

## COVID-19 Impact: Canada Vs the Philippines

Liu Yufei, Queen Fatema Merhin, Zian Md Afique Amin,  
Muhammad Syazmi bin Suhaidi

Department of Computer Science, Kulliyah of Information and Communication  
Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia.  
[arthas1061476137@gmail.com](mailto:arthas1061476137@gmail.com), [mehrin.62783@gmail.com](mailto:mehrin.62783@gmail.com), [afiquezian@gmail.com](mailto:afiquezian@gmail.com),  
[syazmiadi@gmail.com](mailto:syazmiadi@gmail.com)

**Abstract**— This paper is to analyze the covid-19 situation in Canada and the Philippines with the help of big data analytics and machine learning algorithm which is the K-means algorithm. Age, gender, province, death rate, etc would be the attributes that will help us to analyze the condition of coronavirus in these two countries. This inspection would help us to compare coronavirus and its impact on the countries. Besides, it will also help other countries to determine the faults and outstanding performances done by these countries and they can implement these in their agenda.

**Keyword**— Coronavirus, Covid-19, Canada, Philippines, death rate, gender, age, province.

### I. INTRODUCTION

Coronavirus has affected almost every corner of the world. The most destructive factor since the Second World War has been the global pandemic. There is virtually no single nation or society that, medically, socially, economically, and politically has not felt its effect. The pandemic has driven the world economy into a slump, forcing the economy to continue to shrink and inflation to halt.

The COVID-19 pandemic had a profound effect on the economy of Canada, bringing them to a recession. The social distancing laws of governments have had the effect of restricting the country's economic activity. Companies began planning workforce mass-layoffs, which were effectively avoided by the Canada Emergency Wage Subsidy.

Similarly, since early March, when the first case of local transmission was reported, the Philippines has been battling its coronavirus outbreak. Since then, the announcement of the Red Sub-level 2 Code (March 12), the Enhanced Community Quarantine (ECQ) covering all of Luzon (March 16 to April 30) and the introduction of mass monitoring measures (April 14) have been implemented, among others.

For performing the analysis of the impact of coronavirus on Canada and the Philippines and make a comparison between them, we have chosen a few datasets from

<https://www.kaggle.com/vanshjatana/virus-in-canada/notebook#data> and

<https://www.kaggle.com/cvronao/covid19-philippines-e-dataset>. We will choose some datasets from here

which will include the travel history of people in Canada and the Philippines, recovery, mortality, and testing according to age, gender, and states. These data are enough for us to analyze the covid-19 situation in Canada and the Philippines and make a comparison between them.

### II. BACKGROUND

COVID-19 has been a mostly discussed topic since its arrival. Many countries have shown various strategies throughout the months. Some countries handled it well, however some countries are seeing the most dangerous side of coronavirus.

TOTAL (2021)	CONFIRMED	DEATHS	RECOVERED	IN TREATMENT	FATALITY RATE	RECOVERY RATE
United States	24,483,901 ↗ 170,504	407,216 ↗ 1,866	14,429,571 ↗ 84,064	9,647,114	1.66%	58.93%
India	10,572,672 ↗ 12,136	152,456 ↗ 122	10,211,342 ↗ 13,032	208,874	1.44%	96.58%
Brazil	8,488,099 ↗ 31,394	209,868 ↗ 518	7,411,654 ↗ 22,870	866,577	2.47%	87.32%
Russia	3,591,066 ↗ 22,857	66,037 ↗ 471	2,978,764 ↗ 18,333	546,265	1.84%	82.95%
United Kingdom	3,395,959 ↗ 38,598	89,261 ↗ 671	1,534,736 ↗ 15,630	1,771,962	2.63%	45.19%
France	2,910,989 ↗ 16,642	70,283 ↗ 141	208,071	2,632,635	2.41%	7.15%
Turkey	2,387,101 ↗ 6,436	23,997 ↗ 165	2,262,864 ↗ 8,812	100,240	1.01%	94.80%
Italy	2,381,277 ↗ 12,544	82,177 ↗ 377	1,745,726 ↗ 16,510	553,374	3.45%	73.31%
Spain	2,252,164	53,314	N/A	2,198,850	2.37%	0.00%
Germany	2,051,059 ↗ 9,821	47,463 ↗ 258	1,691,700 ↗ 19,700	311,896	2.31%	82.48%

The picture was retrieved from  
<https://www.outbreak.my/world>. Here top 10

coronavirus affected countries are shown. According to this website Canada is on number 22 and Philippines is 32. Therefore, these countries might have applied some better strategies than those of the top 10. Hence this analysis might help those countries understand the faults as well as learn from their good performances.

We are going to apply big data analytics and multilayer perceptron algorithms in order to answer some research questions that will be helpful for the analysis of our data.

### **III. RESEARCH QUESTION/HYPOTHESIS**

- a. What is the growth rate and death rate of total cases in Canada and the Philippines respectively?
- b. Since Canada and the Philippines have different geographical locations and economical development, Are there different recovery rates and death rates per age group between Canada and the Philippines?
- c. Which region in Canada and the Philippines led to an increasing number of cases?
- d. Based on temperature conditions, to what extent does it affect the COVID-19 transmission rates or growth factors in Canada and the Philippines?
- e. Similarly, both Canada and Philippine have the issue that male is more vulnerable than females. Why do the Philippines have a greater gap than Canada?

### **IV. RESEARCH OBJECTIVES**

- a. To compare the growth rate and death rate in Canada and the Philippines.
- b. To research if COVID-19 affects the developing countries more or developed countries.
- c. To identify the main source of the new cases.
- d. To evaluate if the temperature affects the COVID-19 transmission.
- e. To find out the reason that the Philippines has a greater difference based on gender than Canada.

### **V. RESEARCH SIGNIFICANCE**

This research will help us understand the covid situation in Canada and the Philippines and the countries' strategies to cope with such situations. By performing the analysis, it can also be determined whether Canada is performing well or poorly compared to the Philippines in dealing with the covid-19 situation.

Additionally, other countries can take help from the analysis to cope with the covid-19 situation. Which strategy is more effective will be understood by analyzing the recovery rate of both of these countries. Also, we can analyze the mortality rate with age groups to determine which age group is most likely to get highly affected by corona and which age group is least likely to be affected. Moreover, with the help of gender columns, it can also be determined whether males or females are most likely to be affected by coronavirus.

