2017

Program Exit Assessment

guy Beaulieu

Program Exit Assessment

Philip Dumaresq

Contents

[Figures 1](#_Toc501832731)

[Task 1 2](#_Toc501832732)

[Introduction 2](#_Toc501832733)

[Comparing the normalized data from Canada and the US. 2](#_Toc501832734)

[Summarizing the raw data from Canada and the US 6](#_Toc501832735)

[Comparing population groups 7](#_Toc501832736)

[Task 2 9](#_Toc501832737)

[Task 3 11](#_Toc501832738)

[Confidence intervals 11](#_Toc501832739)

[Success proportion values 13](#_Toc501832740)

# Figures

[Figure 1: Normalized homicide data in Canada and the US 2](#_Toc501832719)

[Figure 2: Percentage of firearm homicides committed using a handgun 3](#_Toc501832720)

[Figure 3: Percentage of total homicides committed using a firearm 3](#_Toc501832721)

[Figure 4: Normalized firearm homicide data in Canada and the US 4](#_Toc501832722)

[Figure 5: Normalized total Canadian homicide data and firearm homicides 4](#_Toc501832723)

[Figure 6: Normalized total American homicides and firearm homicides 5](#_Toc501832724)

[Figure 7: Normalized data comparing firearm homicide number in Canada and the US 7](#_Toc501832725)

[Figure 8: Normalized data comparing total number of homicides in Canada and the US. 8](#_Toc501832726)

[Figure 9: Normalized Canadian data for homicides with firearms vs total homicides 9](#_Toc501832727)

[Figure 10: Normalized American data for homicides with firearms vs total homicides 10](#_Toc501832728)

[Figure 11: Relative frequency of Canadian firearm homicides 11](#_Toc501832729)

[Figure 12: Relative frequency of American firearm homicides 12](#_Toc501832730)

# Task 1

## Introduction

The data given to us is pulled from Stats Canada and the FBI website about homicides in Canada and the Unites stated. The data is from 2005 to 2014 and includes the method used to commit the homicide, breaking it down primarily into total homicides, firearm homicides and then more specifically, handgun homicides. There is other data for different types of homicide included as well, but that data will not be the focus. This data was collected when the method of death was confirmed by police officials and once it had been confirmed that the death was a homicide related incident. This data was taken so that the Canadian and US governments know where these crimes are happening in their countries so that they can make an effort into putting money where it’ll make a bigger difference. By providing more funding to police departments where more homicides typically happen, they can try and reduce the amount if fatal crime in different regions.

## Comparing the normalized data from Canada and the US.

Figure 1: Normalized homicide data in Canada and the US

Figure 2: Percentage of firearm homicides committed using a handgun

Figure 3: Percentage of total homicides committed using a firearm

Figure 4: Normalized firearm homicide data in Canada and the US

Figure 5: Normalized total Canadian homicide data and firearm homicides

Figure 6: Normalized total American homicides and firearm homicides

These 5 bar charts represent the normalized data of total homicides, firearm related homicides, and specifically handgun homicides in Canada and the United States. We can see very clearly in the five bar charts that the United States has over twice as many homicides per million people than Canada every year. While the United States maxed out ~50 homicides per 1 million people, Canada only hit ~21. In the next two charts, we start looking at firearm and specifically handgun homicides. We can note two things from this normalized data. The first of which, is that if the number of guns owned, the types of guns owned, and the law controlling the guns were the same in Canada and the United States, we would see this data look very similar to the first bar chart, with Canada having overall less murders, but around the same proportion compared to the United States in each of these three graphs. However, what we see instead is a significant drop in the proportion of homicides in Canada that are being committed with a firearm. This indicates that Canada’s gun control laws and the number of guns that are owned by people in the country has a significant impact on the proportion of homicides that are carried out with firearms. This is then confirmed in the last two bar charts.

In the 4th one, we look at the total number of homicides in Canada in proportion to the total number of firearm homicides in Canada and we can see that only ~30% of homicides are done with firearms. In the United States however, that proportion is significantly different. In the US, the percentage of homicides using firearms is closer to ~68%. That means that if you are murdered in America, you are more than twice as likely to have been killed with a firearm, than if you got murdered in Canada.

