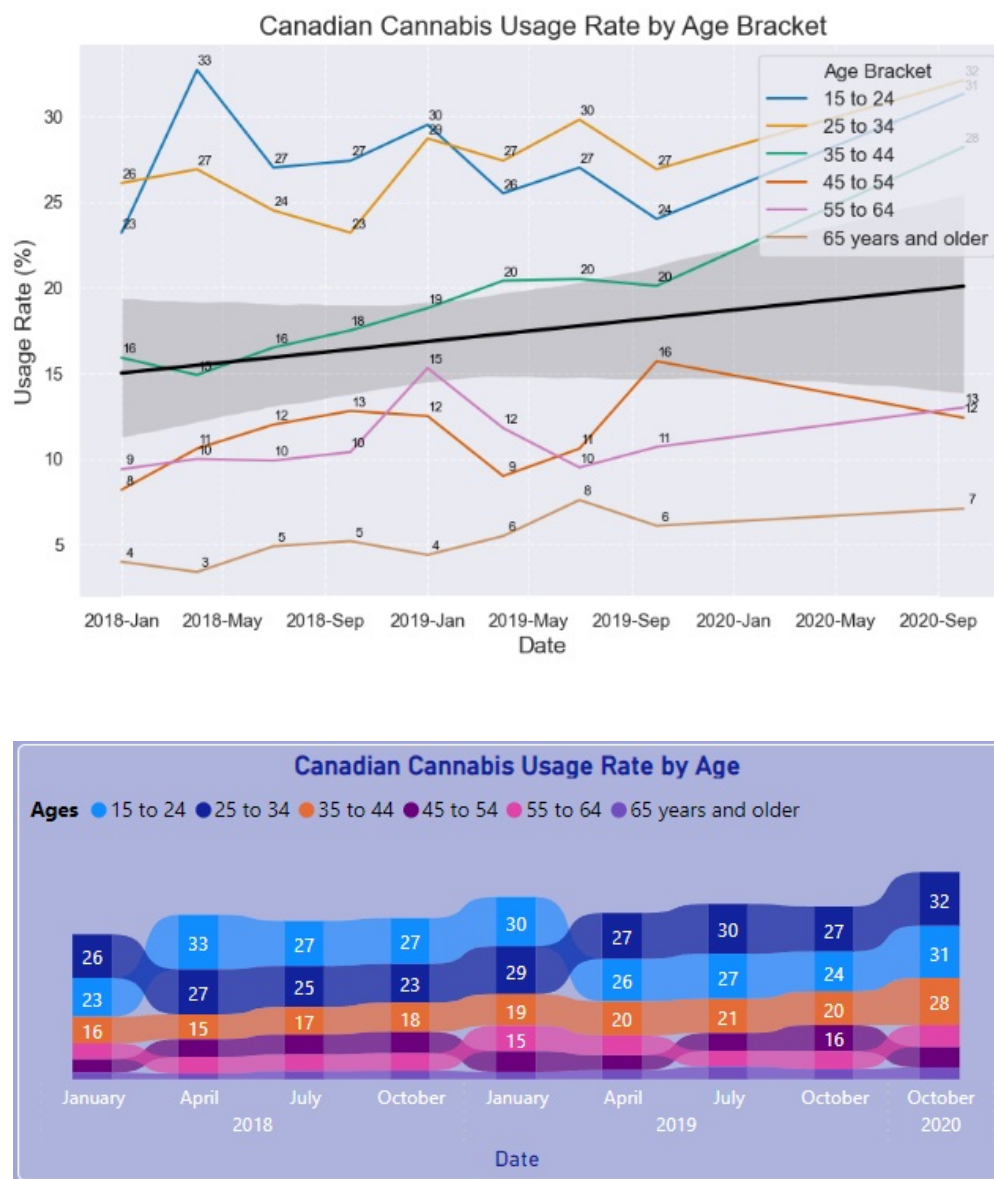

Figure 1.2: Usage Rate by Age Bracket

The next dimension that was analyzed was age. Immediately following legalization, the greatest surge in users was seen in the 25-34 and 55-64 age brackets. One could speculate that the younger crowd was interested in the novelty while the older crowd embraced cannabis for its medicinal properties. Although all age brackets saw an increase in cannabis users since legalization, the number of cannabis users in the three youngest age brackets grew the most. Interestingly, the rankings of the age brackets have stayed relatively consistent throughout the years being analyzed. This data suggests that people younger than 45 years of age are particularly attracted to cannabis. Policy makers interested in influencing cannabis users would likely benefit from tailoring their messages to this demographic.
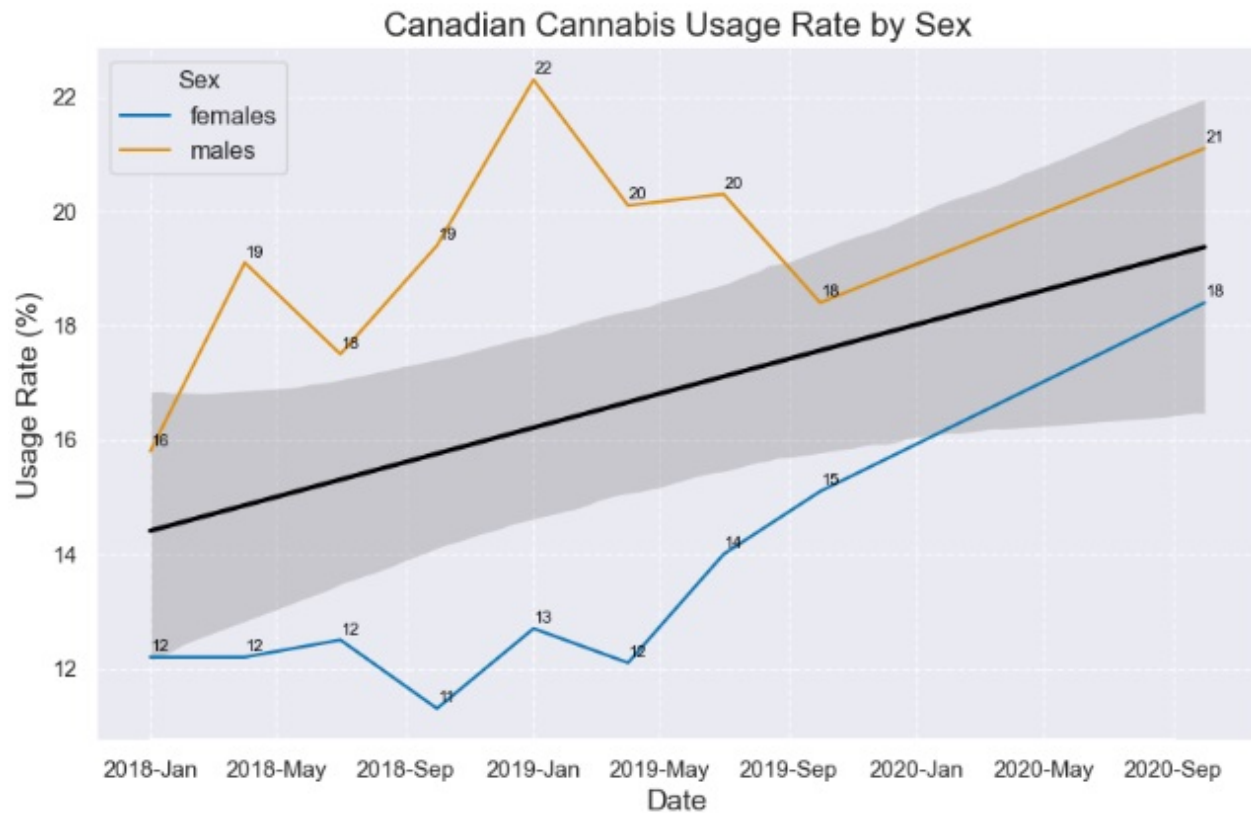

Figure 1.3: Usage Rate by Sex

The final dimension available for analysis was sex. Interestingly, following legalization the proportion of male users increased significantly for several months before pulling back. Females showed a small bump in users post-legalization, indicating some interest, but not nearly as much as the males. However, after the initial surge, females showed consistent growth in the proportion of cannabis users while the proportion of male users dipped before resuming the upward trend. These trends indicate that males are more likely than females to experiment with novelties the moment they become available. It is also possible that males are more willing to take the risk of purchasing cannabis from black market sources leading to an earlier peak while females waited for the dispensary infrastructure to be built. Ultimately

both sexes saw a similar increase in the proportion of users.