Furthermore, the travel history of people will help us analyze from which countries it is most likely to import coronavirus cases in Canada and the Philippines. Hence these aspects will help us to get an overall idea of coronavirus in Canada and the Philippines as well as make a comparison between them.

### **VI. DATASET**

<https://www.kaggle.com/vanshjatana/virus-in-canada/notebook#data>  
<https://www.kaggle.com/cvronao/covid19-philippine-dataset>

We have two datasets obtained from the Kaggle website which are Canada's and Philippines' respectively.

In Canada's dataset, we have four files: First, "Mortality\_Canada.csv", recording the death cases with the people's basic information like id, age, gender, region. There are some missing values of id, age, gender and we filled those values with mean and mode separately. Second, "Public\_COVID-19\_Canada.csv", containing more than three thousand records and recording some Canadian citizens' travel history and if they had close contact. Similarly, we fix the missing value issues with mean on numerical data types and mode on categorical data types. What is more, we have "Recovered\_Canada.csv" recording the recovered cases counted by provinces and "Testing\_Canada.csv" which is about test numbers

in every province. As we can see, the age attribute is defined as a categorical data type as a range of age. Besides the case\_id and gender, the province attribute is recorded frequently so that can be a research part. Below is a part of our data frame:

In the Philippines' dataset, the information collected is not totally the same as that in Canada. For example, "Case\_Information.csv" recorded more detailed and extra data such as the exact number of age, region, city, status, and health\_status. Differently, "status" focuses more on the covid-19 itself, and health\_status concentrates more on personal conditions, and this file contains

case_id	provincial_case_id	age	sex	health_region	province	country	date_report	report_week
1.0	1.0	50-59	Male	Toronto	Ontario	Canada	25-01-2020	19-01-2020
2.0	2.0	50-59	Female	Toronto	Ontario	Canada	27-01-2020	26-01-2020
3.0	1.0	40-49	Male	Vancouver Coastal	BC	Canada	28-01-2020	26-01-2020
4.0	3.0	20-29	Female	Middlesex-London	Ontario	Canada	31-01-2020	26-01-2020
5.0	2.0	50-59	Female	Vancouver Coastal	BC	Canada	04-02-2020	02-02-2020

more than twelve thousand rows of records. Otherwise, we have the resource "Testing\_Lab\_Locations.csv" with laboratory accurate locations and testing capacity as well as "Testing\_Aggregates\_20200513.csv" which involves virus testing data in different laboratories from April to May. It had the amount of positive and negative cases with the percentage of total testing cases without missing values. As shown below:

## VII. LITERATURE REVIEW / RELEVANT WORK

Researchers Ashleigh, David, and Amy(2020) discovered that physical-distancing interventions have slowed down the spread level of coronavirus, though different nanopharmaceuticals inventions are being used. Dynamic physical distancing could maintain health-system capacity and also allow periodic psychological and economic respite for populations.

In another research regarding the COVID-19 pandemic in Canada, it clearly stated that Canada faced different epidemiological situations during the covid-19 pandemic. Data were collected from official government documents whenever possible, supplemented by information from international databases and local media reports. The data was then analyzed to identify common patterns as well as significant divergences across the country, especially in the areas of health policy and technology use. The varying degrees of federalism and regional autonomy across the country highlight the different constraints faced by

national policy-makers within different governance models.

### The presence of Covid-19 in high-level

UNIQUE INDIVIDUALS TESTED	POSITIVE INDIVIDUALS	% positive/unique individuals	NEGATIVE INDIVIDUALS	% negative/unique individuals	EQUIVOCAL	% equivocal/unique individuals
15,197	2,652	17%	11,367	75%	32	0.20%
473	21	4%	452	96%	0	0.00%
115	31	27%	84	73%	0	0.00%
672	37	6%	634	94%	1	0.10%
332	43	13%	279	84%	10	3.00%

income countries like Canada is raising important concerns about effective pandemic responses and also the preparation in terms of health. In Canada, some areas have bad settings for services that do not have access to normal and accurate information sources about infirmity. Symptoms and symptoms can not be detected by people living in the regions. Disease etiology details. This may also be a factor in why Covid has 19 instances in Canadians.

The Philippines is the country that recorded the first death outside China. The first case of novel coronavirus (2019-nCoV, now COVID-19) in the Philippines was confirmed on 30 January 2020, in a 38-year old woman who arrived from Wuhan. Two days later, the Philippines recorded the first death outside China on 01 February 2020.

The 2019 coronavirus disease (COVID-19) pandemic presents a challenge to the mental health of populations. This research investigated the prevalence of depressive symptoms in the Philippines and described the variables contributing to the psychological effects. One-fourth of respondents recorded moderate-to-severe anxiety during the early phase of the pandemic in the Philippines, and one-sixth reported moderate-to-severe depression and psychological effects.

Kenneth fitted a hierarchical Bayesian model (to provincial COVID-19 death count time series) that estimated the effects of social distancing measures on COVID-19 transmissibility. There were, however, certain crucial but unknown input parameters, namely, the provincial COVID-19 infection fatality rates (IFR, defined as the conditional probability of dying of COVID-19

given that one is infected with it). Their theoretically straightforward estimates are simply the provincial ratios of the number of COVID-19 deaths to the true number of COVID-19 infections. Unfortunately, the near-complete lack of knowledge of the latter, especially during the early phase of the pandemic, rendered the IFRs highly uncertain.

Currently, in Canada there are more than 555K cases, among them, 457k cases are recovered cases and 14964 are dead. The number of cases is increasing every day. In the Philippines, current cases are more than 472k, among them 439k are recovered and 9162 are dead.

### VIII. METHODOLOGY

This proposal contains various aspects of Coronavirus and its impact on Canada by reviewing the literature and information. The data collection used for this research is a secondary type taken from Kaggle. Theoretical and empirical articles both were taken into consideration for this research purpose. The datasets were discussed among the group members before proceeding with the proposal. The study case emphasizes the graphical technique while the classical technique emphasizes the quantitative technique. First, we will explore the dataset and clean the dataset(if needed), then we will do feature engineering, and finally to analyze our research question we will use various libraries.

