Coronavirus analysis

Impact of the COVID-19 pandemic on Canadian economy

-604 term project report
-Adaeze Nwigwe, Bowen Li, Gerry Lin, Jialing Cai, John Campbell

1.0 Introduction

2020 has become an extremely hard year for people globally with the impact of the coronavirus disease 2019 (COVID-19). The COVID-19 pandemic is currently a global outbreak for an untreatable strain of coronavirus that was first identified in late 2019. The virus spread to more than 20 countries by 30 January 2020. To minimize human-to-human transmission, the most commonly used tools are isolation and quarantine (A Wilder-smith, 2020). Numerous people's lives have been significantly influenced and changed. People do their work and study virtually from home rather than physically going to workplaces and schools. People are also stopping traveling for their vacation due to safety reasons.

446,361 COVID-19 confirmed cases and 13,222 deaths have been recorded in Canada to date. There are currently 31,725 active cases in Canada. (Covid-19 , 2020). The pandemic spread across Canada rapidly and the Canadian government chose to take action with border closures and requiring self-quarantine. Although these methods are very effective to stop the spread of the virus, there are some consequences we need to face. The unemployment rate increased to 13.7% in May 2020 (Statistic Canada, 2020). Many people from different working sectors have been laid off due to the restriction and economic conditions associated with COVID-19 (Statistic Canada, 2020). Canada is facing a massive challenge at this moment.

Our aim for this project is to analyze and discover the impact of COVID-19 through Canada. First, we want to analyze data on cases and deaths to see how many Canadian are directly impacted by this pandemic. The second thing we are interested in is how much indirect impact it brings to Canada. How have working hours been reduced due to the impact of COVID-19? How bad will the effect be for different industries? We want to analyze through different working sectors' sales volume to see if every industry will experience an economic downturn through 2020. Retail sales are an important component to measure Canada's production and it will also influence the Banks' interest rates and Business investment strategies (Statistic Canada,2020).

One key aspect to evaluate would be province policies and COVID-19 cases versus economic metrics. It could be assumed that provinces with more relaxed pandemic policies may have higher economic recovery but at the expense of higher COVID-19 cases. This thesis may change month to month as provinces with strict pandemic policies and low COVID-19 case counts may in fact have a better economy as life is more stabilized and back to normal. Moreover, under the impact of COVID-19, numerous airlines canceled or reduced the number of their flights. For example, only in September 2020, Air Canada and WestJet have already canceled about 439 flights(CBC,2020). In June 2020, it has been reported that restaurants, bars are operating at almost 40% sales loss (Global News, Jun 2020), but there was some good news in Sep 2020 as the restaurant, bar sales kept rising in July for the third straight month (OttawaMatters, 2020). We want to analyze what will be the change based on this factor.

2.0 Individual Datasets

Dataset # 1: Data on COVID-19 in Canada [2] -- John

Data on COVID-19 is a compilation of COVID-19 data (confirmed, recovered, deaths and active cases) that is provided by the Public Health Infobase – a department of the Health Promotion and Chronic Disease Prevention Branch of the Public Health Agency of Canada (PHAC). The data that is compiled by the Public Health Infobase – is data as provided by individual Canadian province and territory websites.

This dataset was chosen as it is essentially the only source of detailed COVID-19 data for Canada – provided by the Government of Canada. It includes detailed data for each province and territory. The data was well prepared and did not have to be significantly cleaned however it is important to note that both provincial, territory and total Canada data is included. The user would need to understand this data structure before simply aggregating the data to evaluate total COVID-19 trends. To prepare the data to be joined in the group data exploration - a separate table/view was created that converted the data into monthly data.

Figure 2.1 illustrates key data from the COVID-19 Canada dataset: active cases by province from February to November 2020. This dataset highlights that active COVID-19 cases since March 2020 have been nearly exclusive to 6 provinces: BC/AB/SK/MB/ON & QB. Quebec (QB) was purposely greyed out compared to the other provinces as Quebec changed their reporting methodology on recovered and active cases in Wave 1. This complicated the comparison of Wave 1 to Wave 2. The figure also demonstrates how Canada has experienced 2 waves of COVID-19. Wave 2 currently has 5 times as many active cases in key provinces (BC/AB/SK/MB/ON) as compared to wave 1. In contrast to these provinces, the Atlantic provinces and territories have been able to keep active case counts very low. The Atlantic provinces' low case count can be primarily attributed to the 'Atlantic Bubble' where both domestic and international travelers are prohibited from entering the province with few exceptions.

As wave 2 has become an alarmingly increasing trend of COVID-19 transmission across Canada – it is evident that data and learnings from Wave 1 were not studied to address the risks of Wave 2 that is currently gripping Canada at the end of 2020. Canada was given an opportunity in the summer months as case counts mitigated to allow for better pro-active policies to prevent rising COVID-19 cases. The opposite has taken shape as wave 2 cases exponentially rise across Canada.

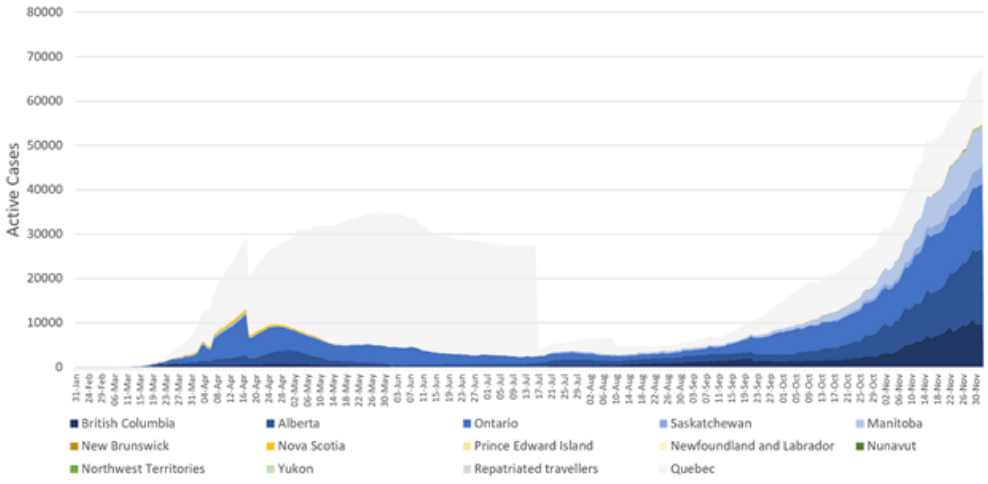


Fig 2.1: Data on COVID-19 in Canada – Active Cases by Province [2]

Dataset #2: Daily number of reported cases/deaths of COVID-19 by country worldwide [8] -- John

The COVID-19 Global data is compiled by the European Centre for Disease Prevention and Control (ECDC). This data is compiled by the ECDC's Epidemic Intelligence team and is based on reports from health authorities worldwide. Data is updated daily from a team of epidemiologists that screens up to 500 relevant sources to collect the latest figures from every country that reports COVID-19 data.

This dataset was chosen as the ECDC is a reputable center that was providing daily updates on global COVID-19 data – confirmed cases and deaths by country. The only data cleaning and preparation that was required on this dataset was to create a separate table/view that converted the data into monthly data. To confirm the accuracy of the global data - this dataset was to compare it to dataset #1 (Data on COVID-19 in Canada). The ECDC dataset was identical and built confidence in the compilation of this dataset.

The most alarming discovery from the individual queries on the global COVID-19 dataset was that global COVID-19 confirmed cases per month were becoming significantly larger in many countries. Figure X – inset chart below shows reported cases per day globally from January to November 2020. Global COVID-19 cases can be defined into 3 waves where wave 1: Mar-May 2020, wave 2: June-Aug 2020, wave 3: Sept-Nov 2020. Global COVID-19 cases in wave 2 were larger than wave 1 however how this value compared to wave 1 varied by country. Wave 3 COVID-19 cases are significantly larger for all countries where global COVID-19 cases appear to be accelerating exponentially.