1.6   Conclusions

The data shows that cannabis usage is on the rise across Canada. Sex and age appear to be more influential in how users approach cannabis than geography. Cannabis users tend to be younger than 45 years of age. Males and females showed equal interest in cannabis although their method of approach differed slightly post-legalization. Policy makers would benefit from targeting their cannabis related messages towards people under the age of 45 in order to effectively reach the largest proportion of users.

1.7   Stakeholders

The legalization of cannabis in Canada involves a range of stakeholders including:

- **Government:** The federal government is responsible for implementing nation-wide cannabis legalization policies and regulations. Provincial governments implement and regulate cannabis laws within their jurisdiction, including distribution and retail frameworks.

- **Law Enforcement Agencies:** The police force is responsible for enforcing cannabis related laws, ensuring compliance with legal requirements.

- **Cannabis Industry:** Producers and retailers are responsible for supplying the populace with cannabis and are required to comply with legal requirements.

- **Consumers:** Medical and recreational users partake in the consumption of cannabis.

- **Healthcare Professionals:** Doctors and Pharmacists may be involved in the prescription and distribution of cannabis to medical users.

- **Public Health Organizations:** Organizations such as Health Canada are responsible for monitoring the health impacts of cannabis on the populace and creating educational campaigns to inform the populace on the effects of cannabis consumption.

- **Researchers:** Study the effects of cannabis use to inform policy makers.

- **Employers:** Businesses need to take cannabis into account when creating safety policies.

1.8   Practices to Ensure Data Integrity

Stats Canada implements a variety of measures to ensure the integrity of their data. For example, all data collected is depersonalized, surveys are designed to be clear and unbiased,

interviewers are trained to follow standardized procedures and remain unbiased, random sampling is implemented to ensure the data is reflective of the entire population, the data is validated by checking for outliers, inconsistencies, and errors. These checks ensure that the data is as accurate as possible by removing all sources of outside influence and then validated to ensure integrity.

2.   *Exploring Alternative Approaches*

In my initial conception of this project, I visualized an interactive PowerBI dashboard with helpful exploratory analysis panels. My alternative approach to communicating my findings takes a more formal reporting approach by using LaTeX to format my explanations and Python to preprocess, analyze, and visualize the data. I have listed the steps in my process below. The Jupyter Notebook file used to process and visualize the data is included separately in this submission.

In my exploration of this approach, I was pleasantly surprised by the professional aesthetics that can be achieved with LaTeX. LaTeX allows for precise formatting to create a polished look that overshadows the friendly, but casual blurbs that I would be partial to using in PowerBI.

