**BUS 5303 - Big Data 2**

**Group Project - The Quality of Life IN Canada**

**Date: Aug 12, 2023,**

**Group Members: JATIN MEHRA   N01530276**

**VOLGA SHAJI    N01531163**

**HUSEYIN ISIK    N01482553**

**NAKUL SHARMA  N01530376**

**RUI SAI N01534536**

**Presentation Link:** [**https://www.youtube.com/watch?v=EQ1E-AE3dps**](https://www.youtube.com/watch?v=EQ1E-AE3dps)

**0. Overview**

The Quality of Life Framework for Canada brings together data for more than 80 key indicators on the well-being of people in Canada. Our group chose 6 indicators in 5 dimensions: Prosperity, Health, Society, Environment, and Good Governance.  Overall, these findings underscore positive trends and concerns in Canada's quality of life.

**HEALTH**

1. **Distribution of death by age group**

This table provides Canadians and researchers with provisional data to monitor weekly death trends by age and sex in Canada. Given the delays in receiving the data from the provincial and territorial vital statistics offices, these data are considered provisional.

Data Source: [Provisional weekly death counts, by age group and sex (statcan.gc.ca)](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310076801)

* **Code**

library(ggplot2)

library(dplyr)

data <- read.csv("C:/Users/jijip/Downloads/1310076801\_databaseLoadingData (1).csv")

summary\_data <- data %>%

group\_by(`Age.at.time.of.death`) %>%

summarize(TotalDeaths = sum(VALUE))

print(summary\_data)

              pie\_chart <- ggplot(summary\_data, aes(x = "", y = TotalDeaths, fill =

              `Age.at.time.of.death`)) +

geom\_bar(stat = "identity", width = 1) +

coord\_polar("y", start = 0) +

theme\_void() +

labs(title = "Distribution of Deaths by Age Group")

print(pie\_chart)

A screenshot of a computer

Description automatically generated

A pie chart with numbers and text

Description automatically generated

**SOCIETY**

1. **Heat map of strong series of belonging to local community**

Below graph illustrates the sense of belonging to the local community, town, province and Canada and trust in people, by groups designated as visible minorities and selected sociodemographic characteristics, 2020.

Data Source: [Sense of belonging to the local community, town, province and Canada and trust in people, by groups designated as visible minorities and selected sociodemographic characteristics, 2020 (statcan.gc.ca)](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=4310006401)

* **Code**

library(ggplot2)

library(dplyr)

data <- read.csv("C:/Users/jijip/Downloads/4310006401\_databaseLoadingData.csv")

# Pivot the data to create a matrix suitable for heatmap

heatmap\_data <- data %>%

select(Visible.minority, Selected.sociodemographic.characteristics, Indicators, VALUE) %>%

spread(Indicators, VALUE)

heatmap\_plot <- ggplot(heatmap\_data, aes(x = Selected.sociodemographic.characteristics, y = Visible.minority)) +

geom\_tile(aes(fill = `Strong sense of belonging to the local community`), color = "white") +

scale\_fill\_viridis\_c() +

theme\_minimal() +

labs(title = "Heatmap of Strong Sense of Belonging to the Local Community",

x = "Selected Sociodemographic Characteristics",

y = "Visible Minority Group")

print(heatmap\_plot)

A screenshot of a computer

Description automatically generated

A graph of different colored squares

Description automatically generated

**LIFE SATISFACTION**

1. **Bubble chart of indicators over time**

Below graph depicts life satisfaction by gender and other selected sociodemographic characteristics.

Data Source: [Life satisfaction by gender and other selected sociodemographic characteristics (statcan.gc.ca)](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310084401&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2021&cubeTimeFrame.endMonth=01&cubeTimeFrame.endYear=2022&referencePeriods=20210101%2C20220101)

* **Code**

library(ggplot2)

library(dplyr)

data <- read.csv("C:/Users/jijip/Downloads/1310084401\_databaseLoadingData.csv")

data <- data %>%

filter(!is.na(VALUE))

bubble\_chart <- ggplot(data, aes(x = REF\_DATE, y = VALUE, size = VALUE, color = Indicators)) +

geom\_point(alpha = 0.7) +

scale\_size\_continuous(range = c(3, 20)) +

scale\_color\_brewer(palette = "Set1") +

labs(title = "Bubble Chart of Indicators Over Time",

x = "Year",

y = "Value",

caption = "Bubble size represents value") +

theme\_minimal()

print(bubble\_chart)

**A graph of different colored circles

Description automatically generated with medium confidence**

**PROSPERITY**

1. **GDP of different provinces of Canada in 2018 and 2022**

The provided graph illustrates the Gross Domestic Product (GDP) of various Canadian provinces in the years 2018 and 2022. This comparative analysis over a five-year period offers valuable insights into the developmental and growth trajectories of these provinces. By examining GDP as a key parameter, we gain valuable insights into the quality of life for residents of Canada, enabling a nuanced assessment of their socio-economic progress.