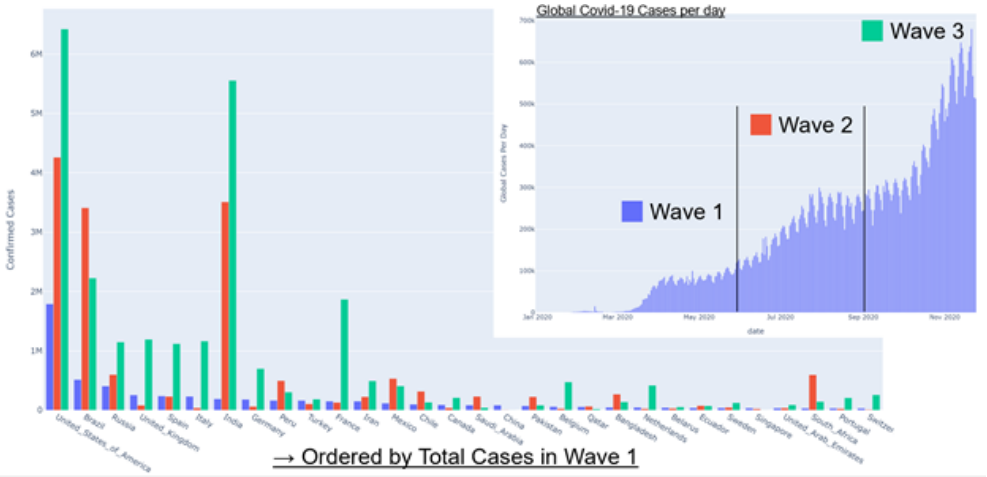


Fig 2.2: COVID-19 Global Total Cases by Wave. Inset: COVID-19 Global Cases per Day – ECDC [8]

Figure 2.2 demonstrates how case counts in each subsequent wave are higher for many countries. Wave 3 is especially concerning where the top 10 countries that had the highest cases in wave 3 were: USA (6,423,709), India (5,556,595), Brazil (2,225,297), France (1,866,717), UK (1,193,028), Italy (1,163,577), Russia (1,148,502), Spain (1,119,758), Argentina (973,122), Poland (809,463). Cases per capita for the top 10 range between 4,100 to 28,000 cases per million.

Canada in comparison has had 209,615 confirmed cases and 5,600 cases per million in wave 3. These statistics rank Canada poorly in terms of overall COVID-19 cases observed in wave 3.

Dataset #3: Retail Trade Sales by Province and Territory [5] – Jialing

The data set I choose contains the information for retail trade sales by province and territory From Statistic Canada. The data set from Statistic Canada under Open Government License Canada. It is an open resource for people to use.

REF_DATE	GEO	DGUID	North American Industry Classification System (NAICS)	Adjustments	UOM	UOM_ID	SCALAR_FACTOR	SCALAR_ID	VECTOR	COORDINATE	VALUE	STATUS	SYMBOL	TERMINATED	DECIMALS
1991-01	Canada	2016A000011124	Retail trade [44-45]	Unadjusted	Dollars	81	thousands	3	v52367096	1.1.1	12588862				0
1991-01	Canada	2016A000011124	Retail trade [44-45]	Seasonally adjusted	Dollars	81	thousands	3	v52367097	1.1.2	15026890				0
1991-01	Canada	2016A000011124	Motor vehicle and parts dealers [441]	Unadjusted	Dollars	81	thousands	3	v52367118	1.2.1	2484133				0
1991-01	Canada	2016A000011124	Motor vehicle and parts dealers [441]	Seasonally adjusted	Dollars	81	thousands	3	v52367119	1.2.2	3124582				0
1991-01	Canada	2016A000011124	New car dealers [44111]	Unadjusted	Dollars	81	thousands	3	v52367142	1.4.1	2085763				0
1991-01	Canada	2016A000011124	New car dealers [44111]	Seasonally adjusted	Dollars	81	thousands	3	v52367143	1.4.2	2563093				0
1991-01	Canada	2016A000011124	Furniture and home furnishings stores [442]	Unadjusted	Dollars	81	thousands	3	v52367150	1.8.1	379906				0
1991-01	Canada	2016A000011124	Furniture and home furnishings stores [442]	Seasonally adjusted	Dollars	81	thousands	3	v52367151	1.8.2	468950				0

Fig 2.3: Retail Trade Sales by Province and Territory - Statistics Canada [5]

The data set includes the monthly retail sales performance from Jan 1991 to Aug 2020. The values are separated by different provinces, the territory with respect to different industry sectors. The data shows the overall performance at the Canada level as well as the province level. The product ID includes retail trade, automobile dealers, clothing stores, etc. The feature I am interested in is the performance of different sectors. I want to see if every sector's economic performance will decrease or there are exceptions. The column I will use later in the project will be REF_DATE, GEO, North American Industry Classification System (NAICS), adjustment and VALUE.

The data size is 20MB with 121,375 rows and 17 columns in CSV format. This is the only file I will use for this project. Since the domain of my data set is to analyze the economic impact of COVID-19. I will mainly focus on data from 2020 and 2019 by doing the comparison. Also, I will filter out the less useful columns like UOM_ID, SALAR_ID, vector, coordinate and decimals.

From the individual query, there are a couple of things that need to be aware of : March, April and May are the three months that Canada has lower retail trade performance. We can see a significant volume drop during April. The retail sale trade dropped from 47,028,013 thousand dollars to 35,352,146 thousand dollars from March 2020 to April 2020. The trend for different provinces is about the same. For Canada, the retail trade industry with maximum volume is Food and beverage stores. The total income during 2020 is 188,797,731 thousand dollars. On the other hand, in the provinces like Alberta and Newfoundland, Labrador and Yukon, etc. the maximum value occurs at motor vehicle and parts dealers. The difference may be caused by different lockdown policies.

geo	max_volume	indutries
Canada	188797731	Food and beverage stores [445]
MontrÃ©al, Quebec	11277334	Food and beverage stores [445]
Quebec	21468815	Food and beverage stores [445]
Nova Scotia	2550638	Food and beverage stores [445]
Vancouver, British Columbia	6289878	Food and beverage stores [445]
British Columbia	14874628	Food and beverage stores [445]
Northwest Territories	236794	Food and beverage stores [445]
Gatineau, Quebec	967706	Food and beverage stores [445]
Ottawa, Ontario	2509752	Food and beverage stores [445]
Alberta	13668661	Motor vehicle and parts dealers [441]
Newfoundland and Labrador	1556229	Motor vehicle and parts dealers [441]
Saskatchewan	3141103	Motor vehicle and parts dealers [441]
Edmonton, Alberta	5471558	Motor vehicle and parts dealers [441]
Ontario	35052434	Motor vehicle and parts dealers [441]
Yukon	105905	Motor vehicle and parts dealers [441]
New Brunswick	2160055	Motor vehicle and parts dealers [441]
Quebec, Quebec	1988466	Motor vehicle and parts dealers [441]
Winnipeg, Manitoba	2048278	Motor vehicle and parts dealers [441]
Calgary, Alberta	4250908	Motor vehicle and parts dealers [441]
Manitoba	3273958	Motor vehicle and parts dealers [441]
Prince Edward Island	401027	Motor vehicle and parts dealers [441]
Toronto, Ontario	15440303	Motor vehicle and parts dealers [441]
Nunavut	286525	Supermarkets and other grocery (e…

Figure 2.4: Provinces with the max volume retail sector

One noteworthy challenge from my dataset is how to filter out the useless information, get the valuable information I need. There is a lot of information included in the table with nested values. Otherwise, the information provided by Statistic Canada is very clean. There is no column with a null value. The data are organized well and easy to understand.

Dataset # 4: Monthly Survey of Food Services and Drinking Places [9] -- Gerry

The data set Monthly survey of food services and drinking places comes from Statistics Canada. The data is collected by Monthly Survey of Food Services and Drinking Places program.

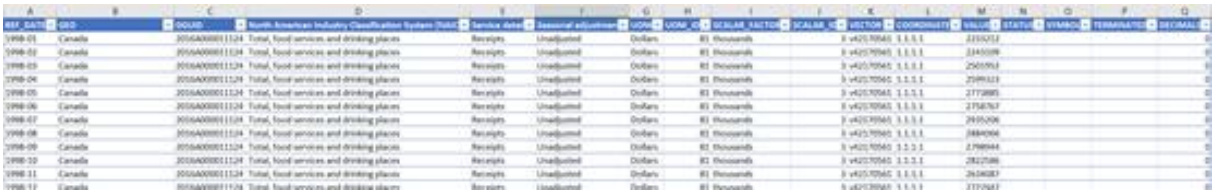


Fig 2.5: Monthly Survey of Food Services and Drinking Places [9]

It is a monthly basis data set from 1998 to 2020; it covers all Canadian provinces and territories. It includes four different service categories: Drinking places (alcoholic beverages) [7224], Full-service restaurants [722511], Limited-service eating places [722512], Special food services [7223] and it also comes with a total sum row of all these four categories. All categories have receipts type service detail data which is service dollar amount. Some categories have location type service detail data which is service counts. We want to compare the data category over category, month over month and year over year. We also want to know if there is a difference between provinces and territories. Most crucial, we want to analyze how COVID-19 affected food and drink service by linking the above data and analysis results with COVID-19 cases and trend data set. The size of the data file is 5.73 MB with CSV format. It has17 columns and 34012 rows. Its data type includes date, varchar and number.

