2023

THE DREAM DESTINY OF MIGRANTS SUFFERS
Inflation in Housing (Canada)

PROJECT – STRATIFIED SAMPLING -USING R-MARKDOWN

A.MAYURI(2348133)

INDEX

Canada - THE DREAM DESTINY OF MIGRANTS SUFFERS-Inflation in Housing

A.Mayuri(2348133)

2023-09-24

CASE STUDY

Canada the most promising dream of migrants over the decades currently undergoing the most tragic financial instability. This is due to the high inflation that is sky-rocketing the daily commodities and specifically the housing rents of the denizens.

INTRODUCTION

Canada has always aimed to be immigrant friendly country. specifically, the country has been a safe heaven for migrants who have suffered with wars, forced migration etc. It has played a significant part in changing the lives of migrants and skilled laborers. However, over the years, the situation has been changing due to the inflation, which is questioning the choices made by the migrants and the validity of these friendly public policies.

CAUSES FOR THE CRISIS

High Migration rate:

The annual immigration in Canada amounts to around 500,000 which is the highest rates per population of any country in the world. and almost 7.5 million immigrants live in Canada as permanent citizens. These statistics reveal the intensity of migration inflow that is practiced by the Canadian nation.

Lack of Infrastructure:

The high demand for housing due to the tremendous inflow of migrants over the year does not match with the houses that are built. Thus, accommodation is demanded commodity. This leads to a surge in demand for rental houses while the supply is close to be fixed thus resulting in the housing inflation.

Investment Pattern:

more in houses than in business rely on natural resources which is sold more over 55% to the USA, strict laws make it harder Canadian start ups to raise resources to build businesses. Thus small start-ups get absorbed/brought by the USA richer counterparts.This makes Canada a very dependent economy on the USA

OBJECTIVE:

1)Descriptive Data Analysis

2)Stratified Sampling

FORMULA:

$$n_h = \frac{N_h}{N} n$$

nh =	Sample size using proportionate stratified random sampling
Nh =	Total stratum population
N =	Total population
n =	Sample size (calculation results using the slovin formula)

Variance:

	Sample size to estimate a proportion	Sample size to estimate an average
Simple random sampling	$\frac{Z^2 p(1-p)}{e^2}$	$\frac{Z^2 \sigma^2}{e^2}$
Proportional stratified sampling	$\frac{Z^2 \sum_{h=1}^{L} W_h p_h (1-p_h)}{e^2}$	$\frac{Z^2 \sum_{h=1}^{L} W_h \sigma_h^2}{e^2}$
Best stratified sampling	$\frac{Z^2 (\sum_{h=1}^{L} W_h \sqrt{p_h (1-p_h)})^2}{e^2}$	$\frac{Z^2 (\sum_{h=1}^{L} W_h \sigma_h)^2}{e^2}$

VARIABLE OF INTEREST:

1) Year (1990/2016)-Time period (used to segregate the data set into different stratum)

2) region - The different districts/county that are considered for the data collection/census.

3)one_/two_/ three bedrooms(rent): The payment that is been charged to temporarily stay at an accommodation of 1/2/3 BHK apartment. the rent is charged at an monthly basis and are collected in Canadian dollars.

DEFINITIONS

Inflation: It is when prices for goods and services rise and purchasing power falls.

When inflation goes up, people and businesses must spend more money to buy the same amount of goods and services. “More money chasing fewer goods”.

Migration: The temporary or permanent change of residence of a set of people/animals in search of a better living standard, or any other reason that can be caused by push or pull factors.

Immigrant(s): A person(people) moving to another country by shifting his/her residence

METHODOLOGY

The project has followed a quantitative approach to understand the hike in rents of the Canadian housing. The project has utilized sources such as a secondary data set delivered by the Kaggle website for analysis, and research papers, books, past case studies, examples to understand the background and history of today situation.

SAMPLING TECHNIQUE:

We have used stratified sampling technique where we aim to segregate the dataset into different stratums or subsets.

STRATUM SELECTION:

Stratum segregates the whole population set by being heterogeneous between and homogeneous within.