Public Covid-19 cases in Canada, totally recovered patients, total testing, and mortality of the corona affected Canada from a certain date will be analyzed for this study case to accomplish the objective of this study case. Sample plots and output will be generated with the Dataplot software program. Line graphs will present the depicted change or trend of variables of covid cases in Canada over a period of time. Certain libraries that will be used for this study case are matplotlib, NumPy, panda, seaborn, and other libraries that will be considered when the study case is conducted.

Multilayer Perceptron Regression will be applied for this study to solve these cases and for prediction. Generally, it is a supervised learning algorithm within the neural network. This kind of regression is a class of feedforward artificial neural networks. An MLP consists of three layers: an input layer, a hidden layer, and an output layer. The

node in layers is a neuron using a non-linear activation function except input nodes. With a hidden layer and activation function, we can have a much more powerful model. Compared to this regression, Linear Regression and the simple neural network can only model linear functions. It has a closed form solution. We can achieve the optimal model directly and efficiently. However, the normal Linear Regression may not be able to solve quite a multi-variable research. Therefore, we plan to utilize MLP as our approach.

### IX. ANALYSIS

#### CANADA:

##### a. Data cleaning:

##### 1. Filling the missing values:

**Mortality before filling out the missing categorical and numerical values:**

```
[ ] Moratlity.isnull().sum()

death_id          0
province_death_id 0
case_id           29
age               21
sex               20
health_region      0
province           0
country            0
date_death_report  0
dtype: int64
```

**Mortality after filling out the missing categorical and numerical values:**

```
[ ] Moratlity.isnull().sum()

death_id          0
province_death_id 0
case_id           0
age               0
sex               0
health_region      0
province           0
country            0
date_death_report  0
dtype: int64
```

**Public before filling out the missing categorical and numerical values:**

```
[ ] public.isnull().sum()
```

```
case_id           500
provincial_case_id 500
age               500
sex               500
health_region     500
province          500
country           500
date_report       500
report_week       500
travel_yn         500
travel_history_country 3547
locally_acquired  3762
dtype: int64
```

**Public after filling out the missing categorical and numerical values:**

```
[ ] public.isnull().sum()
```

```
case_id           0
provincial_case_id 0
age               0
sex               0
health_region     0
province          0
country           0
date_report       0
report_week       0
travel_yn         0
travel_history_country 0
locally_acquired  0
dtype: int64
```

**Recovered before filling out the missing categorical and numerical values:**

```
[ ] recovered.isnull().sum()
```

```
date_recovered     0
province           0
cumulative_recovered 339
province_source     546
dtype: int64
```

**Recovered after filling out the missing categorical and numerical values:**

```
[ ] recovered.isnull().sum()
```

```
date_recovered     0
province           0
cumulative_recovered 0
province_source     0
dtype: int64
```

**Testing before filling out the missing categorical and numerical values:**

```
[ ] testing.isnull().sum()
```

```
date_testing       0
province           0
cumulative_testing 12
province_source     129
dtype: int64
```

**Testing after filling out the missing categorical and numerical values:**

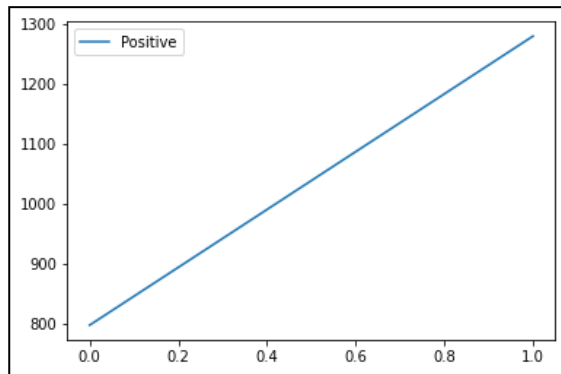
```
[ ] testing.isnull().sum()
```

```
date_testing       0
province           0
cumulative_testing 0
province_source     0
dtype: int64
```

## b. Research Questions

**Question 1:** What are the growth rate and death rates of total cases in Canada?

We generated the growth rate of corona based on the dataset.

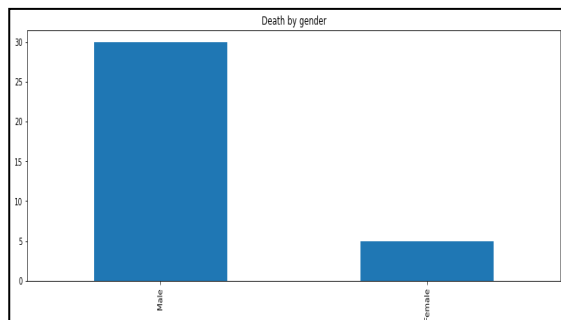


Here we can see that coronavirus had a high growth rate of positive cases in a short amount of time in Canada.

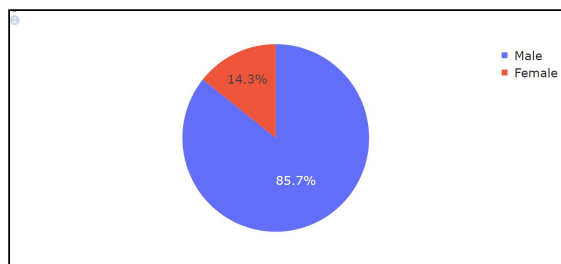
```
[ ] print("Total Cases in Canada : " + str(public.case_id.max()))
    print("Maximum Cases in Province : " + str(public.provincial_case_id.max()))

Total Cases in Canada : 3409.0
Maximum Cases in Province : 1339.0
```

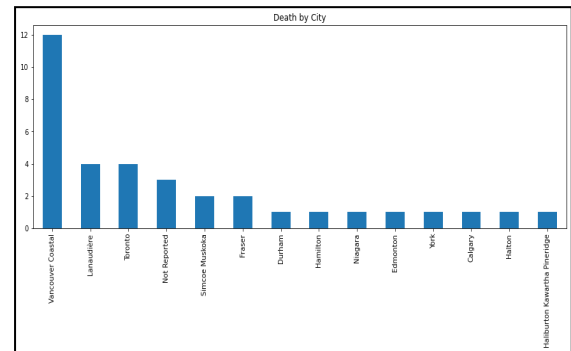
We have further analyzed the dataset for the death rate based on gender, province, and city.