The data is licensed under the Statistics Canada Open License.

Dataset # 5: Data on working hour in Canada by Industry [10] -- Bowen

This dataset is about working hours at the main job and North American Industry Classification System (NAICS), from January 1976 to September 2020 for all the provinces in Canada. Data is also available for sexes and classes of workers. The data is open to public license by Open Government License – Canada. It is available at https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410003601

Working hours indicates the employment situation which also implies the growth of the economy for the country. In this dataset, there are a lot of nested relationships in the columns. For example, it has data for males, females and both sexes. Before exploring data, data cleaning is conducted in the Jupyter notebook. The data is taken after 2000-01 for total actual hours and total employed. Also, DGUID, UOM_ID, SCALAR_FACTOR, ACALAR_ID, VECTOR was dropped since they are not going to be used in the following data exploration.

From the individual queries, we find out that in the past 20 years, total working hours for all the industries increase except for the year 2020. In 2020, wholesales and retail trades, health care and social assistance, manufacturing, professional, scientific and construction are the top industries with working hours. Compared with 2019, even monthly total working hours in health care and social assistance dropped. For provincial data, Alberta has the most working hours in Forestry, fishing, mining, quarrying, oil and gas

while Ontario has the most working hours in the rest of the industries. The working hours of males are more than females in all the industries except for health care and social assistance and educational services.

Dataset #6: Non-resident travelers entering Canada, by country of residence, seasonally adjusted [11] -- Adaze

The data set was obtained from the Statistics Canada website and is licensed under the Statistics Canada Open License. The data source of the Statistics Canada data is the Frontier Counts data which provide statistics on the number of international travelers by selected category and by type of transportation to many different organizations such as the World Tourism Organization and the Tourism Satellite Account. This helps us trust the integrity of the data since many organizations rely on this data source.

The data set contains monthly information on the number of non-resident travelers entering Canada by country of residence, excluding the United States. Data is seasonally adjusted to measure and remove the influences of predictable seasonal patterns. Dataset is presented in a tabular format and consists of data from January 1990-August 2020. The CSV file is 163kb.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	REF_DA	GEO	DGUID	Country	UOM	UOM_I	SCALAR	SCALAR	VECTOR	COORD	VALUE	STATUS	SYMBO	TERMIN	DECIMA	
2	Jan-72	Canada	2016A000(	Total non-	Persons	249	units	0	v126275	1.1	18907				0	
3	Jan-72	Canada	2016A000(	North Ame	Persons	249	units	0	v126443	1.168	2047				0	
4	Jan-72	Canada	2016A000(	Jamaica	Persons	249	units	0	v126444	1.169	420				0	
5	Jan-72	Canada	2016A000(	Mexico	Persons	249	units	0	v126445	1.17	388				0	
6	Jan-72	Canada	2016A000(	Trinidad a	Persons	249	units	0	v126446	1.171	265				0	
7	Jan-72	Canada	2016A000(	South Ame	Persons	249	units	0	v126482	1.207	907				0	
8	Jan-72	Canada	2016A000(	Brazil	Persons	249	units	0	v126483	1.208	173				0	
9	Jan-72	Canada	2016A000(	Europe, to	Persons	249	units	0	v126276	1.2	11902				0	
10	Jan-72	Canada	2016A000(	Belgium	Persons	249	units	0	v126277	1.3	185				0	
11	Jan-72	Canada	2016A000(	Denmark	Persons	249	units	0	v126329	1.4	153				0	
12	Jan-72	Canada	2016A000(	France	Persons	249	units	0	v126330	1.5	1725				0	

Fig 2.6. Non-Resident Travellers entering Canada, by country of residence, seasonally adjusted [11]

The data set was well organized and even though it contained many years of data it was very complete. When downloaded into CSV format, the dataset was also accompanied by a metadata file with some definitions around the dataset. The metadata file was not very complete and did not provide clear definitions for all the fields in the dataset making some of the columns of the dataset ambiguous. The most relevant columns that we needed to use for data analysis and exploration purposes were clearly defined so we were still able to continue with the data analysis. The data was of good integrity and the format was easy to work with so not much clean-up or manipulation was needed.

To analyze travel trends over time we calculated the percent difference in travel from each country into Canada between 2019 and 2020. Travel trends show a steady decrease in travelers from all countries to Canada between 2019 and 2020 starting in March for most countries.

In March and April travel trends show a percent decrease in travelers between 2019 and 2020 of up to 99%.

Most countries in Europe, North America, South America and Africa show a slight percentage increase in travelers between 2019 and 2020 between January and March, but then a sharp decrease is noticed in the months after that.

Eastern Asian countries on the other hand show a small percentage decrease in travel trends between 2019 and 2020 starting from early February and the percent decrease reaches number as high and 99% in the subsequent months of the year.

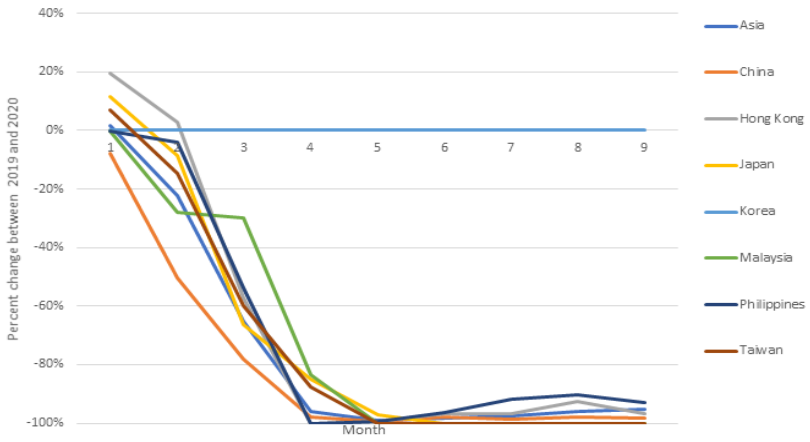


Fig 2.7: Non-Resident Travellers from country of origin - Percent change between 2019 and 2020

3.0 Proposed methodology

This project's primary method is the relational database model, SQL and Python. We will save all data into a central MySQL database. Once completed, we will analyze the data with SQL and Python. We complete the whole process with the following stages:

1. Extract the above data sets from different data sources with proper format. We want to use CSV format files here mainly. We will use Python and Pandas to do the first-round data clean if it is necessary. For example, remove unnecessary header or tail, convert the file to a proper encoding character.

2. Load each data set into a dedicated MySQL database table with the MySQL *LOAD DATA LOCAL INFILE* command or Python.

3. Use SQL/Python moves setup 2 data into another set table. We will do data clean during the process. It includes:

    a. Drop unnecessary fields

    b. Convert data to proper data type, like string to date, decimal to integer.

    c. Delete duplicate records

    d. Unify table column names

    e. Create aggregation tables or views. For example, we may have data sets with daily data. We would need to make a separate table/view to store the monthly data

    f. Find the common key(s)/column(s) between different data tables. We may add a new column and combined values of other columns to a table if necessary.

    g. Create necessary indexes

4. Create additional views or tables to store the combined/join result between tables

5. Use SQL and Python to analyze the data and use Python or Excel to create data visualization.

    a. We will compare 2020 travel trends to travel trends from the past 10 years to see how the trends have changed in the past 10 years and to see how different the trends have been since the Covid-19 pandemic. We will also explore how travel trends vary between countries and observe any possible correlation between COVID-19 cases in various countries and their travel patterns to Canada

    b. We will compare working hours in 2020 to working hours from the past 10 years. Geography and industries are also being considered in order to get the impact of COVID-19 pandemics. Which province or industry has the largest working hour cut? Which industry has the largest working hour increase? Does the working hour have an impact on air qualities?

    c. How were the retail sales affected by COVID-19? Will every sales industry have the same economic performance during the outbreak of the pandemic? Which province or industry will be affected the most? Will the most affected province have related to the province which has the most death cases?
    d. The impact of foreign travel into Canada and its provinces from countries with higher COVID-19 cases will be evaluated to determine their impact on active cases and deaths in each province.
    e. Did Canada shut-down its borders to other countries as communicated? How many travelers and from which countries came into Canada in the months leading up to the first COVID-19 case in Canada? Do borders continue to be locked down?
    f. How do working hours, retail sales and food service data compare in provinces with high vs low active COVID-19 cases? How has this data changed throughout the COVID-19 pandemic as each province experienced wave 1 in March 2020 and the current second wave being experienced since September 2020.