For our study we have chosen 1990 and 2016 data as two strata since the price index and inflation will be constant in those years but are different for the two years given the duration difference.

ANALYSIS

STEP1: IMPORTING DATASETS

```
library(readxl)
PROJECT_1990 <- read_excel("C:/Users/mayur/Desktop/Mstat/tri sem 1/R/dataset/PROJECT_1990.xlsx")
View(PROJECT_1990)
attach(PROJECT_1990)


library(readxl)
PROJECT_2016 <- read_excel("C:/Users/mayur/Desktop/Mstat/tri sem 1/R/dataset/PROJECT_2016.xlsx")
View(PROJECT_2016)
attach(PROJECT_2016)
```

```
## The following objects are masked from PROJECT_1990:
##
##     employment_change, labour_participation_rate, migration,
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
```

```
library(readxl)
PROJECT <- read_excel("C:/Users/mayur/Desktop/Mstat/tri sem 1/R/dataset/PROJECT.xlsx")
View(PROJECT)
attach(PROJECT)
```

```
## The following objects are masked from PROJECT_2016:
##
##     employment_change, labour_participation_rate, migration,
```

```
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
##
## The following objects are masked from PROJECT_1990:
##
##     employment_change, labour_participation_rate, migration,
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
```

Step 2: DESCRIPTIVE DATA ANALYSIS

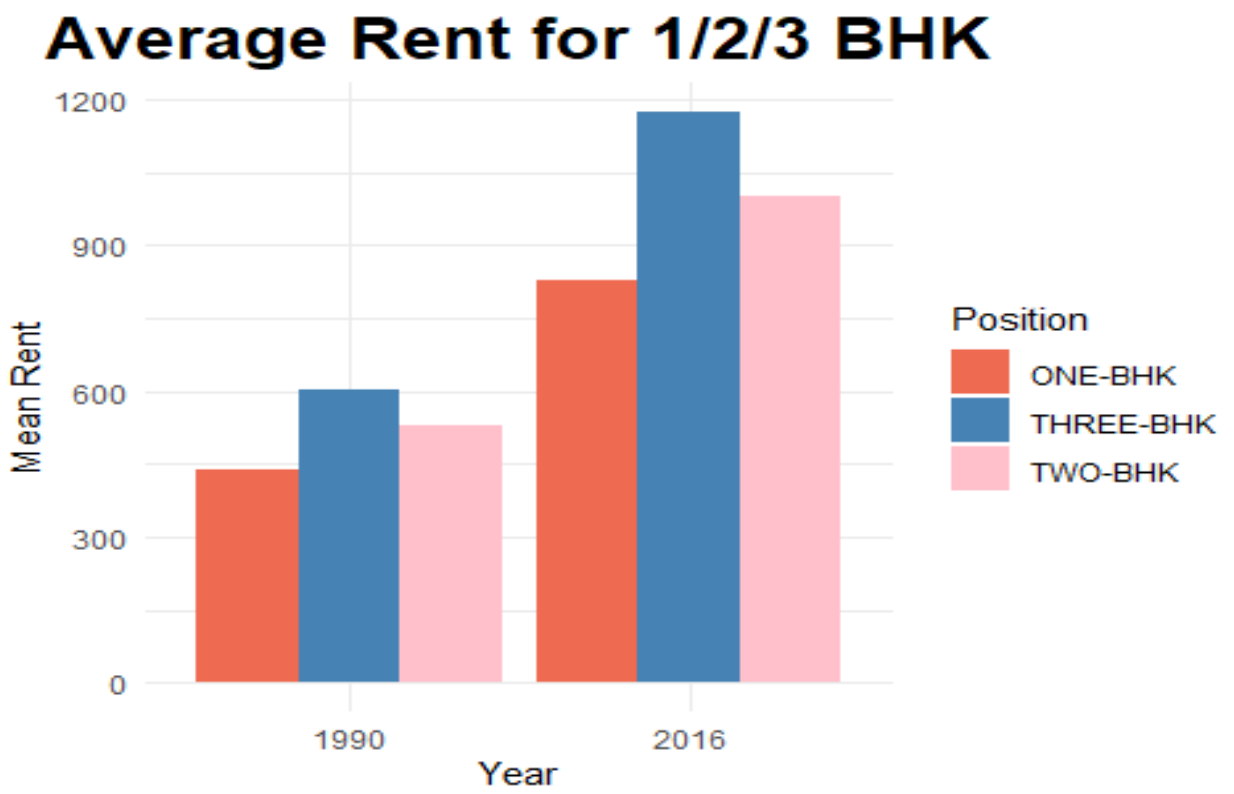