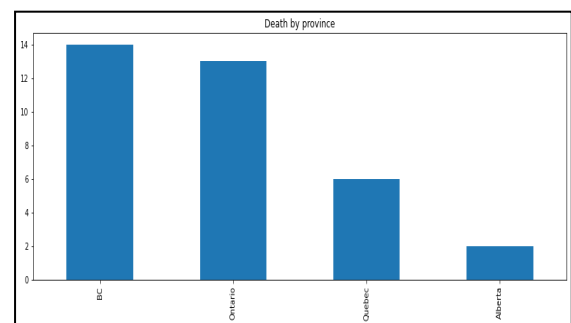
Here it can be observed that males have a much higher death rate than females.



To understand it better, we have generated a pie chart where it can be seen that the male death rate is 85.7% whereas the female death rate is only 14.3%. Hence males are more prone to be affected by the coronavirus in Canada.



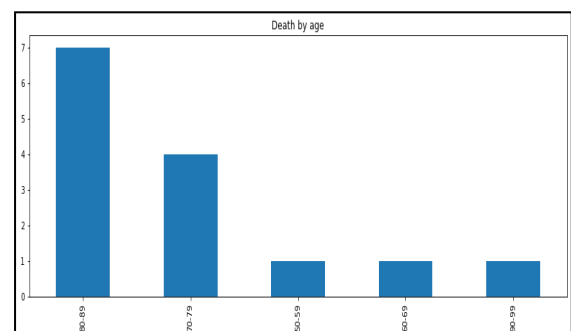
Furthermore, we analyzed the death rate based on the city where it is seen that the Vancouver coastal has the maximum death rate.



And, while analyzing the death rate based on province it was found that British Columbia has the highest death rate followed by Ontario.

**Question 2:** Different recovery rates and death rates per age group in Canada?

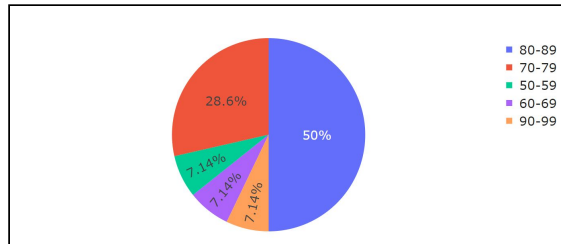
We used a bar plot to analyze the death rate according to age. However, the recovery rate according to age is not mentioned in the dataset.



Here it can be observed that people from the 80-89 age group have the highest death rate followed by the age group from 70-79.

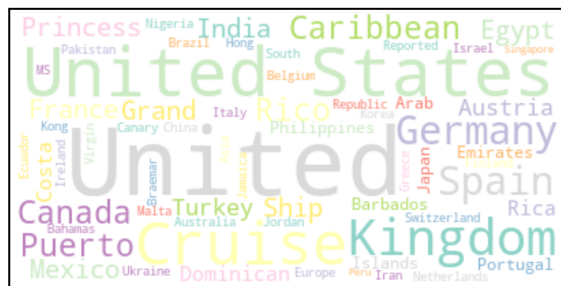


To understand it better, we have generated a pie chart, where 50% of the people from the age group 80-89 have died.

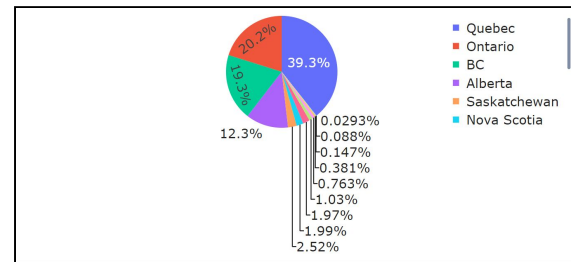
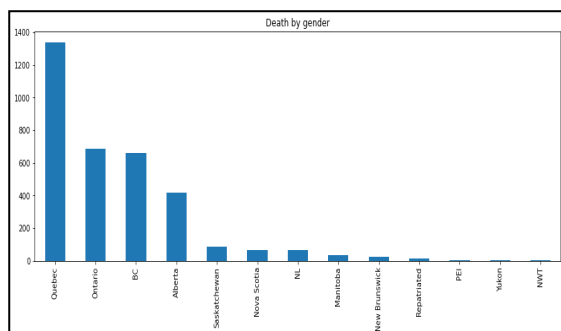


**Question 3:** Which region in Canada led to an increasing number of cases?

Through a bar graph and pie chart, it has been seen that Quebec is the region in Canada that led to an increasing number of cases. We have also analyzed the sources of coronavirus in the regions of Canada.

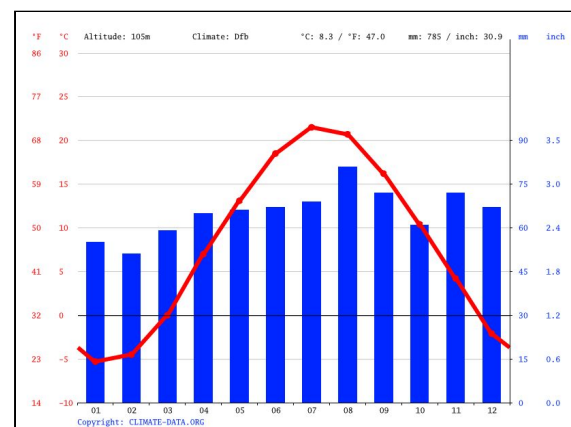


It is seen that the affected People mostly travel to the United States, UK, and Spain. So it could be a cause of the CoronaVirus.

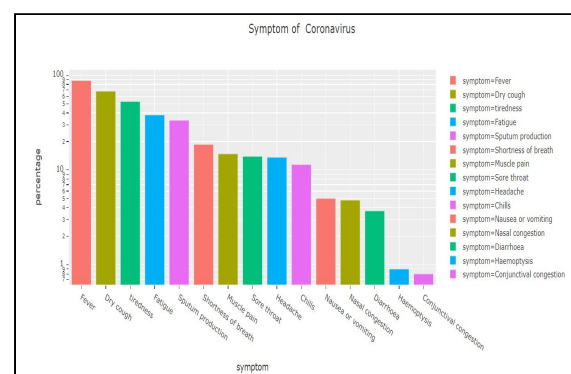


Here it can be seen that 39.3% of cases were reported from the province Quebec followed by 20.2% from Ontario and 19.3% from British Columbia.