In a similar vein, I appreciated the freedom that comes along with using the Python environment. Although the application is effective, I find that using solely PowerBI can limit your artistic freedom when creating visualizations. I find most of the plot visualizations I created with Python to be much more visually pleasing while providing more information than the plots created in PowerBI. The only exception I found was PowerBI's ribbon plot. Ultimately I may choose LaTeX over PowerBI, but I am likely to use a combination of visualization tools in my final submission. Below you can see the ribbon plot, followed by several direct comparisons of PowerBI to Python visuals.
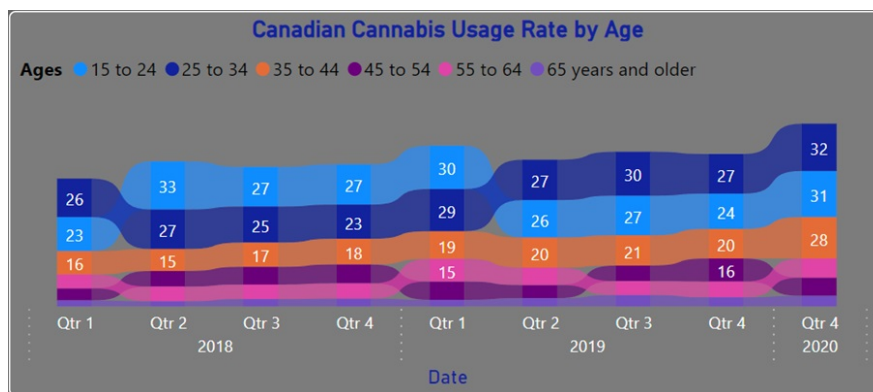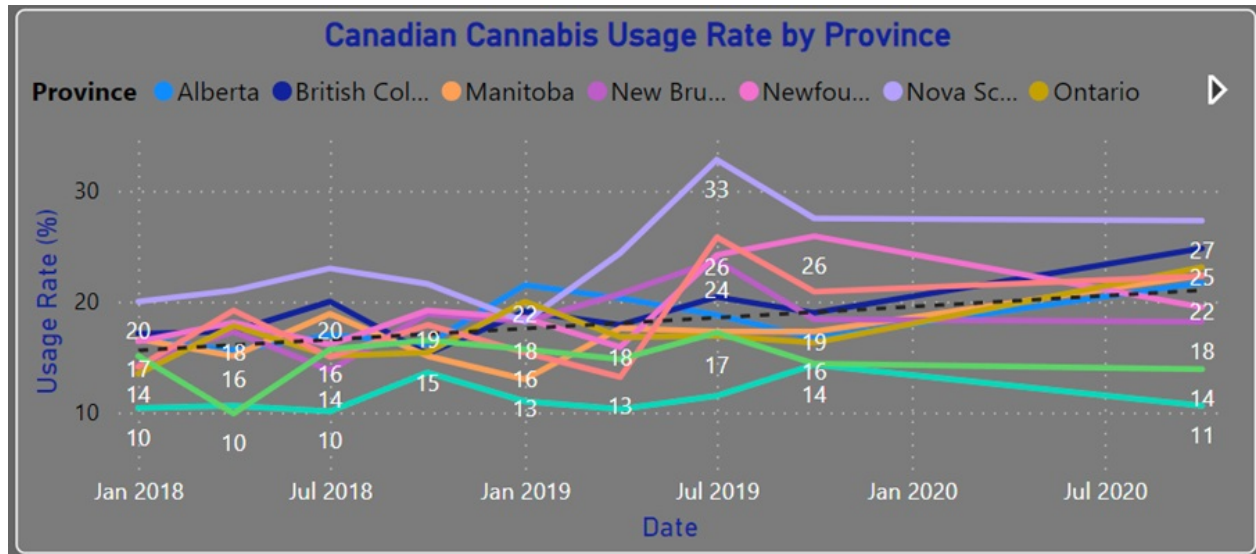