Data Source: Gross domestic product (GDP) at basic prices, by industry, monthly, growth rates <https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610043402>

* **Code**

library(ggplot2)

library(dplyr)

# Read the data from the CSV file

data <- read.csv("GDP.csv")

data

#applying necessary filter to data  to get only columns for 2022 GDP and Provinces

data\_filtered <- data %>% select("Geography", "X2022", "X2018")

ggplot(data = data\_filtered, aes(x = Geography, group = 1)) +

  geom\_line(aes(y = X2018, color = "2018"), size = 1) +

  geom\_line(aes(y = X2022, color = "2022"), size = 1, linetype = "dashed") +

  geom\_point(aes(y = X2018, color = "2018"), size = 3) +

  geom\_point(aes(y = X2022, color = "2022"), size = 3) +

  labs(title = "GDP of Different Provinces of Canada in 2018 and 2022",

       x = "Province",

       y = "GDP") +

  scale\_color\_manual(values = c("2018" = "blue", "2022" = "red"),

                     labels = c("2018", "2022")) +

  theme\_minimal() +

  theme(axis.text.x = element\_text(angle = 45, hjust = 1))

A graph on a computer screen

Description automatically generated

**POPULATION**

1. **Percentage of people living within 500 meters of public transport**

The below graph indicates the percentage of the population living within 500 meters of a public transport stop  in  different provinces in canada

Data Source: Canada Indicators For The Sustainable Development Goals  [(sdgcif-data-canada-oddcic-donnee.github.io)](https://sdgcif-data-canada-oddcic-donnee.github.io/11-4-1/)

* **Code**

install.packages("ggplot2")

install.packages("tidyr")

library(tidyr)

library(ggplot2)

data = read.csv("C:\\Users\\Nakul\\Downloads\\BUSDATA.csv")

data

ggplot(data, aes(x = NAME, y = VALUE, fill = NAME)) +

  geom\_bar(stat = "identity") +

  geom\_text(aes(label = VALUE), vjust = 3.0) +

  theme\_dark() +

  theme(axis.text.x = element\_text(angle = 45, hjust = 1)) +

  labs(title = "Percentage of the population living within 500 meters of a public transport stop", x = "Province/Territory", y = "Value") +

  scale\_fill\_discrete() +

  ylim(0, 100)

A graph of different colored bars

Description automatically generated

**GOOD GOVERNANCE**

1. **Crime Severity Index in 10 years**

       The Crime Severity Index (CSI) measures changes in the level of severity of crime in Canada from year to year. In the index, all crimes are assigned a weight based on their seriousness. More serious crimes are assigned higher weights, less serious offences lower weights. As a result, more serious offences have a greater impact on changes in the index. The PRCSI is standardized to 100 in base year 2006.

Data Source: Crime severity index and weighted clearance rates, Canada, provinces, territories and Census Metropolitan Areas <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510002601>

* **Code**

library(tidyverse)

library(skimr)

library(RColorBrewer)

data <- read.csv("C:/Users/Rui/Documents/Humber 2023 Summer/Big Data 2/GroupPJ/CrimeIndex.csv")

data

#Restructure the data table

crimedata <- data %>%

  select(c('REF\_DATE', 'Statistics', 'VALUE')) %>%

  filter(Statistics %in% c('Crime severity index', 'Violent crime severity index'))

print(crimedata)

#Modify the data type for graphs

crimedata$VALUE <- as.numeric(crimedata$VALUE)

crimedata$REF\_DATE <- as.Date(paste0(crimedata$REF\_DATE, "-01-01"))

#Create line graphs

theme\_set(theme\_gray(base\_size = 12, base\_family = "Verdana"))

crimedata %>%

  ggplot() +

  geom\_line(aes(x = REF\_DATE,

                y = VALUE,

                group = Statistics,

                color = Statistics),

            size = 1) +

  labs(x = "Years", y = "Index") +

  scale\_x\_date(date\_labels = "%Y", date\_breaks = "1 year")+

  ggtitle("Crime Severity Index")

A graph showing the growth of a company

Description automatically generated

1. **Conclusion**

* Canada, the age of death groups 65 to 84 and 85 and over are significantly major groups in terms of age with 81.4%.
* Public transport accessibility varies across provinces. On the other hand, the major cities have at least 50% access to public transportation in 500 m. which indicates a very good ratio.
* GDP Growth in 2022 is observed in 8 regions compared to 2018, while decline in 3 regions.
* Life satisfaction displays fluctuations with recent minor declines.
* Sense of belonging varies among visible minority groups, highlighting differing levels of community connection.
* In the last decade, the crime severity index has significantly increased, especially in terms of violent crimes.

Our findings result highlighted that Canada has continued to improve its standard of living over the last few decades. However, there are several important factors (e.g., sense of belonging and crime rates) that indicate that the quality of life in Canada has begun to deteriorate in recent years.

**infographic**

**A close-up of a chart

Description automatically generated**

**8. References**

Government Canada. “Provisional weekly death counts, by age group and sex.” *Statistics Canada*, 2023, https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310076801. Accessed 12 8 2023.

Government of Canada. “Crime severity index and weighted clearance rates, Canada, provinces, territories and Census Metropolitan Areas.” *Statistics Canada*, 2023, https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510002601. Accessed 12 8 2023.

Government Canada. “Gross domestic product (GDP) at basic prices, by industry, monthly, growth rates” *Statistics Canada*, 2023, https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610043402 Accessed 12 8 2023.

Government Canada. “Life satisfaction by gender and other selected sociodemographic characteristics” *Statistics Canada*, 2023, https://open.canada.ca/data/en/dataset/965f3214-de52-4ce3-b115-365f5f4eefe7 Accessed 12 8 2023.

Government Canada. “Sense of belonging to the local community, town, province and Canada and trust in people, by groups designated as visible minorities and selected sociodemographic characteristics” *Statistics Canada*, 2022, https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=4310006401 Accessed 12 8 2023.

Government Canada. “*Canada Indicators For The Sustainable Development Goals*.” (n.d.). <https://sdgcif-data-canada-oddcic-donnee.github.io/> Accessed 12 8 2023.