4.0 Data Exploration

4.1 Data Cleaning

-John

Since the COVID-19 data is daily data, the reset is monthly, to avoid writing subquery and aggregating data whenever we need to join COVID-19 data with another dataset. We create a COVID-19 monthly view which includes monthly COVID-19 data: TOTAL CASES, DEATHS, ACTIVE CASES

-Jialing Cai
My original dataset is downloaded from Statistic Canada with file number 20100008. The original data contains the monthly retail performance information from 1991 January to 2020 August. In this project, I will use the value between 2019 and 2020. While transforming the data into MySQL, I changed the column name from capitalized to lowercase, using underscores to separate the column name. Simplified the column name, make it more readable and easier to understand. Moreover, I only choose the meaningful columns with their corresponding value to use in my project. In this case, the value I use includes the reference date, NAICS, geo, volume, and adjustment type.

-Bowen Li
To reduce files' size, some columns such as DGUID, UOM_ID, SCALAR_FACTOR, ACALAR_ID, VECTOR are dropped since they are not going to be used during data exploration. There are some nested

relationships on the working hours data, total working hours and total employed are kept. Those operation has been done in the individual queries.

--Zhan (Gerry) Lin

All data which is older than 2010 have been deleted. We exclude adjusted data and counting from the query. Indexes have been added on date and province columns.

--Adaeze Nwigwe

The columns of the data set that were relevant for the purpose of our data exploration were: ref_date (monthly date), country (country from which travellers are coming from), geo (country travellers are entering into) and value (number of travellers). The only columns that needed some clean up were: country and geo. Some update queries were used to clean up and standardize the names in the country field to allow for grouping and analysis of the data. The geo column contained some incomplete data and while it was supposed to provide details around the province that travellers were entering, the only provinces reported were BC, Ontario and Quebec and even those provinces were not reported consistently. We decided to exclude province specific numbers from the analysis and only include data for overall travel into Canada. We also included data from the past 10 years in the analysis

4.2 Group query summaries

Group Query (John Campbell)

The objective was to review the impact of COVID-19 on overall retail sales (unadjusted and seasonally adjusted), food sales (unadjusted and seasonally adjusted) and working hours specific to those industries.

This query intends to combine high level - specific columns from 4 key data tables outlined above to illustrate how these datasets trend with one another. These tables included: retail_sales_data , food_drinking_service, working_hour_data covid_canada. The query is a set of nested subqueries where food and retail sales were joined – then joined with working hours – then joined with COVID-19 – Canada data. The objective of the query was to join the tables based on dates – using inner joins to maintain as much data from each table as possible. The final result is a combined table shown in Table 4.2.1.

Table 4.2.1: Group Query - Retail, Food Sales (unadjusted and adjusted) , Hours and COVID-19 data

date	year	month	geo	Retail_Hours	RETAIL - Unadjusted Sales ($M)	RETAIL - Seasonally Adjusted Sales ($M)	Food_Hours	FOOD - Unadjusted Sales ($M)	FOOD - Seasonally Adjusted Sales ($M)	COVID Cases	COVID Deaths	COVID Active Cases
2020-01-00 00:00:00	2020	1	Canada	86056	43646873	51891454	30281	5702234	6468869	4	0	4
2020-02-00 00:00:00	2020	2	Canada	87219	42981961	52262166	31179	5682623	6482288	11	0	11
2020-03-00 00:00:00	2020	3	Canada	75015	43641006	47028013	17327	4002082	4152415	8533	95	1700
2020-04-00 00:00:00	2020	4	Canada	61025	34869392	35352146	10905	2442264	2508885	44688	3088	22020
2020-05-00 00:00:00	2020	5	Canada	70531	47342177	42852368	13077	3448650	3252320	37711	4111	32632
2020-06-00 00:00:00	2020	6	Canada	80235	56342104	52488186	20216	4416741	4182268	13257	1296	30882
2020-07-00 00:00:00	2020	7	Canada	82010	57254433	52996949	23620	5290717	4759937	12108	344	16809
2020-08-00 00:00:00	2020	8	Canada	82046	55077784	53189033	26571	5570711	5062309	12636	191	5294
2020-09-00 00:00:00	2020	9	Canada	82672	NULL	NULL	26856	NULL	NULL	29810	171	8848
2020-10-00 00:00:00	2020	10	Canada	81215	NULL	NULL	25071	NULL	NULL	76686	839	21342

This group query was also run using a SQL-connection in JUPYTER to take advantage of python visualization modules to illustrate the trends of the combined data table. The result was very powerful where retail and food sales (adjusted/unadjusted) vs working hours vs COVID-19 cases could be compared based on the combined table.

Figure 4.2.1 illustrates how COVID-19 in Canada significantly impacted retail sales and hours in the first wave (Mar-May 2020). The latest data shows how retail sales have significantly rebounded whereas retail

working hours have not rebounded to the same degree. This is likely due to the impact of online sales where consumers are still buying goods but are avoiding physical locations that have impacted the need for retail workers.

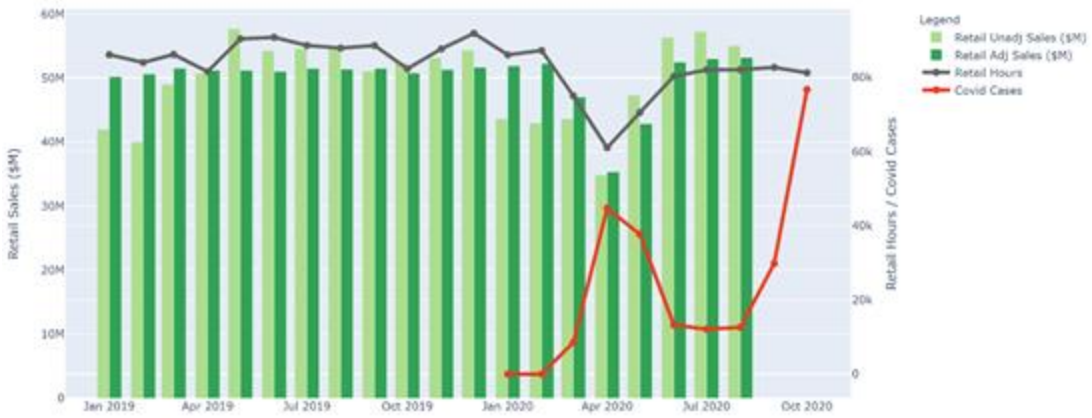


Fig 4.2.1: Retails Sales vs Retail Employment Hours vs COVID-19 Cases (2019-2020) – Canada

Figure 4.2.2 illustrates how COVID-19 in Canada significantly impacted food sales and hours in the first wave (Mar-May 2020). Food service sales and hours are improving since wave 1 but have not rebounded to the same degree as the retail data. Food service sales and hours have also been equally impacted. This can be explained as food service relies less on online sales as compared to the retail sector

Fig 4.2.2: Food Sales vs Food Employment Hours vs COVID-19 Cases (2019-2020) – Canada

Group queries (Jialing Cai):

Dataset:

There are three datasets used for my queries, including the COVID-19in Canada, retail trade sales by province and territory, and the data on working hours in Canada by industry. The data for COVID-19in Canada is found on the open.canada.ca website with Open Government License. The link for this source is https://open.canada.ca/data/en/dataset/261c32ab-4cfd-4f81-9dea-7b64065690dc. The retail sales data is obtained from Statistic Canada with file number 20-10-0008-02. The DOI link for the retail data is https://doi.org/10.25318/2010000801-eng. The working hours in Canada by industry is also obtained from Statistic Canada. The file number for this data set is 141000666 with link: https://open.canada.ca/data/en/dataset/223517c9-3759-45ce-bd83-c2e39ca8ffed?=undefined&wbdisable=true

Query summaries:

In this query, I aggregate and join the COVID-19table and retail sales table. For the COVID-19table, we sum together the original daily value to monthly value. Besides the COVID-19dataset, the other datasets we have are all monthly values. Then choose the value for the Canada range for COVID-19. Next, join the information for retail sales in 2020 and 2019 in the same month and Canada. The last thing is to get the difference in retail sales volume from 2019 to 2020 in the same month, and also the percentage value. The final result is included in the previous report. The main purpose of this query is to see how COVID-19 influences Canada's economy by comparing the retail sales volume difference from last year to this year. Is the retail sales volume related to the COVID-19cases? Which are the months that get the most influence.