Figure 2.1: PowerBI Ribbon Chart
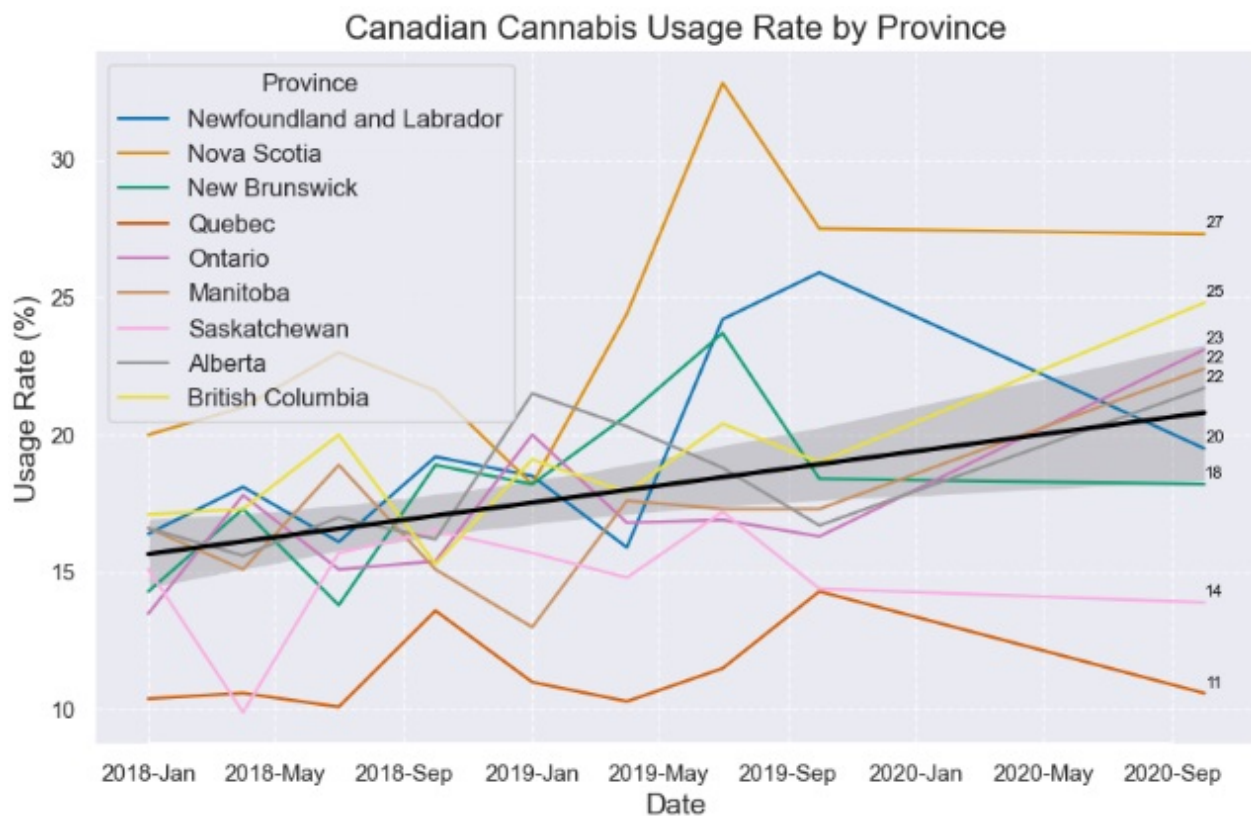
Figure 2.2: PowerBI Line Plot
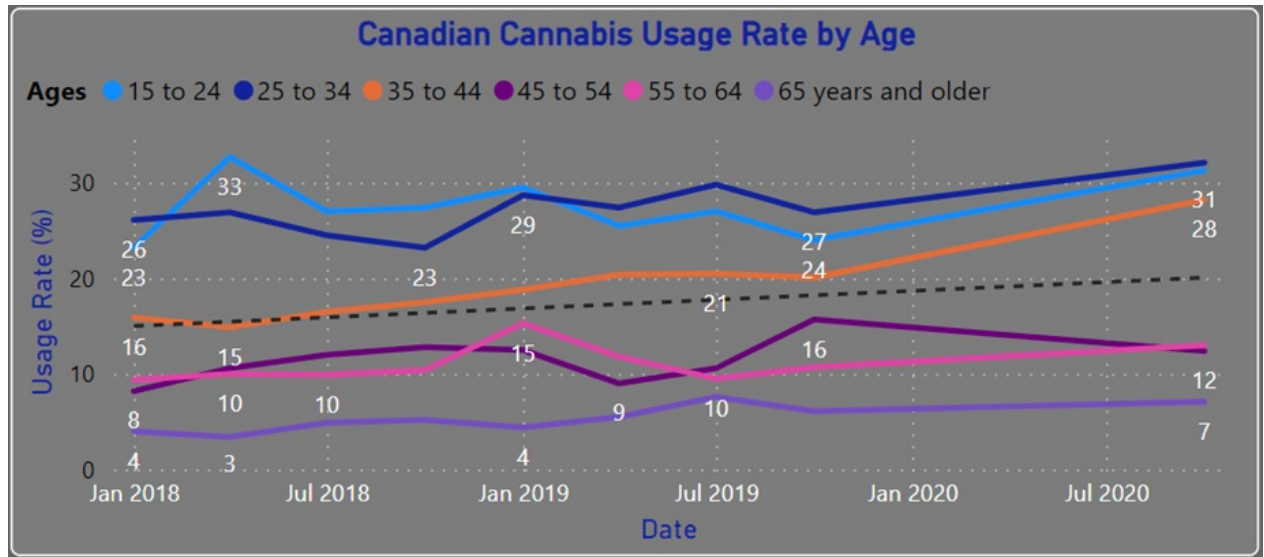