Let us try to understand the average(mean) rent for different sized houses over the two years. And we understand that the grouped bar graphs give the best depiction for the illustration of the same.

```
#mean for 1990 is in series of x
#mean for 2016 is in series of y

x1=mean(PROJECT_1990$one_bedroom)
x2=mean(PROJECT_1990$two_bedroom)
x3=mean(PROJECT_1990$three_bedroom)

y1=mean(PROJECT_2016$one_bedroom)
y2=mean(PROJECT_2016$two_bedroom)
y3=mean(PROJECT_2016$three_bedroom)
c1=c(x1,x2,x3)
c2=c(y1,y2,y3)
#Creating data-frame to do grouped bar graph

df <- data.frame(Year=rep(c('1990', '2016'), each=3),
                 BHK=rep(c('ONE-BHK', 'TWO-BHK', 'THREE-BHK'), times=2),
                 Mean_Rent=c(x1,x2,x3,y1,y2,y3))
df
```

```
##   Year       BHK Mean_Rent
## 1 1990   ONE-BHK  438.8286
## 2 1990   TWO-BHK  528.4286
## 3 1990 THREE-BHK  601.6286
## 4 2016   ONE-BHK  829.0588
## 5 2016   TWO-BHK  998.4706
## 6 2016 THREE-BHK 1174.9118
```

```
library(ggplot2)
```

```
ggplot(df, aes(fill=BHK, y=Mean_Rent, x=Year)) +
 geom_bar(position='dodge', stat='identity') +
 theme_minimal() +
 labs(x='Year', y='Mean Rent', title='Average Rent for 1/2/3 BHK') +
 theme(plot.title = element_text(hjust=0.5, size=20, face='bold')) +
 scale_fill_manual('Position', values=c('coral2', 'steelblue', 'pink'))
```