Table 4.2.8: COVID-19 case vs Retail Trade performance in Canada

prname	cvmonth	total_cases	NAICS	volume_2020	volume_2019	difference	difference_perc...
Canada	1	4	Retail trade [44-45]	51891454	50195398	-1696056	-3.3800
Canada	2	11	Retail trade [44-45]	52262166	50626765	-1635401	-3.2300
Canada	3	8533	Retail trade [44-45]	47028013	51538062	4510049	8.7500
Canada	4	44688	Retail trade [44-45]	35352146	51212247	15860101	30.9700
Canada	5	37711	Retail trade [44-45]	42852368	51203886	8351518	16.3100
Canada	6	13257	Retail trade [44-45]	52488186	51028664	-1459522	-2.8600
Canada	7	12108	Retail trade [44-45]	52996949	51488286	-1508663	-2.9300
Canada	8	12636	Retail trade [44-45]	53189033	51375824	-1813209	-3.5300

The table above represents the difference in the overall retail performance change from 2019 to 2020 with the change of COVID-19 cases. The difference column is the retail sales volume of 2019 minus the retail sales volume of 2020. In the last column, the percentage difference is calculated as the difference divided by the retail sales volume of 2019. In January and February, the COVID-19 case is very low. Canada increased by approximately 3% of retail sales volume from 2019 to 2020. However, from February to March, the COVID-19 cases increased significantly from 11 cases to 8,533 cases. In March the retail sales volume decreased by 8.7%. In May the difference reaches the peak. The retail sales volume drops 30.9% compared to 2019. From May to June, the COVID-19 cases finally started to decrease from 37,711 to 13,257. As the retail sales volume bounces back to the normal level generally. March, April, and May are the three months that Canada has a lower retail sales volume in 2020. June, July and August are the three months that the volume goes back to usual.

For the next query, I join the COVID-19table and the retail sales table in the Alberta range. The procedure is the same as my group query one. I calculated the difference from last year to this year's retail sales volume in the same month. By comparing the percentage change and difference, it is easy to see the retail sales performance in Alberta. The purpose of this query is to discuss the overall doing for Alberta. How much the retail sales performance changes from last year to this year under the impact of COVID-19.

Table 4.2.9: COVID-19 case VS Retail Trade performance in Alberta

prname	cvmonth	total_cases	NAICS	volume_2020	volume_2019	difference	difference_percent
Alberta	3	747	Retail trade [44-45]	5967712	6895136	927424	13.4500
Alberta	4	4601	Retail trade [44-45]	4923680	6952206	2028526	29.1800
Alberta	5	1655	Retail trade [44-45]	5884701	6747223	862522	12.7800
Alberta	6	1098	Retail trade [44-45]	6776950	6731052	-45898	-0.6800
Alberta	7	2735	Retail trade [44-45]	6853786	6795187	-58599	-0.8600
Alberta	8	3059	Retail trade [44-45]	6742688	6782347	39659	0.5800

The table above shows the COVID-19cases and also the retail sales performance in Alberta. Same as Canada, March, April, and May are the three months that the retail volume of 2020 is lower than in 2019. From March to April, the Covid-19 cases increased rapidly from 747 cases to 4,601 cases. As the percentage difference is increased from 13.4% to 29.18%. However, while comparing Alberta's COVID-19cases with Canada's COVID-19cases. We can see the confirmed COVID-19cases in Alberta are still increasing from June to August. This fact leads to a terrible consequence that the retail sales volume in Alberta is still lower in August compared to the level of 2019. The retail sales performance in August is 0.58% lower in 2020 compared to 2019. If the Alberta government cannot control the spread of coronavirus, there is a big chance that the economy will get worse in the following months.

For the last query, I join and aggregate three tables. Based on my work from query one. I added the working hour's table in. For this table, I choose the working hours for the retail sales sector. Also, I choose gender as both sexes. These three tables are combined in the same month. For the last column, I divide the retail sales volume by the working hour. The purpose of this query is to see the relationship between working hours and retail sales. when the lock-down is over, everyone goes back to their work. Will the retail sales bounce back immediately to the normal level?

Table 4.2.10: Retail Sales Volume VS working hour during COVID-19

prname	cvmonth	total_cases	retail_volume	work_value	RW_ratio
Canada	1	4	51891454	86056	602.9964
Canada	2	11	52262166	87219	599.2062
Canada	3	8533	47028013	75015	626.9148
Canada	4	44688	35352146	61025	579.3060
Canada	5	37711	42852368	70531	607.5678
Canada	6	13257	52488186	80235	654.1807
Canada	7	12108	52996949	82010	646.2254
Canada	8	12636	53189033	82046	648.2831

The table above shows the retail sales performance and working hours during the outbreak of COVID-19. For the working hours data, I set the gender as both sexes, reference date as 2020 in the range of Canada. Also, I use the 'Wholesale and retail trade' as the NAICS for working hours to match the value with retail trade volume. From the table, we can find that as the COVID-19cases increased, the retail sales and working hours decreased. In general, under the same amount of working hours. If the RW_ratio is higher, indicates better retail sales performance. If the retail sales volume is about the same, RW_ratio will increase as the working hour decreases. From the table, we can see the lowest RW_ratio occurs in April (579.3). April is also the month that the COVID-19 case volume reaches the peak. Under the influence of COVID-19, the maximum RW_ratio occurs in June (654.18). June is also the month we can see a significant drop for COVID-19 cases. By comparing June and July, we can see the working hours in June are less than July. However, the retail sales volume in June is almost the same. For this output, we can see the COVID-19 case has a negative relationship between retail sales and working hours. There is a positive relationship between the working hours and the retail sales volume.

Group queries (Zhan Lin):

We want to find out how COVID-19 19 impacts Alberta food services and drinking places business. More specifically, we compare 2020 to 2019 on the business monthly income and check the COVID-19 cases and business income month to month in 2020. To answer these questions, we use data from two sources: COVID-19 Canada monthly cases view and Food Services and Drinking Places table. We get the result with the SQL SELECT statement with MySQL Workbench.

Table 4.2.7: Group query - Food Services and Drinking Places and COVID-19 Canada Monthly

last_month	current_month	last_month_value	current_month_value	service_change_pct	current_month_19_value	current_month_year_change_value	current_month_year_change_pct	last_month_total_case	current_month_total_case	case_change_pct
2020-03	2020-04	532162	$ 344,930.00	-35.18	$ 793,355.00	$ (448,425.00)	-56.52	747	4601	515.93
2020-04	2020-05	344930	$ 496,748.00	44.01	$ 855,555.00	$ (358,807.00)	-41.94	4601	1655	-64.03
2020-05	2020-06	496748	$ 651,668.00	31.19	$ 865,884.00	$ (214,216.00)	-24.74	1655	1098	-33.66
2020-06	2020-07	651668	$ 707,853.00	8.62	$ 889,756.00	$ (181,903.00)	-20.44	1098	2735	149.09
2020-07	2020-08	707853	$ 725,413.00	2.48	$ 893,418.00	$ (168,005.00)	-18.8	2735	3059	11.85

1. Food services and drinking places business keeps recovering since April
2. 3. The business got 44%, and 31% increased from April to May, and from May to Jun, COVID-19 cases got decreased by 64% and 33.66% in the same period
3. COVID-19 case surge since Jun, however, the business still increases slowly but still drops around 20% if we compare to the same month in 2019.
4. Year 2020's dollar number is less than Year 2019's dollar number when we compare monthly by monthly but their gap keeps reducing.

Group queries (Bowen Li):

Working hours & COVID-19 cases

In this part, we would like to answer how the COVID-19 cases influence the working hours in Canada and Alberta. the **total_work_hours** and **covid_canada_montly** are joined together. **Total_work_hour** table is available at the schema after clean-up in Bowen's queries. As for **covid_cananda_monthly** is available at the schema under views. It has been converted to a monthly format with active cases and total cases. Table 4.2.4 and Table 4.2.5 shows the result of COVID-19 cases and total working hours in Canada and Alberta. The result is obtained from the SQL SELECT statement with MySQL Workbench.