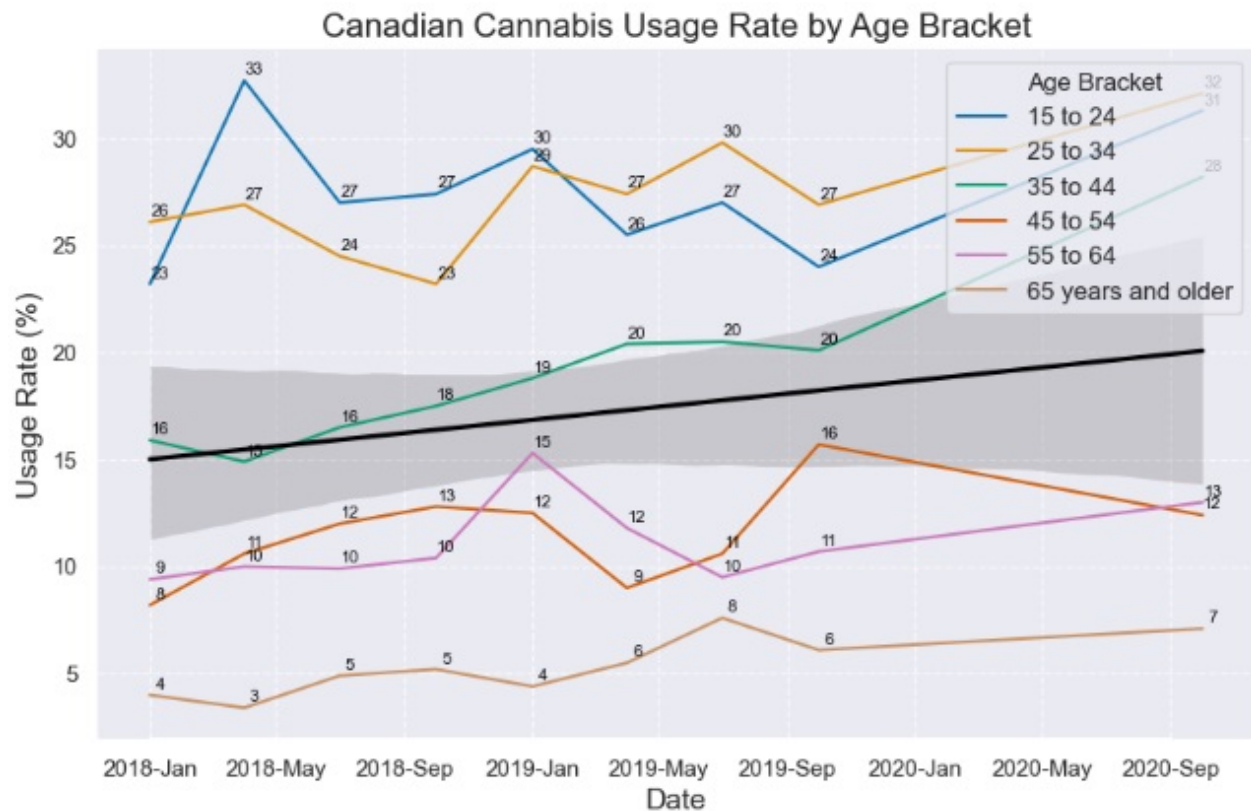Figure 2.3: Python Line Plot

Figure 2.4: PowerBI Line Plot

Figure 2.5: Python Line Plot

2.1   Actions and processes to support implementation of alternative detailed

I have chosen a more traditional reporting approach in lieu of the initially envisioned PowerBI dashboard. LaTeX will be employed to structure and format the report, providing a polished and professional document. Python will be used to handle data preprocessing, analysis, and visualization, via a Jupyter Notebook.