Interpretation of the graph:

We observe that the rent of 1/2/3BHK houses have raised far above over the years.Let us examine the same in much more specifics in the following analysis

STRATIFIED DATA ANALYSIS:

STEP3 : INITIALIZATION

```
library(samplingbook)
```

```
## Loading required package: pps
```

```
## Loading required package: sampling
```

```
## Loading required package: survey
```

```
## Loading required package: grid
```

```
## Loading required package: Matrix
```

```
## Loading required package: survival
```

```
##
## Attaching package: 'survival'
## The following objects are masked from 'package:sampling':
##
##     cluster, strata
##
## Attaching package: 'survey'
## The following object is masked from 'package:graphics':
##
##     dotchart
attach(PROJECT)
## The following objects are masked from PROJECT (pos = 11):
##
##     employment_change, labour_participation_rate, migration,
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
## The following objects are masked from PROJECT_2016:
##
##     employment_change, labour_participation_rate, migration,
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
## The following objects are masked from PROJECT_1990:
##
##     employment_change, labour_participation_rate, migration,
##     one_bedroom, population, region, three_bedroom, two_bedroom,
##     unemployment_rate, year
```

STEP4 : CREATING STRATUMS

For the stratum we have considered year 1990 as ‘0’ and year 2016 as ‘1’

```
#obtaining stratum size and stratum standard deviation (N1,N2,S1,S2)
stratum1=PROJECT[PROJECT$year==0, ]
stratum1 #creating stratum 1 for year 1990
```

```
## # A tibble: 35 × 10
##    year region       one_bedroom two_bedroom three_bedroom population
##   <dbl> <chr>              <dbl>       <dbl>         <dbl>      <dbl>
##  1     0 manitoba             418         524           608      1105.
##  2     0 vancouver            566         751           876      1608.
##  3     0 winnipeg             419         531           614       666.
##  4     0 calgary              456         584           591       750.
##  5     0 prince_edward        387         479           524       130.
##  6     0 ottawa               509         640           761       725.
##  7     0 oshawa               533         604           690       243.
##  8     0 saguenay             365         390           413       163.
##  9     0 hamilton             432         523           653       615.
## 10     0 abbotsford           450         588           608       143.
## # ℹ 25 more rows
## # ℹ 4 more variables: labour_participation_rate <dbl>, employment_change <dbl>,
## #   unemployment_rate <dbl>, migration <dbl>
```

```
stratum2=PROJECT[PROJECT$year==1, ]
stratum2 #creating stratum 1 for year 2016
```

```
## # A tibble: 34 × 10
##    year region       one_bedroom two_bedroom three_bedroom population
##   <dbl> <chr>              <dbl>       <dbl>         <dbl>      <dbl>
##  1     1 manitoba             823        1033          1289      1318.
##  2     1 vancouver           1159        1450          1631      2549.
##  3     1 winnipeg             836        1068          1327       812.
##  4     1 calgary             1050        1258          1258      1469.
##  5     1 prince_edward        696         852           922       149.
##  6     1 ottawa               982        1201          1457      1019.
```

```
##  7     1 oshawa              979       1109        1198       394.
##  8     1 saguenay            469        587         645       160.
##  9     1 hamilton            869       1037        1232       778.
## 10     1 abbotsford          744        915        1019       187.
## # ℹ 24 more rows
## # ℹ 4 more variables: labour_participation_rate <dbl>, employment_change <dbl>,
## #   unemployment_rate <dbl>, migration <dbl>
```

```
# Finding the number of observations for each stratum
N1=sum(stratum1$year==0)
N1
```

```
## [1] 35
```

```
N2=sum(stratum2$year==1)
N2
```

```
## [1] 34
```

```
# Finding the mean
MN_1B_1990=mean(stratum1$one_bedroom)
MN_2B_1990=mean(stratum1$two_bedroom)
MN_3B_1990=mean(stratum1$three_bedroom)

MN_1B_2016=mean(stratum2$one_bedroom)
MN_2B_2016=mean(stratum2$two_bedroom)
MN_3B_2016=mean(stratum2$three_bedroom)

MN_1B_1990
```

```
## [1] 438.8286
```

```
MN_2B_1990
```

```
## [1] 528.4286
```

```
MN_3B_1990
```

```
## [1] 601.6286

MN_1B_2016

## [1] 829.0588

MN_2B_2016

## [1] 998.4706

MN_3B_2016

## [1] 1174.912

# Finding the standard deviations
SD_1B_1990=sqrt(var(stratum1$one_bedroom))
SD_2B_1990=sqrt(var(stratum1$two_bedroom))
SD_3B_1990=sqrt(var(stratum1$three_bedroom))

SD_1B_2016=sqrt(var(stratum2$one_bedroom))
SD_2B_2016=sqrt(var(stratum2$two_bedroom))
SD_3B_2016=sqrt(var(stratum2$three_bedroom))

SD_1B_1990

## [1] 59.40315

SD_2B_1990

## [1] 84.55882

SD_3B_1990

## [1] 114.6458

SD_1B_2016

## [1] 170.2419

SD_2B_2016

## [1] 204.9877
```

```
SD_3B_2016
```

```
## [1] 260.7633
```

STEP5 : USING OPTIMIMUM ('OPT') FOR SELECTING THE SAMPLES.

```
#let the sample of size n=10 has to be drawn using proportional allocation
sample_size=stratasamp (n=10, Nh=c(N1, N2), Sh=c(SD_1B_1990, SD_1B_2016), type="opt"
)
sample_size
```

```
##
## Stratum 1 2
## Size    3 7
```

```
stratasize(e=.1, Nh=c(N1, N2), Sh=c(SD_1B_1990, SD_1B_2016), type="opt" )
```

```
##
## stratamean object: Stratified sample size determination
##
## type of sample: opt
##
## total sample size determinated: 56
```