To get the output as below, first, the data on Canada are extracted from the **covid_canada_monthly** table as A. Secondly, we aggregate the work_value to get toal_work_hour in Canada for both sexes group by date and this would be table B. And then we join A with B based on the same year and month. As a result, we would get the total working hours and COVID-19 cases information for 2020.

Table 4.2.4: Total working hours vs total COVID-19 cases – Canada

prname	cvyear	cvmonth	total_cases	deaths	active_cases	total_work_hour
Canada	2020	1	4	0	4	608009
Canada	2020	2	11	0	11	614465
Canada	2020	3	8533	95	1700	505471
Canada	2020	4	44688	3088	22020	440440
Canada	2020	5	37711	4111	32632	501030
Canada	2020	6	13257	1296	30882	561271
Canada	2020	7	12108	344	16809	544301
Canada	2020	8	12636	191	5294	545267
Canada	2020	9	29810	171	8848	607252
Canada	2020	10	76686	839	21342	570162

In Canada, May has the most COVID-19 active_cases that is probably because total working hours in May started to grow. In September, the total working hours reached the maximum which is about the same as January. We reckoned that could be the reason that COVID-19 cases exploded again in October.

A similar procedure is conducted for Alberta. From table 4.2.5, we can tell June is the best month for Albertan handling procedures. It is the month that COVID-19 cases hit the bottom with the working hours going back to normal. In September, the total working hours reached a peak while COVID-19 cases increased dramatically.

Table 4.2.5: Total working hours vs total COVID-19 cases – Alberta

prname	cvyear	cvmonth	total_cases	deaths	active_cases	total_work_hour
Alberta	2020	3	747	9	238	66621
Alberta	2020	4	4601	80	1826	57758
Alberta	2020	5	1655	54	1417	63742
Alberta	2020	6	1098	11	445	69175
Alberta	2020	7	2735	42	916	66359
Alberta	2020	8	3059	43	1153	66877
Alberta	2020	9	4160	28	1483	75276
Alberta	2020	10	10183	56	3014	69783

In this query, the information on 2019 and 2020 are extracted separately as **d19** and **d20** and aggregate work_value to get monthly total working hours in 2019 and 2020 at the same time from **work_hour_data**. We join **d19** with **d20** and name it as table **A**.

Select **covid_canada_montly** table that geography is Canada and name it as table **B**. We join table **A** and table **B** based on month.

Select month, working hours information in 2019 and 2020, calculate the growth rate from 2019 to 2020 at the same month and COVID-19 cases information after joining A and B.

Table 4.2.6: Total working hours growth rate from 2019 to 2020 vs total COVID-19 cases

monthv	wh20	wh19	rate	active_cases
1	38000.5625	37927.6250	0.00192307	4
2	38404.0625	37363.6875	0.02784455	11
3	31591.9375	37355.4375	-0.15428811	1700
4	27527.5000	35213.3125	-0.21826440	22020
5	31314.3750	40076.0000	-0.21862524	32632
6	35079.4375	41066.9375	-0.14579855	30882
7	34018.8125	37746.7500	-0.09876181	16809
8	34079.1875	37331.3125	-0.08711521	5294
9	37953.2500	40394.6875	-0.06043957	8848
10	35635.1250	36436.6250	-0.02199710	21342

We found that COVID-19 cases have an immense impact on total working hours in both Alberta and Canada. To eliminate the factor of seasonal change, we compare the working hours in 2019 and 2020 monthly. In January and February, working hours increased since there are only several COVID-19 cases in Canada. After March, total working hours dropped consistently. What is interesting is that in April and October, the COVID-19 cases are about the same while the working hours in April decreased by 20 percent while in October, it only decreased by about 2 percent. That is probably because people are getting used to the style and less restriction compared with the first wave of COVID-19 outbreak.

Group queries (Adaeze Nwigwe):

Travel trends & COVID-19 cases

The purpose of the analysis was to compare 2020 travel trends to travel trends from the past years to see how the trends have changed and to see how different the trends have been since the COVID-19 pandemic. We will also explore how travel trends vary between countries and observe any possible correlation between COVID-19 cases in various countries and their travel patterns to Canada.

In order to carry out this analysis the two tables that were joined were the COVID-19 monthly data view table and the travelers to Canada table.

Travel trends show a steady decrease in travelers from all countries to Canada between 2019 and 2020 starting in March. The drop in number of travellers in March corresponds to the gradual increase in COVID-19 numbers starting from March 2020.Percent decrease in travelers between 2019 and 2020 of up to 99% is correlated to increase in COVID-19 cases and introduction of travel restriction measures. Canada introduced travel restrictions and closed its international borders to non-Canadians as early as March 18th 2020, in an attempt to reduce the spread of COVID-19 and to keep the population safe.

Table 4.2.2: Travellers & COVID-19 by country

month(REF_DATE)	country	2019	2020	percent_change	total_cases
3	Argentina	1845	954	-48.2900	1932
4	Argentina	2156	12	-99.4400	6612
5	Argentina	1821	6	-99.6700	23858
6	Argentina	1892	29	-98.4700	92108
7	Argentina	1815	99	-94.5500	233456
8	Argentina	1928	36	-98.1300	444486
9	Argentina	1862	18	-99.0300	670740
1	Australia	33171	31454	-5.1800	14
2	Australia	31878	31944	0.2100	36
3	Australia	31752	12816	-59.6400	9064
4	Australia	31405	0	-100.0000	4378
5	Australia	31887	74	-99.7700	878
6	Australia	30045	145	-99.5200	1164
7	Australia	31343	364	-98.8400	17072
8	Australia	29754	562	-98.1100	18734
9	Australia	30771	523	-98.3000	2786

The introduction of the travel restrictions is very evident in the data. Although travel restrictions were introduced in early March and the travel trends to Canada show a major drop starting in March 2020, COVID-19 cases in Canada and other countries continue rising through the Spring. This shows that although travel restrictions may have helped slow the spread of the virus, they were not sufficient to stop the spread. A general drop in COVID-19 cases numbers is only seen in the summer months.

Despite the general drop in COVID-19 cases numbers during the months of June and July, travel trends still show a high percentage decrease in the summer months as travel restrictions stayed in place.

Table 4.2.3: Travellers & COVID-19 by country

month(REF_DATE)	country	2019	2020	percent_change	total_cases
4	China	62115	1199	-98.0700	3406
5	China	63271	410	-99.3500	368
6	China	61954	1321	-97.8700	1304
7	China	60145	832	-98.6200	5418
8	China	62249	1205	-98.0600	4812
9	China	58781	1079	-98.1600	1266
1	Denmark	3032	3144	3.6900	0
2	Denmark	3571	3223	-9.7500	4
3	Denmark	3594	1653	-54.0100	5150
4	Denmark	3468	147	-95.7600	12862
5	Denmark	3782	326	-91.3800	5250
6	Denmark	3647	0	-100.0000	2236
7	Denmark	3466	0	-100.0000	1948
8	Denmark	3528	0	-100.0000	5950
9	Denmark	3550	0	-100.0000	21528
1	Finland	1558	1546	-0.7700	2
2	Finland	1320	1609	21.8900	4
3	Finland	1498	575	-61.6200	2620
4	Finland	1567	142	-90.9400	7186
5	Finland	1604	0	-100.0000	3840
6	Finland	1597	0	-100.0000	766
7	Finland	1608	0	-100.0000	428
8	Finland	1398	8	-99.4300	1308

5.0 Conclusion

COVID-19 pandemic continues to impact Canada and the rest of the world where cases are significantly increasing in multiple waves with each subsequent wave worse than the last. In Canada - the first wave has significantly impacted key economic metrics represented by retail, food and employment hours.

Overall retail sales in Canada have been significantly impacted where year over year sales have dropped from 52,262,166 to 47,028,013. The highest percentage drop -30.9% from 2019 to 2020 occured in April which was during the peak of the first wave. Employment was similarly impacted with year over year working hours decreasing from 614,465 to 505,471.

Alberta has trended similarly to Canada with respect to rebounding retail sales after the first wave however the latest data in August 2020 has shown sales falling. As COVID-19 cases increase in Alberta it will be important to continue to track how Alberta is performing compared to Canada.