- **Data Preprocessing:** The same modifications were made to the data using Python's Pandas library instead of Microsoft's Power Query tool.

- **Data Analysis:** The Seaborn library was used to calculate a regression line, providing a quick visual indication of the direction the data is trending.

- **Data Visualization:** The Matplotlib and Seaborn libraries were used to create and stylize the visualizations.

- **Reporting:** LaTeX was used to create a crisp, clean document to communicate findings.

2.2   More visualizations used along with data exploration using different dimensions

I believe that line and ribbon plots are the best visualization tools to communicate the data available. For comparison purposes I have added a bar chart below. The data currently encompasses the dimensions of geography, age, and sex, all of which have been explored. Ideally the data collection process would be reworked to include a user classification system that describes the amount and/or frequency of cannabis consumed. I believe evolving the data in this manner would provide considerable depth to the amount of available insights.
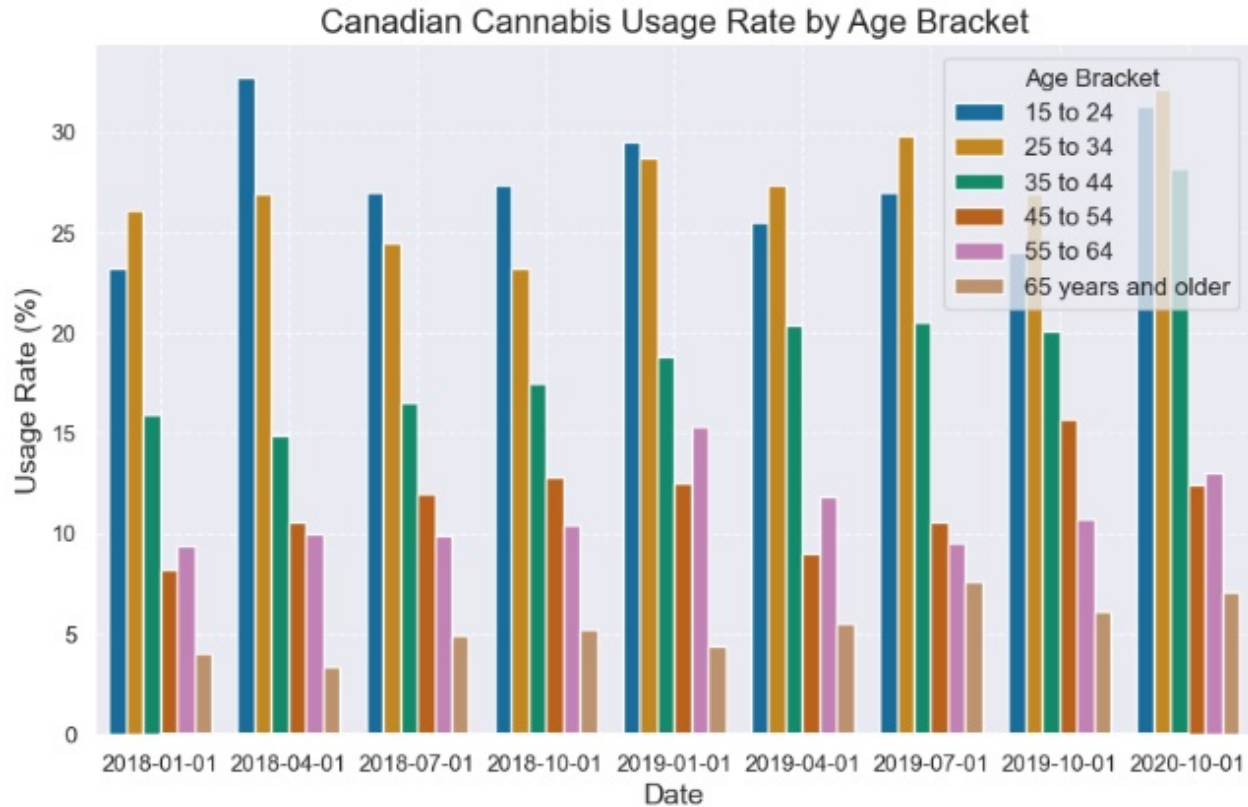
Figure 2.6: Python Bar Chart

2.3   Gap Analysis

2.3.1   Initial Plan vs. Revised Approach:

- **Communication Method:**

  - **PowerBI Dashboard (Initial Plan):** I originally envisioned an interactive dashboard providing a visually engaging experience with the option to explore further.