STEP6: DETERMINING SAMPLE SIZE OF 7 FROM BOTH STRATUMS

```
#collect a random sample of size 5,5 from both strata
sample1=stratum1[sample(1:nrow(stratum1), 7, replace=FALSE), ]
sample1 # sample 1 collected from stratum 1
```

```
## # A tibble: 7 × 10
##    year region     one_bedroom two_bedroom three_bedroom population
##   <dbl> <chr>          <dbl>       <dbl>         <dbl>      <dbl>
## 1     0 moncton          377         444           464       122.
## 2     0 saguenay         365         390           413       163.
## 3     0 sudbury          399         489           547       160.
## 4     0 saint_john       340         402           427       128.
## 5     0 calgary          456         584           591       750.
```

```
## 6     0 guelph           486        561         606      129.
## 7     0 montreal         419        463         541     3270.
## # ℹ 4 more variables: labour_participation_rate <dbl>, employment_change <dbl>,
## #   unemployment_rate <dbl>, migration <dbl>
```

```
sample2=stratum2[sample(1:nrow(stratum2), 7, replace=FALSE), ]
sample2 # sample 1 collected from stratum 2
```

```
## # A tibble: 7 × 10
##    year region     one_bedroom two_bedroom three_bedroom population
##   <dbl> <chr>          <dbl>       <dbl>         <dbl>      <dbl>
## 1     1 halifax          845        1063          1288       426.
## 2     1 victoria         912        1188          1485       371.
## 3     1 sudbury          776         990          1111       166.
## 4     1 ottawa           982        1201          1457      1019.
## 5     1 sherbrooke       502         622           764       216.
## 6     1 abbotsford       744         915          1019       187.
## 7     1 kingston         942        1119          1798       171.
## # ℹ 4 more variables: labour_participation_rate <dbl>, employment_change <dbl>,
## #   unemployment_rate <dbl>, migration <dbl>
```

STEP7 : BINDING THE SAMPLES TOGETHER

```
total_sampled_data=rbind(sample1, sample2)
total_sampled_data # total sample collected using stratified random sampling
```

```
## # A tibble: 14 × 10
##    year region     one_bedroom two_bedroom three_bedroom population
##   <dbl> <chr>          <dbl>       <dbl>         <dbl>      <dbl>
## 1     0 moncton          377         444           464       122.
## 2     0 saguenay         365         390           413       163.
## 3     0 sudbury          399         489           547       160.
## 4     0 saint_john       340         402           427       128.
## 5     0 calgary          456         584           591       750.
```

```
##  6    0 guelph          486       561         606      129.
##  7    0 montreal        419       463         541     3270.
##  8    1 halifax         845      1063        1288      426.
##  9    1 victoria        912      1188        1485      371.
## 10    1 sudbury         776       990        1111      166.
## 11    1 ottawa          982      1201        1457     1019.
## 12    1 sherbrooke      502       622         764      216.
## 13    1 abbotsford      744       915        1019      187.
## 14    1 kingston        942      1119        1798      171.
## # ℹ 4 more variables: labour_participation_rate <dbl>, employment_change <dbl>,
## #   unemployment_rate <dbl>, migration <dbl>
```

```
# Estimation of population mean using stratified random sample
nh=as.vector(table(total_sampled_data$year))
nh
```

```
## [1] 7 7
```

```
wh=nh/sum(nh)
wh
```

```
## [1] 0.5 0.5
```

```
#calculating strata means for all three forms of houses
ST_mean_1B=stratamean(y=total_sampled_data$one_bedroom, h=as.vector(total_sampled_da
ta$year),
wh=wh, eae=TRUE)
ST_mean_1B   #1bhk
```

```
##            Mean       SE      CIu      CIo
## 0      406.0000 19.51556 367.7502 444.2498
## 1      814.7143 61.50057 694.1754 935.2532
## overall 610.3571 32.26134 547.1261 673.5882
```

```
ST_mean_2B=stratamean(y=total_sampled_data$two_bedroom, h=as.vector(total_sampled_da
ta$year),
```

```
wh=wh, eae=TRUE)
ST_mean_2B   #2bhk

##             Mean       SE      CIu       CIo
## 0       476.1429 28.08273 421.1017  531.1840
## 1      1014.0000 76.01879 865.0059 1162.9941
## overall  745.0714 40.52005 665.6536  824.4893

ST_mean_3B=stratamean(y=total_sampled_data$three_bedroom, h=as.vector(total_sampled_d
ata$year),
wh=wh, eae=TRUE)
ST_mean_3B  #3bhk

##             Mean       SE       CIu       CIo
## 0       512.7143  29.46727  454.9595  570.4691
## 1      1274.5714 129.51133 1020.7339 1528.4090
## overall  893.6429  66.41066  763.4804 1023.8054

stratamean(y=total_sampled_data$one_bedroom, h=as.vector(total_sampled_data$year),
wh=wh)

##
## stratamean object: Stratified sample mean estimate
## Without finite population correction.
## Mean estimate: 610.3571
## Standard error: 32.2613
## 95% confidence interval: [547.1261,673.5882]

stratamean(y=total_sampled_data$two_bedroom, h=as.vector(total_sampled_data$year),
wh=wh)

##
## stratamean object: Stratified sample mean estimate
## Without finite population correction.
## Mean estimate: 745.0714
```

```
## Standard error: 40.52
## 95% confidence interval: [665.6536,824.4893]

stratamean(y=total_sampled_data$three_bedroom, h=as.vector(total_sampled_data$year),
wh=wh)

##
## stratamean object: Stratified sample mean estimate
## Without finite population correction.
## Mean estimate: 893.6429
## Standard error: 66.4107
## 95% confidence interval: [763.4804,1023.805]
```