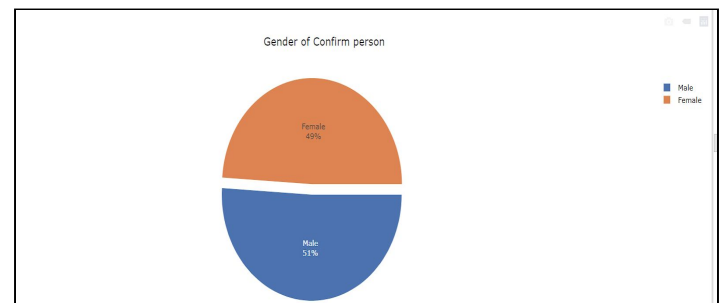
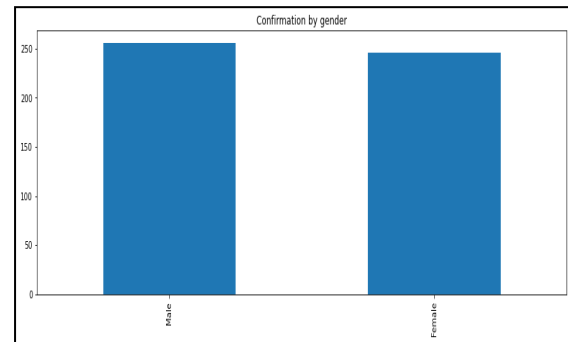
**Question 4:** Based on temperature conditions, to what extent does it affect the COVID-19 transmission rates or growth factors in Canada?



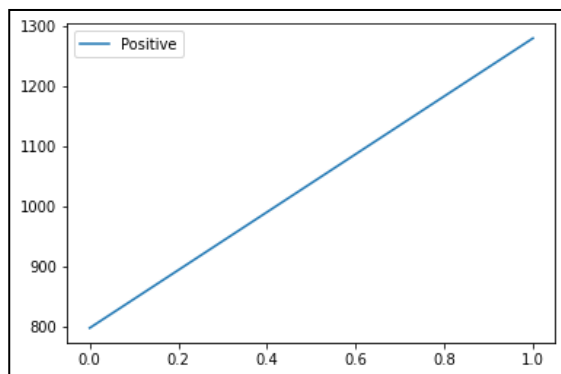
This graph was retrieved from <https://en.climate-data.org/north-america/canada-116/> where it can be seen that Canada is a relatively cold country.



	symptom	percentage
0	Fever	87.9
1	Dry cough	67.7
2	tiredness	52.8
3	Fatigue	38.1
4	Sputum production	33.4
5	Shortness of breath	18.6
6	Muscle pain	14.8
7	Sore throat	13.9
8	Headache	13.6
9	Chills	11.4
10	Nausea or vomiting	5.0
11	Nasal congestion	4.8
12	Diarrhoea	3.7
13	Haemoptysis	0.9
14	Conjunctival congestion	0.8



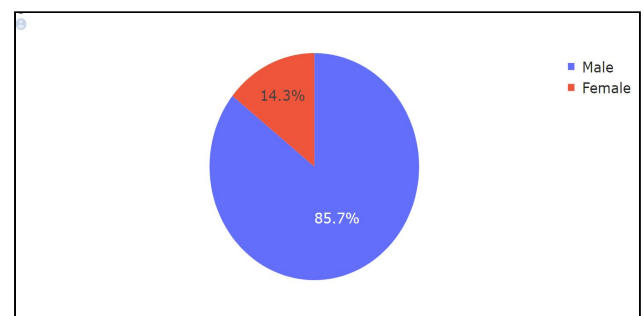
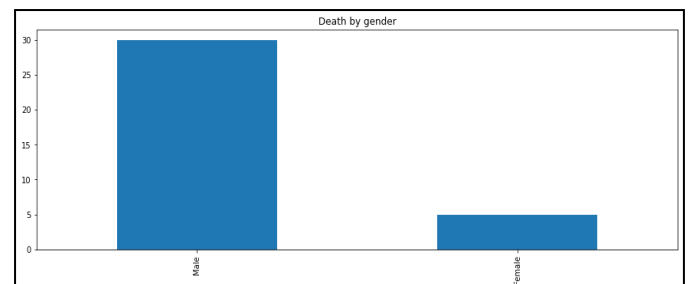
By studying the symptoms of COVID-19 in people in Canada, fever has been reported as the most common symptom followed by a dry cough.



Hence from the figure above, it is seen that coronavirus had a drastic increase in Canada.

**Question 5:** Which gender contributes the most cases in Canada?

From the above graphs, we can see that Male are more prone to get infected by coronavirus than females. Although the difference in case of getting infected is lesser, only 2%, the difference increases to a notable amount when it comes to death by gender.



Here it can be found that 85.7% Male died due to coronavirus whereas Female death rate is only



14.3%. Hence males contribute more to most cases in Canada.

### c. MLP Regression

For multilayer perceptron regression we have used a dataset on the impact of COVID-19 on Canada.

```
[ ] data.head()
```

	Province/State	Country/Region	Lat	Long	Date	Confirmed	Deaths	Recovered	Active	IHO Region
0	NaN	Afghanistan	33.93911	67.709953	2020-01-22	0	0	0	0	Eastern Mediterranean
1	NaN	Albania	41.15330	20.168300	2020-01-22	0	0	0	0	Europe
2	NaN	Algeria	28.03390	1.659600	2020-01-22	0	0	0	0	Africa
3	NaN	Andorra	42.50630	1.521800	2020-01-22	0	0	0	0	Europe
4	NaN	Angola	-11.20270	17.873900	2020-01-22	0	0	0	0	Africa

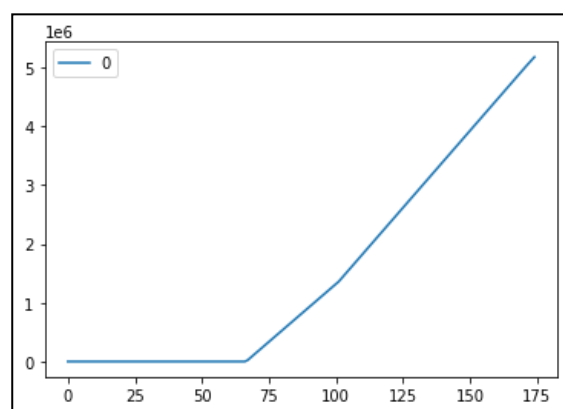
The dataset requires cleaning since it has NaN values for Province/State.

```
data.head()
```

	Date	Confirmed
0	2020-01-22	0
1	2020-01-23	0
2	2020-01-24	0
3	2020-01-25	0
4	2020-01-26	1

Hence for our analysis, we will take only 2 columns, the date, and the confirmed cases.