There’s a definite decrease in the homicide rates over the course of the years that these number were recorded.

In Science vs., there’s a lot of discussion about how firearms are managed by the governments in different countries. In the US, buying a firearm from a registered gun dealer means filling out a simple form that can be returned within days, where the buyer is background checked by poorly maintained data. On top of that, anyone can go out and legally sell a gun privately. In Canada however, there’s no such thing as private purchases like that. All firearms must be purchased through a registered dealer, all owners must carry a license to own that type of gun, and the pool of types of weapons you can purchase is significantly smaller. In Canada, you aren’t likely to find much that isn’t used for hunting purposes or for sport. From this information, we can draw the conclusion that the regulation of guns in Canada and the types of guns you can purchase, decreases the number of firearm related homicides in the country.

## Summarizing the raw data from Canada and the US

The following summaries represent the homicide data that was recorded in Canada and the United States. The data is broken down into six different summaries, 3 for each country, showing the total homicides, the firearm homicides, and the handgun specific homicides in each country. These summaries show us what the least amount of homicides we saw in a single year was, what the maximum amount we saw was, the average across the years, and the first and third quartiles of the data, so we can see where 68% of the data lies.

These summaries tie in to what Science vs. had to say about the proportion of homicides that were committed using firearms in the US. They said that over 50% of homicides in the US are caused by firearms, which holds up in the data that we see here. According to this data set, firearms accounted for on average 68% of all homicides in the US. This is compared to Canada where firearms only accounted for approximately 31% of all homicides from 2005 to 2014. This is a significant different when Canada is one of the leading countries in the world in terms of number of guns per capita.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Minimum | 1st quartile | Mean | 3rd quartile | Maximum |
| Canadian Handgun Homicides | 2.56 | 2.96 | 3.274 | 3.8 | 4.07 |
| Canadian Firearm Homicides | 3.85 | 4.61 | 5.292 | 5.87 | 6.96 |
| Total Canadian Homicides | 14.54 | 15.74 | 17.262 | 18.4 | 20.62 |
| American Handgun Homicides | 17.44 | 19.77 | 21.563 | 24.53 | 26.24 |
| American Firearm Homicides | 25.47 | 27.76 | 30.024 | 33.59 | 34.35 |
| Total American Homicides | 37.51 | 41.03 | 44.26 | 49.46 | 50.6 |

## Comparing population groups

The following box plot represents the normalized data of the firearm homicides in Canada and in the US. We can see here that even though we’re looking at the normalized data, Canada’s maximum number of firearm homicides in a single year never comes close to hitting the minimum firearm homicides observed in the US in one year. The mean number of firearm homicides in the US is 30/1000000 people, while in Canada it’s only down at 5/1000000 people. If we were to make histograms of these datasets, we would expect to see Canada being unimodal and nearly symmetric – indicating no change, and the US having a slight skew to the right indicating a decreasing amount of homicides.

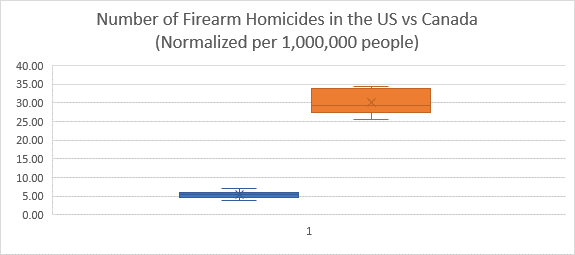


Figure 7: Normalized data comparing firearm homicide number in Canada and the US

In the next boxplot, we observe the mean number of total homicides in Canada and the US. The US averages around 43 homicides per 1000000 people, while Canada is at ~18/1000000. This means that Canada, although it has significantly less firearm homicides than the US, it has a little less than half as many total homicides. This ties in to something that was mentioned in the Science vs. podcast. While we don’t see the crime rates appear in these data sets at all, we can say that increased homicide rate might be because of the fact that a gun is far more likely to be between two people during potentially dangerous encounter, whether it’s a burglary, an assault, or any other type of crime. Having the gun between the two people automatically increases the odds of that crime turning fatal, and in Canada you can’t open carry a weapon, and at home it must be under lock and key. So, while Canada can have plenty of other weapons between an attacker and a victim, a firearm is far less likely, potentially leading to lower homicide rates.