CONCLUSION

It is observable that the mean rent has been increased significantly in all three forms of general accommodation (1/2/3 BHK) over the years. 406 to 804 Canadian dollars for 1BHK which is approximately twice, and a similar pattern for the 2/3 BHK accommodation. It is vital to build in more infrastructure for the situation to be reduced of its burden.

FUTURE SCOPE AND PROSPECTIVE ANALYSIS USING THE DATASET

-*Time series analysis*: We have considered only two-year data set since we are doing a stratified sampling however we can do a time series analysis if we had to understand the trend of increase over the years.

-*Simulation*: We can do a simulation using the historical data set, adjust it to current inflation to predict how the trend might look at the future. However it is to note that simulation will give all possibilities with different changes but not the exact truth of the future nor the optimal solution.

REFORMS/SUGGESTIONS FOR THE CANADIAN INFLATION

The government and Banks are a pivot player in maintaining the stability of a country's economy. It is vital to have a healthy international understanding with other countries and their

people. However, it is the government's responsibility to forecast towards the future and understand the pros and cons before implementation of the public policies.

Moreover, It is a truth that countries cannot engage without international cooperation on trade and commerce, However they should analyze their capacity towards self-sufficiency, especially in having and creating start-ups and adapting to the requirements of the future.

Countries should also be flexible with external and internal policies, without compromising the security and identity of the nation.

REFERENCE

https://livingcost.org/cost/canada#:~:text=The%20cost%20of%20living%20in%20Canada%20is%20%241901%2C,enough%20to%20cover%20living%20expenses%20for%201.7%20months.

https://www.ctvnews.ca/canada/millions-of-people-will-move-to-canada-in-the-next-20-years-but-where-will-they-live-1.5464205

https://cbc.ca/news/business/canada-inflation-rate-1.6526060

https://www.statista.com/topics/2917/immigration-in-canada/#topicOverviewhttps://www.cbc.ca/news/politics/inflation-political-parties-solutions-1.6586736

https://globalnews.ca/news/9970539/canada-inflation-august-2023/

https://www.bankofcanada.ca/2023/03/understanding-the-reasons-for-high-inflation/


THE END


100