Hence we got the following prediction where it can be seen that the cases were stable from 0-70(approx) and they had sudden increments gradually.



## PHILIPPINES

### a. Data Cleaning

#### 1. Changing the data types for all the columns into the correct data type

##### Data Type before changes

```
[ ] df.dtypes
```

case_id	object
age	float64
age_group	object
sex	object
date_announced	object
date_recovered	object
date_of_death	object
status	object
date_announced_as_removed	object
province	object
muni_city	object
health_status	object
home_quarantined	object
date_of_onset_of_symptoms	object
pregnant	object
region	object
dtype:	object

##### Data Types after changes

```
case_id
```

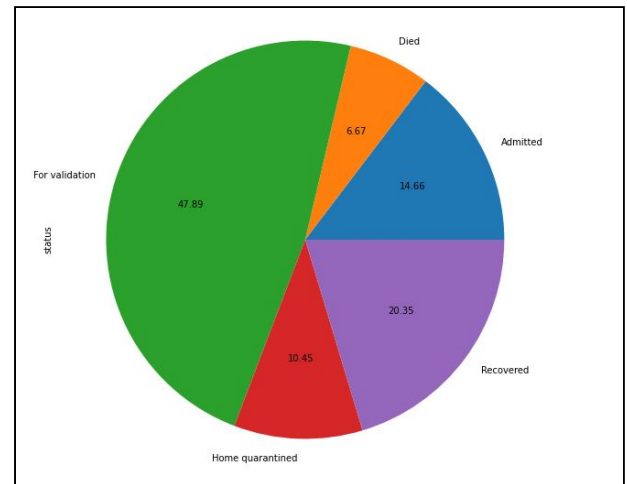
case_id	category
age	float64
age_group	category
sex	category
date_announced	datetime64[ns]
date_recovered	datetime64[ns]
date_of_death	datetime64[ns]
status	category
date_announced_as_removed	datetime64[ns]
province	category
muni_city	category
health_status	category
home_quarantined	category
date_of_onset_of_symptoms	datetime64[ns]
pregnant	category
region	category
dtype:	object

#### 2. Remove any NaN data in the data frame

```
hq= hq.dropna()
hq= hq.reset_index(drop=True)
hq
```

	home_quarantined
0	No
1	No
2	No
3	No
4	No
...	...
5305	Yes
5306	Yes
5307	Yes
5308	Yes
5309	Yes

5310 rows × 1 columns



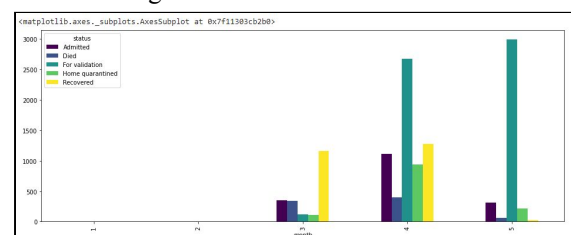
From the dataset of the Philippines, the death rate is recorded with 6.7% Dead Cases. Unfortunately, this dataset cannot be confirmed as the 'For Validation' dataset has the most percentage. Therefore this cannot conclude that the death rate in the Philippines is low.

### 3. Drop the duplicate data in age\_group column

	age	age_group	count
0	38.0	35 to 39	1061
1	44.0	40 to 44	967
2	60.0	60 to 64	921
3	48.0	45 to 49	1002
4	62.0	60 to 64	921
...	...	...	...
12086	35.0	35 to 39	1061
12087	37.0	35 to 39	1061
12088	62.0	60 to 64	921
12089	18.0	15 to 19	282
12090	50.0	50 to 54	973

12091 rows × 3 columns

We have further analyzed the dataset for the death rate according to the month



The bar plot showed that the death rate on Mac to April had a slight increase and in May the death rate dramatically dropped.

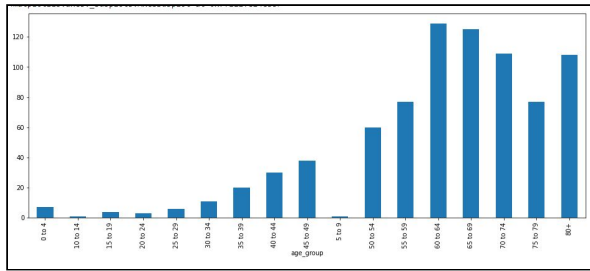
**Question 2:** Different recovery rates and death rates per age group in the Philippines?

We used the unstack function to divide the status\_health into 5 columns. From there we create a new column that is 'Died Percentage', 'Recovered Percentage', and 'Total' for each age\_group.