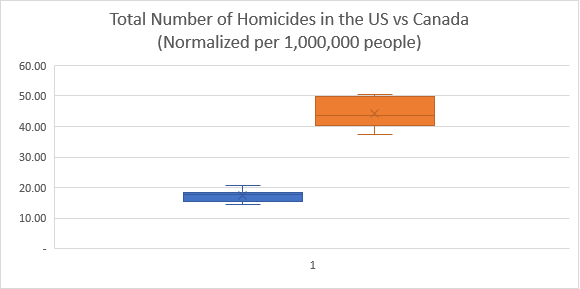


Figure 8: Normalized data comparing total number of homicides in Canada and the US.

# Task 2

Figure 9: Normalized Canadian data for homicides with firearms vs total homicides

The scatterplot above represents the total number of homicides compared to the number of firearm homicides with each year from 2005 to 2014 as a data point. The correlation between the total number of homicides and the number of homicides is quite strong. The correlation is positive and linear, so as the number of total homicides go up, we also see the number of firearms homicides go up with the proportion of total to firearm homicides staying relatively similar. The data does meet the correlation conditions for this scatter plot. The data points are both quantitative, the plot does meet the straight enough, and there are no outliers here. The data point for 2011 is the one data point here that could potentially be an outlier, but I don’t believe that the data is far enough away from the trend line to be able to be considered one. The coefficient of determinants for this scatterplot is .8429, which is pretty high. Having a high R2 value means that there are very few data points in this scatterplot that deviate significantly from the regression line. The correlation coefficient for this scatter plot is .9181, which means that there is a very strong correlation between the number of homicides and the number of firearm homicides every year.

Figure 10: Normalized American data for homicides with firearms vs total homicides

This scatterplot represents the total number of homicides compared to the number of firearm homicides in the US with each year from 2005 to 2014 as a data point. The correlation between the total number of homicides and the number of homicides is very strong. The correlation is positive and linear, so as the number of total homicides go up, we also see the number of firearms homicides go up with the proportion of total to firearm homicides staying relatively similar. The data does meet the correlation conditions for this scatter plot. The data points are both quantitative, the plot does meet the straight enough, and there are no outliers here. The coefficient of determinants for this scatterplot is .9929, which is extremely high. Having such a high R2 value means that almost all the points in this scatterplot fall very close to the regression line, with virtually no deviation from it. The correlation coefficient for this scatter plot is .9964, which means that there is a very strong correlation between the number of homicides and the number of firearm homicides every year.

Both Canada and the US have a strong correlation between the number of firearm homicides and the total number of homicides. However, while Canada’s is pretty high, the correlation for the US homicides is extremely strong. This suggests that the US has a gun problem, seeing as the that correlation is so strong, you would think that if there were regulations imposed against guns, a lot of these homicides wouldn’t happen, thus reducing this very strong correlation between the number of homicides and firearm homicides.

# Task 3

## Confidence intervals

Performing a two-proportion z-interval requires meeting certain conditions and assumptions before hand. The first condition to meet is the randomization condition. While we cannot reasonably prove that either of these two data sets meets the randomization condition, we can trust the organization who collected this data. The Canadian government and FBI are trustable sources of information, and so we can assume that this data has not been skewed in any way.

The second condition to statisfy is the nearly normal condition. Looking at a relative frequency histogram of the data, we can see that the histogram is symmetric, unimodal, and has no outliers or anything else unusual about it. Thus, we can conclude that the Canadian data satisfies the nearly normal condition.