To avoid human-to-human infection, the Canadian government issued official travel restrictions in March 2020. Only Canadian citizens or permanent residents of Canada have the permit to enter Canada (Government of Canada, 2020). Most of the non-essential travel to Canada is stopped under the restriction. The impact of the travel restrictions is evident where we see up to 99% drop between 2019 and 2020 travel trends to Canada globally. Despite the travel restriction COVID-19 cases continue to rise across the world, showing that travel restrictions were necessary to slow the spread of the virus but not the only necessary measure to overcome the pandemic

Key findings from the study:

1. Between March and April was the most significant percentage decrease in food services and drinking places business and also the biggest percentage increase in COVID 19 cases. Food services and drinking places business has been recovering since April.
2. Food business sales increased 44%, and 31% from April to May, and from May to Jun, COVID 19 cases got decreased by 64% and 34% in the same period. COVID 19 case surge since Jun, however,

the food business sales still increased slowly but still drops around 20% if we compare to the same month in 2019.

3. Year 2020's dollar number is less than Year 2019's dollar number when we compare monthly by monthly but this gap keeps tightening.
4. Canada's retail sales dropped rapidly in March, April, and May. After the COVID-19 cases decreased, the retail sales increased to comparable 2019 values in June, July and August.
5. The maximum retail sales volume decrease for Alberta is in April. The retail sales volume dropped 29% compared to 2019 on the same month. The retail sales volume does not jump back to normal level in August as the COVID-19cases keep increasing.
6. Working hours dropped dramatically during the first outbreak of COVID-19 in Canada. In September, working hours for both Alberta and Canada went back to normal with less COVID-19 cases. However, with the second wave of outbreak, the working hours barely dropped in October.
7. Travel restrictions due to COVID-19 caused a major drop in the travel trends starting in March 2020. The drop in travel trends show a decrease around 40-60% in March and April and a decrease of up to 99% in the subsequent months as COVID-19 cases increase globally.
8. Eastern Asian countries' travel trends show a percent decrease in 2020 starting from February and this corresponds to the appearance of the first COVID-19 cases in Easter Asia in January and February.

While COVID-19 cases are on the rise in Canada and around the world - there are still positive signs that show the economy is capable of recovering quickly. Economic indicators in retail, food and working hours for Canada and individual provinces such as Alberta demonstrated strong rebounds as COVID-19 case numbers were able to stabilize. There is hope that if Canada and its respective provinces can keep cases under control - the economy in Canada might see the light at the end of the tunnel in 2021.

6.0 Discussion

Commentary from John P. Campbell

The datasets used in evaluating the impact of COVID-19 were dependable and little data cleaning and wrangling was required which helped drive confidence in evaluating the data. There were two key techniques that were learned in the individual and group data exploration queries. First – the SQL interface via mySQL to query the data was powerful in quickly modifying data tables to an appropriate format including processes to aggregate, filter, or sort. Converting the COVID-19 data from daily to monthly is a very good example where the creation of a monthly 'view' of this data table was an easy task. The SQL connection in JUPYTER was extremely powerful to import the data tables created by each query and use modules such as plotly to visualize these data tables in a meaningful way.

I would have used Jupyter earlier on in the individual data exploration as this was a more impactful platform to investigate the result of the data query. Jupyter would have also been useful in individual and group exploration where queries could be interactive. For example – a user could select the appropriate province and date timeframe to refresh a query based on these inputs. Various queries could be imported into separate dataframes to then contrast to each other.

For future analysis – contrasting the impact of COVID-19 on each province would be of particular interest. Each province has addressed the COVID-19 pandemic with very different policies and restrictions. How have Provinces such as those within the Atlantic bubble performed on retail and food sales while keeping COVID-19 cases down as compared to other Provinces? In the latest wave – have these Provinces better positioned themselves economically?

Commentary from Jialing Cai

Throughout the project, there are a couple of things that are new and interesting to me. First, I learn how to import raw data into MySQL. The data types for different columns need to be analyzed first before we put the dataset into MySQL workbench. Especially for DateTime, the 'STR_TO_DATE' command will be very helpful to change the DateTime from string to date. For the individual queries, I get familiar with the MySQL subquery. It helps me to get a query nested with other queries, and only output the valuable information. From the beginning of the course, I do not understand how nested queries work. However, right now it is easy for me to use the filter and join information from selected subqueries. The other thing I find useful is that I can do the calculation on MySQL for different columns. The column values can be added, multiplied, subtracted by using a simple MySQL command.

There are a couple of things I would do differently. I will find the retail sales performance not only in Canada but also in other countries. Because it is a worldwide outbreak of COVID-19. It will be more meaningful if I can compare Canada's economic situation with other countries. Perhaps I can predict the feature economic performance for Canada. For future work, I will analyze the data in a more detailed way. I will look through different retail sectors and see which one dropped the most. The other thing I will do differently is I will try to look at the online sales data and combine it with my current data. Since the data I obtained from Statistic Canada are all for physical stores. While people quarantined at home, most of them may choose to shop online instead of going out. Using both the online sales volume and the retail sales volume can describe Canada's economic performance in a more comprehensive way.

For future work. The next thing I will do for this project is that I will add more visualization graphs for the result. Right now, the outputs are all directly generated from MySQL workbench. The format of the output is all tables. If I can put these values in Python, use Plotly and Matplotlib to draw the observation. The result can be more attractive and easier to understand. Moreover, I will try to connect the retail sales data with the traveler's dataset. Theoretically, retail sales, especially for the sectors related to tourists will be affected badly while the country experiences a lockdown. In the year 2020, Calgary's annual stampede festival was canceled. By connecting these two tables, I can analyze the economic loss due to the lock-down period. The other thing I will do is updating my retail sales dataset. I noticed on Statistic Canada; the last retail trade sales amount is updated to October. My current data is only until August 2020.

Commentary from Zhan(Gerry) Lin

Through the data exploration of this project, it confirms that do not make any conclusion before you put enough time and effort into analyzing the data. Naturally, we might think the business performance trend will go in the opposite direction with the trend of the COVID-19 cases. After we dug into the data we found out we were right for the first/begin phase of the COVID-19 pandemic but wrong after that.

More importantly, the project found these basic analysis steps are still true. First of all, check the data quality before using it. Most of our data come from Statistics Canada which has been clean but we still found data problems, for example, missing data in the three territories, some period data is wrong. Second of denormalizing database tables can make data query easily but it may cause us to neglect data problems. The Statistics Canada dataset includes nest relationship like we get total Canada, provinces and territories data in the same table, or same like categories and subcategories. This design simplifies the steps of getting aggregate data for parent nodes in a hierarchy. However, it may hide the data problem at some child levels. Last, of all, SQL is the easier and effective way to analyze data but it can't tell us data problems directly or see the big picture, data visualization tools can help us to fill in the gap.

Commentary from Bowen Li

In this project, the result verifies our assumption at the beginning that COVID-19 hit the economies dramatically. The cases decrease a lot when the government executes strict restrictions. And in some cases, it might mean less working hours and less retail sales. As for the technical aspect, I found out that SQL is relatively easier than other programming languages while it can deliver data selection and combination, etc at the same time. However, it would be limited when it comes to visualization.

I would connect MySQL Workbench with JUPYTER notebook to get access to visualization for better effect of presentation.

In the future, I would compare the working hour drop rate in the previous pandemic with COVID-19 and predict the recovery time.

Commentary from Adaeze Nwigwe

Through the data exploration in this project I learned the importance of using reliable and well-structured data. The data exploration was made easy and simple by the reliability and the good data organization which made manipulation and analysis of the data to answer questions quite easy and straightforward.

Even though the data set was quite large and contained thousands of records, my SQL workbench was an excellent and powerful tool to be able to handle the data and slice it and observe the different parts of the data that I was interested in. It was also nice to be able to share database with my team members so we were able to work with each other's tables and utilize each other's view tables when appropriate. The only downfall with my SQL workbench is that it does not provide data visualization ability, since that is not the intent of the tool, so it needs to be coupled with any data visualization tool to truly get a better picture of the data that is being explored. I used excel for some basic data visualization, but in the future I would probably use Jupyter or other data visualization tools that provide more interactivity.

In the future, as the pandemic comes to an end and hopefully travel trend numbers start increasing, I would like to conduct the analysis again to see if due to the lack of travel in the year 2020 if we see a major rise in travel trends once the pandemic is over.

7.0 Acknowledgement

We would like to express our special thanks and gratitude to Dr. Leanne Wu and Apoorve Chokshi for their help and guidance in this meaningful data exploration project.