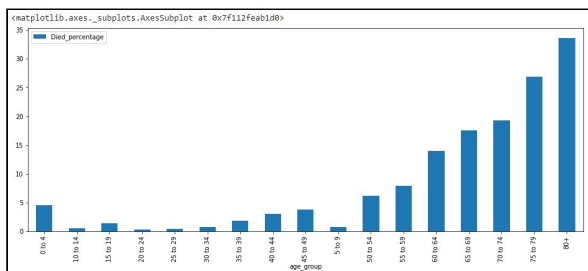
age_group	status	Admitted	Died	for_validation	home_quarantined	Recovered	total	Died_percentage	Recovered_percentage
0 to 4		22	7	100	11	14	154	4.55	9.09
10 to 14		8	1	142	11	7	169	0.59	4.14
15 to 19		25	4	228	15	10	282	1.42	3.55
20 to 24		97	3	451	90	126	767	0.39	16.43
25 to 29		140	6	718	184	241	1289	0.47	18.70
30 to 34		185	11	745	230	339	1510	0.73	22.45
35 to 39		129	20	563	130	219	1061	1.89	20.64
40 to 44		128	30	506	114	189	967	3.10	19.54
45 to 49		151	38	483	103	227	1002	3.79	22.65
5 to 9		5	1	98	6	12	122	0.82	9.84
50 to 54		137	60	441	101	234	973	6.17	24.05
55 to 59		172	77	416	82	220	967	7.96	22.75
60 to 64		168	129	348	68	208	921	14.01	22.58
65 to 69		153	125	193	50	192	713	17.53	26.93
70 to 74		123	109	163	44	126	565	19.29	22.30
75 to 79		55	77	88	9	58	287	26.83	20.21
80+		72	108	90	14	37	321	33.64	11.53

#### b. Research Questions

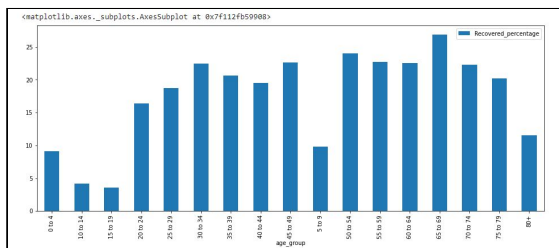
**Question 1:** What are the growth rate and death rates of total cases in the Philippines?



First, we analyze the total death for each of the age groups, and patients from 60 years old to 64 years old recorded the highest with 129 death recorded.



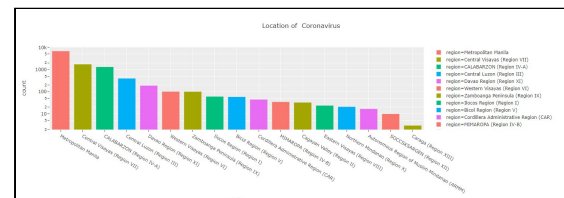
From there we analyzed the Percentage of Death Rate with the total cases in Age\_group. Patients older than 80 years old reported the highest percentage of death rate compared to the other age group.



This Bar Chart represents the Recovered percentage rate for the patient in the Philippines, 65 years old until 69 years old patients recorded the highest rate of recovery with a total of 713 patients

**Question 3:** Which region in the Philippines led to an increasing number of Covid-19 cases?

	region	count
0	Central Visayas (Region VII)	1750
3	Metropolitan Manila	6873
4	CALABARZON (Region IV-A)	1327
12	Central Luzon (Region III)	401
24	Bicol Region (Region V)	59
64	Cordillera Administrative Region (CAR)	45
124	Davao Region (Region XI)	191
151	Autonomous Region of Muslim Mindanao (ARMM)	17
231	MIMAROPA (Region IV-B)	35
245	Western Visayas (Region VI)	102
256	Cagayan Valley (Region II)	33
264	Ilocos Region (Region I)	61
496	Eastern Visayas (Region VIII)	24
548	Northern Mindanao (Region X)	21
556	Zamboanga Peninsula (Region IX)	101
629	SOCSCSARGEN (Region XII)	10
3727	Caraga (Region XIII)	3



We have further analyzed the dataset with the most cases in the Philippines based on Province and Muni City

```
province_max = data4.max()
province_max

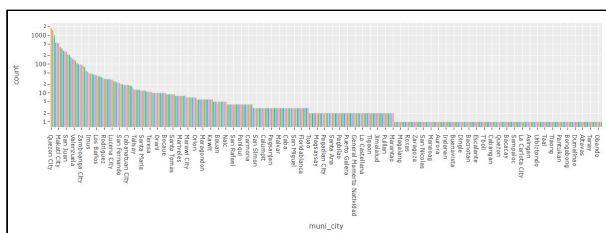
province      Zamboanga del Sur
count                6873
dtype: object
```

```
# Finding the lowest cases
province_min = data4.min()
province_min

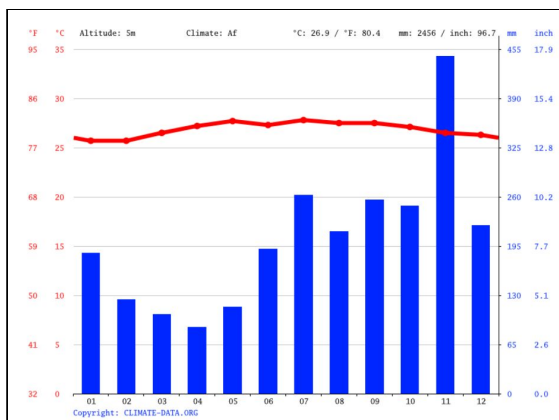
province      Abra
count                1
dtype: object
```

	muni_city	count
0	Dumaguete City	3
2	Panglao	1
3	Taguig	352
4	Cainta	99
6	Makati City	537
...	...	...
10840	Valladolid	1
10906	Bago City	1
10955	Hinigaran	1
10959	Minalabac	1
11040	Hagonoy	1

363 rows x 2 columns

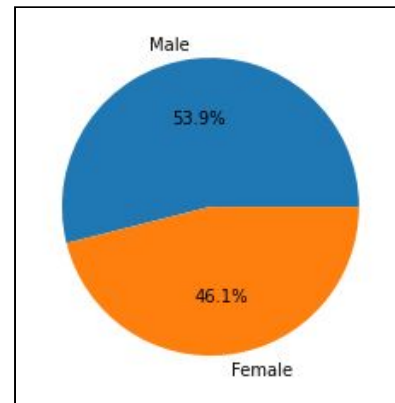


**Question 4:** Based on temperature conditions, to what extent does it affect the COVID-19 transmission rates or growth factors in the Philippines?



This graph was retrieved on the website: <https://en.climate-data.org/asia/philippines/leyte/or-moc-3489/#climate-graph> where it can be seen that the Philippines is a humid country. The average temperature is 26.9 celsius | 80.4 Fahrenheit and the rainfall is around 2456mm | 96.7 inches per year.