  - **LaTeX and Python (Revised Approach):** The concept shifted towards a more formal and structured reporting approach using LaTeX for document formatting and Python for data preprocessing and visualization.

- **Interactivity:**

  - **PowerBI Dashboard (Initial Plan):** Enables the use of interactive panels for exploratory analysis, allowing users to dynamically engage with the data.

  - **LaTeX and Python (Revised Approach):** Limited to a static but well-documented report format suitable for detailed analysis and formal presentation.

- **Presentation Style:**

  - **PowerBI Dashboard (Initial Plan):** The application is geared towards a dynamic experience with a focus on visual appeal.

  - **LaTeX and Python (Revised Approach):** Using these tools allows for the creation of a polished and professional document to communicate findings.

2.3.2   Tool Selection:

- **LaTeX:**

  - **Strengths:**
    * **Professional Aesthetics:** LaTeX offers fringeless font-formatting freedom, enabling users to customize their typeset for a polished and professional look.
    * **Document Structuring:** Allows for precise document structuring, simplifying the process of creating formal reports with sections, subsections, and captions.

  - **Weaknesses:**
    * **Learning Curve:** Has a steeper learning curve than PowerBI.
    * **Limited Interactivity:** LaTeX is designed for static documents, limiting user/reader interactivity.

- **Python:**

  - **Strengths:**
    * **Data Processing Power:** Python's libraries offer robust capabilities for data preprocessing and visualization.
    * **Flexibility:** Python's diverse library base enables programmers to create customized visualizations that would not be possible with PowerBI alone.

  - **Weaknesses:**
    * **Visual Appeal:** Matplotlib and Seaborn offer great visualization customization but are limited in some cases such as ribbon charts.

- **Consideration Factors:**

  - **Flexibility:** Python provides a high degree of flexibility, making it easy to analyze and visualize data. LaTeX's document formatting is unmatched but it does not permit dynamic data exploration. PowerBI allows for interactive analysis.

  - **Reproducibility:** The use of Jupyter Notebooks in Python makes it easy to follow the steps taken to reach the final product. Similarly, LaTeX code can be read to understand the process that was followed. In contrast, it is much more difficult to follow the development process of a PowerBI report as there is no explicit log of actions taken.

  - **Communicating Findings:** Python's libraries enable programmers to create intricate visualizations and perform in-depth data analysis which can be communicated clearly in a LaTeX document. This is a significant advantage when conveying complex findings. PowerBI may be more reliant on a presenter to convey complex insights to a non-technical or unfamiliar audience.

2.3.3   Methods to reach the target goal applied to overcome gaps

2.3.4   Skill Development:

- I have and will continue to practise using Python for data analysis and visualization.

- I have some previous but limited experience with LaTeX and will continue to actively develop my skills to ensure the documents I create are of professional quality.

2.3.5   Collaboration and Feedback:

- I have discussed my possible approaches with peers to determine public sentiment.

2.3.6   Iterative Development:

- I have embraced an iterative approach to this project in all aspects of data analysis and reporting.

By adopting a LaTeX and Python-based approach, I aim to provide a robust and well-documented analysis of cannabis usage in Canada, fostering a deeper understanding of the trends and factors influencing post-legalization consumption. The detailed Jupyter Notebook accompanying this report serves as a transparent and reproducible foundation for the presented findings.

3.   *References*

"Prevalence of Cannabis Use in the Past Three Months, Self-Reported." Government of Canada, Statistics Canada, Government of Canada, Statistics Canada, 21 Apr. 2021, `www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310038301&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2018&cubeTimeFrame.endMonth=10&cubeTimeFrame.endYear=2020&referencePeriods=20180101%2C20201001`.