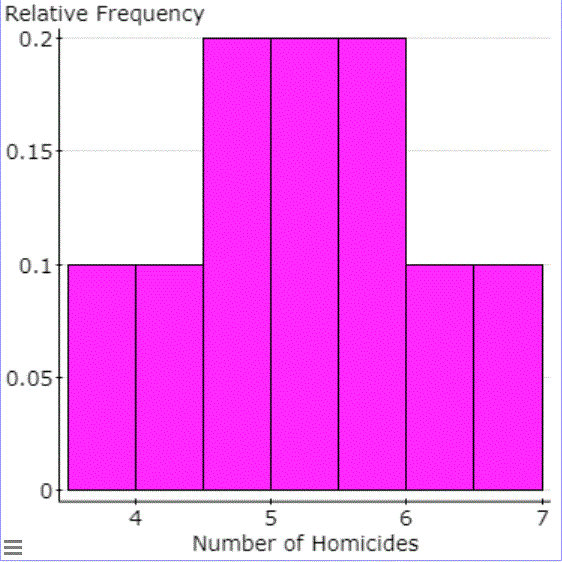


Figure 11: Relative frequency of Canadian firearm homicides

The American data is slightly harder to interpret than the Canadian data since the relative frequency histogram here is symmetric, and it could be argued to be bi-modal. However, because the possible second mode is at the same level as the column sitting directly next to the clear mode, with not that much space between them, we can argue that this is not a true separate mode in the graph. If we applied the central limit theorem to this, we would end up with a histogram that appeared nearly normal.

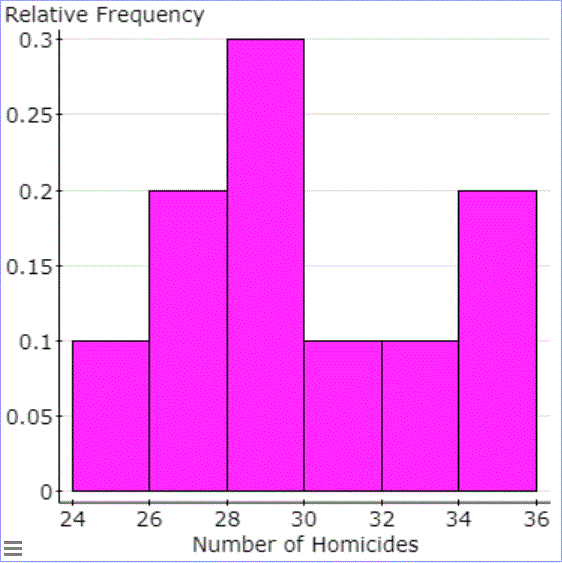


Figure 12: Relative frequency of American firearm homicides

Using the normalized data, and treating all the years of Canadian firearm statistics as a single sample, we can calculate the mean of this data to be . Knowing the mean of this data, we can then calculate the standard error to be .2859. These numbers will create a 95% confidence interval of . This interval means that we are 95% confident that that the true mean of firearm homicides in Canada, per year, lies within the interval presented.

Switching over using the normalized American firearm homicides as our sample data, we find a mean of . Again, we can use the mean of this data to calculate our standard error, this time coming out to 1.0128. These numbers will create a 95% confidence interval of . This interval means that we are 95% confident that that the true mean of firearm homicides in Canada, per year, lies within the interval presented.

## Success proportion values

Performing a two-proportion z-test requires meeting certain assumptions and conditions. The first condition that must be met is the randomization condition. As was stated previously, while we cannot reasonably prove that either of these two data sets meets the randomization condition, we can trust the organization who collected this data. The Canadian government and FBI are trustable sources of information, and so we can assume that this data has not been skewed in any way. The next assumption that must be met is the independence assumption. While a small number of homicides in both groups would likely be related; due to mass shootings, serial killing, etc., there is no reason to believe that a significant amount of the homicides presented had any effect on other homicides that occurred. One condition that must be met is the 10% condition. This condition states that our sample size must be less than 10% of the population that we are trying to represent. In both cases, our sample size is the entirety of what we wish to represent. While the 10% condition is not met, in both situations we know what god knows, and therefore we can proceed. The success failure condition is also met in both cases since the number of success in both cases is well over 10.

|  |  |
| --- | --- |
| H0 : µ1 = µ2 | (Null Hypothesis) |
| HA : µ1 ≠ µ2 | (Alternate Hypothesis) |

Proportion of Canadian Firearm Homicides:

Proportion of American Firearm Homicides:

Once you have the proportion of Canadian and American firearm homicides, you can proceed to calculate the P-Value.

Such a large z-value, one unavailable on a z-score table, will yield a p-value of statistical insignificance; regardless of whether we use a significant level of 10% or 5%, this is too high. This would show that H0 is not true, and therefore we will reject the null hypothesis.