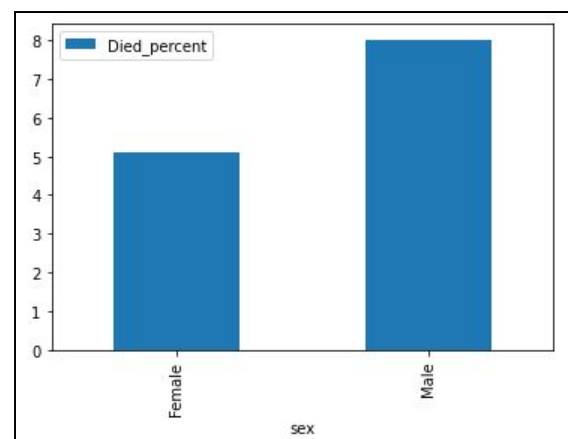
**Question 5:** Which gender contributes the most cases in the Philippines and which gender has the highest rate of death?



This pie chart represents the number of male patients and female patients for the Covid 19 cases in the Philippines. Male patients have the most contiguous gender compared to female patients.

We further our study by analyzing the death rate of the patient according to each gender

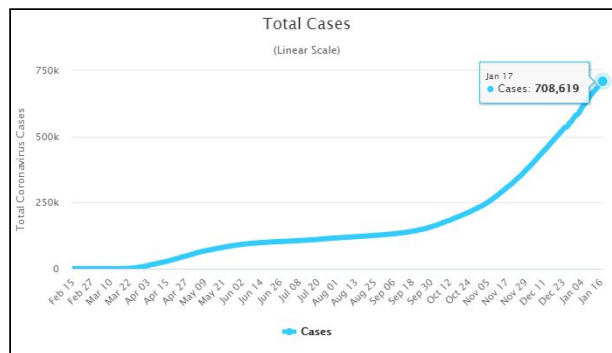
status	Admitted	Died	for_validation	home_quarantined	Recovered	total	Died_percent
sex							
Female	842	283	2645	669	1131	5570	5.080790
Male	930	523	3145	594	1329	6521	8.020242



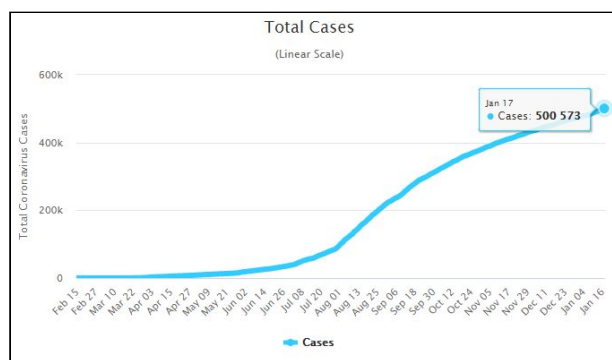
Based on our analysis the bar chart showed that Male patients recorded the highest Death rate compared to Female patients with 8% and 5% respectively. This due to the lifestyle of citizens, it stated that Men in the Asia Pacific smoke five times more than women. Therefore having a healthy lifestyle can reduce the possibility of death by this virus.

## X. DISCUSSIONS

### The comparison of Covid-19 Cases between Canada and the Philippines



Total cases of Covid-19 in Canada on 17 Jan 2021



Total Cases of Covid-19 in the Philippines on 17 Jan 2021

Both of these graphs were taken from <https://www.worldometers.info/coronavirus/> to find the latest update of cases for both countries. Canada recorded the highest number of cases compared to the Philippines. This is because the Philippines government started its partial lockdown on 12 March in Manila to reduce the number of cases in their country. The death rate between both countries showed that Canada also has the highest number of deaths compared with the Philippines. On 17 January 2021, Canada recorded 18,014 deaths while 9,090 deaths in the Philippines.

From this study, we also can conclude that this epidemic does not differentiate between developing countries and developed countries. This is because Canada is a developed country therefore in logical view the country should handle this situation more efficiently compare to the Philippines who is not a developed country but manage to handle this situation effectively

For both countries the temperatures of the country are different. Canada can be considered a cold country while the Philippines is a humid country. From our observation, this Virus does not

discriminate between any country's temperature. This Epidemic affects all over the world and does not affect the transmission.

Finally, both countries had recorded a high number of Male patients compared to Female patients. In general, both genders are equal to contact with Coronavirus, but why both countries have more Male patients, and why more Male patients have died because of this Virus. Based on Biajibswas (2020) stated that women contain more antibodies compared to men therefore it can boost up their immune system. Besides that, female own estrogen as sex hormone this give them more advantage compare to male. Furthermore, individual lifestyles also play an important role in the outcome of this virus. The number of cigarettes that male smoke each day is higher than female. This puts them at a bigger possibility of lung cancer or other cardiovascular disease.

## XI. FUTURE WORKS

This study has some weaknesses and can be improved in the future. Firstly, we can choose other more typical country examples instead of current ones since there are some data frames with different attributes and domains that can not be unified to help us make comparisons between countries. Secondly, there are some other algorithms on regression we could explore and model like Linear Regression and Polynomial Regression. On the other hand, model performance scores can be supplied to show accuracy. To achieve this goal, Mean Squared Error (MSE) can be used to evaluate the performance. The lower the MSE value it is, the better model we have. Otherwise, we could attempt the K-Nearest Neighbors(K-NN) algorithm since it is in another branch of machine learning: supervised learning and it is a nonparametric algorithm.

## REFERENCES

- [1] "Data.org," *Climate*. [Online]. Available: <https://en.climate-data.org/north-america/canada-116/>. [Accessed: 18-Jan-2021].
- [2] Even in the Philippines, Men are More Vulnerable to COVID-19 Than Women . (2021). Retrieved 18 January 2021, from <https://www.esquiremag.ph/life/health-and-fitness/men-vs-women-covid-a00304-20200324>

[3] COVID-19 Outbreak Live Updates, 2019.

[online] Available at:

<<https://www.outbreak.my/world>> [Accessed 18 January 2021].

[4] Biajibswas. (2020). Are Men More Vulnerable to Covid-19 as Compared to Women ? Biomed J Sci & Tech Res, 27(2), 20645–20646.

<https://doi.org/10.26717/BJSTR.2020.27.004481>