8.0 Appendix

Group Query - John P. Campbell

In the group query outlined below – the objective was to review the impact of COVID-19 on overall retail sales (unadjusted and seasonally adjusted), food sales (unadjusted and seasonally adjusted) and working hours specific to those industries.

Datasets used in the group query are: Dataset #1: COVID-19 in Canada [2], Dataset #3: Retail Trade Sales by Province and Territory [5], Dataset # 4: Monthly Survey of Food Services and Drinking Places [9] and Dataset # 5: Data on working hour in Canada by Industry [10].

The query is a set of nested subqueries where the process involves creating a join of two tables of key columns and then nesting that query with another. Each table was filtered to geo = 'Canada' so that the key joins between each table only had to be on date.

- To setup the query - retail sales table had to be joined with itself to create separate columns for adjusted and unadjusted sales. This table was named - **'Canada_Retail'**. This was also done for the food table - named as **'Canada_Food'**.

- The **'Canada_Retail'** and **'Canada_Food'** tables were the first two tables joined using inner join – named as **'Canada_Sales'** on a common date. The retail sales and food data have the longest history of data so it was structured as such to maintain pre-2020 data.

- Working hours had to be joined with itself as well to create separate columns for both Retail and Food Hours - named **'Canada_Retail_Hours'** and **'Canada_Food_Hours'**. This combined table was named **'Canada_Hours'**.

- **'Canada_Sales'** and **'Canada_Hours'** tables were then joined on date - to create a new table named **'data_tb1'**.

- This table, '**data_tb1'** was then joined with the COVID-19 Canada data – simply named as **'covid_tb1'** to create the final query. The key columns joined in the final query are shown below:

Fig 1: Group Query- Combined Data Table

date	year	month	geo	Retail_Hours	RETAIL - Unadjusted Sales ($M)	RETAIL - Seasonally Adjusted Sales ($M)	Food_Hours	FOOD - Unadjusted Sales ($M)	FOOD - Seasonally Adjusted Sales ($M)	COVID Cases	COVID Deaths	COVID Active Cases
2020-01-00 00:00:00	2020	1	Canada	86056	43646873	51891454	30281	5702234	6468869	4	0	4
2020-02-00 00:00:00	2020	2	Canada	87219	42981961	52262166	31179	5682623	6482288	11	0	11
2020-03-00 00:00:00	2020	3	Canada	75015	43641006	47028013	17327	4002082	4152415	8533	95	1700
2020-04-00 00:00:00	2020	4	Canada	61025	34869392	35352146	10905	2442264	2508885	44688	3088	22020
2020-05-00 00:00:00	2020	5	Canada	70531	47342177	42852368	13077	3448650	3252320	37711	4111	32632
2020-06-00 00:00:00	2020	6	Canada	80235	56342104	52488186	20216	4416741	4182268	13257	1296	30882
2020-07-00 00:00:00	2020	7	Canada	82010	57254433	52996949	23620	5290717	4759937	12108	344	16809
2020-08-00 00:00:00	2020	8	Canada	82046	55077784	53189033	26571	5570711	5062309	12636	191	5294
2020-09-00 00:00:00	2020	9	Canada	82672	NULL	NULL	26856	NULL	NULL	29810	171	8848
2020-10-00 00:00:00	2020	10	Canada	81215	NULL	NULL	25071	NULL	NULL	76686	839	21342

Group Query (Part 1 - Left ; Part 2 - Right)

```
select
       data_tbl.date,
       year(data_tbl.date) as year,
       month(data_tbl.date) as month,
       data_tbl.geo,
       data_tbl.Retail_Hours,
       data_tbl.RETAIL_Unadj as 'RETAIL - Unadjusted Sales ($M)',
       data_tbl.RETAIL_Adj as 'RETAIL - Seasonally Adjusted Sales ($M)',
       data_tbl.Food_Hours,
       data_tbl.Food_Unadj as 'FOOD - Unadjusted Sales ($M)',
       data_tbl.Food_Adj as 'FOOD - Seasonally Adjusted Sales ($M)',
       covid_tbl.total_cases as 'COVID Cases',
       covid_tbl.deaths as 'COVID Deaths',
       covid_tbl.active_cases as 'COVID Active Cases'
from
(select Canada_Hours.date,
       Canada_Hours.geo,
       Canada_Hours.Retail_Hours,
       Canada_Sales.RETAIL_Unadj,
       Canada_Sales.RETAIL_Adj,
       Canada_Hours.Food_Hours,
       Canada_Sales.FOOD_Unadj,
       Canada_Sales.FOOD_Adj
from
(select Canada_Retail_Hours.date as date,
        Canada_Retail_Hours.geo as geo,
        Canada_Retail_Hours.hours as Retail_Hours,
        Canada_Food_Hours.hours as Food_Hours
from
(select geo,reference_date as date,work_value as hours from working_hour_data
    where sex = 'Both sexes'
    and geo = 'Canada'
    and reference_date >= '2010-01-00 00:00:00'
    and naics like '%Wholesale and retail%') as Canada_Retail_Hours
inner join
    (select geo,reference_date as date,work_value as hours from working_hour_data
    where sex = 'Both sexes'
    and geo = 'Canada'
    and reference_date >= '2010-01-00 00:00:00'
    and naics like '%Food Service%') as Canada_Food_Hours
on (Canada_Retail_Hours.date = Canada_Food_Hours.date)
)
as Canada_Hours
left join
(select Canada_Retail.date as date,
       Canada_Retail.geo as geo,
       Canada_Retail.RETAIL_Unadj as RETAIL_Unadj,
       Canada_Retail.RETAIL_Adj as RETAIL_Adj,
       Canada_Food.FOOD_Unadj as FOOD_Unadj,
       Canada_Food.FOOD_Adj as Food_Adj
from
(select Canada_Unadj_Retail.date,
        Canada_Unadj_Retail.geo,
        Canada_Unadj_Retail.volume as RETAIL_unadj,
        Canada_SeasonAdj_Retail.volume as RETAIL_Adj
from
(select reference_date as date,geo,volume from retail_sales_data
     where geo='Canada' and
           NAICS = 'retail trade [44-45]' and
           Adjustments = 'Unadjusted' and
           reference_date >= '2010-01-00 00:00:00')
     as Canada_Unadj_Retail
inner join
(select reference_date as date, geo,volume from retail_sales_data
     where geo='Canada' and
     NAICS = 'retail trade [44-45]' and
     Adjustments = 'Seasonally adjusted' and
     reference_date >= '2010-01-00 00:00:00') as Canada_SeasonAdj_Retail
on (Canada_Unadj_Retail.date = Canada_SeasonAdj_Retail.date)
) as Canada_Retail
```

Query Continued (Part3)

```
inner join
(select Canada_Unadj_Food.date,
       Canada_Unadj_Food.geo,
       Canada_Unadj_Food.data_value as Food_Unadj,
       Canada_SeasonAdj_Food.data_value as Food_adj
from
(select ref_date as date,geo,data_value from food_drinking_service
     where geo='Canada' and
           NAICS = 'Total, food services and drinking places' and
           seasonal_adjustment = 'Unadjusted' and
           ref_date >= '2010-01-00 00:00:00')
     as Canada_Unadj_Food
inner join
(select ref_date as date, geo,data_value from food_drinking_service
     where geo='Canada' and
     NAICS = 'Total, food services and drinking places' and
     seasonal_adjustment = 'Seasonally adjusted' and
     ref_date >= '2010-01-00 00:00:00') as Canada_SeasonAdj_Food
on (Canada_Unadj_Food.date = Canada_SeasonAdj_Food.date)
) as Canada_Food
on (Canada_Retail.date = Canada_Food.date)
)
as Canada_Sales
on (Canada_Hours.date = Canada_Sales.date)) as data_tbl
left join
(select prname as geo,
        cast(concat(cvyear,'-',cvmonth,'-00 00:00:00') as datetime) as date,
        total_cases,
        deaths,
        active_cases
        from covid_canada_monthly
        where prname = 'Canada'
        ) as covid_tbl
on (data_tbl.date = covid_tbl.date) order by data_tbl.date;
```

Group Query – Jialing Cai

Dataset information

There are three datasets used for my queries, including the Covid-19 in Canada, retail trade sales by province and territory, and the data on working hours in Canada by industry. The data for Covid-19 in Canada is found on the open.canada.ca website with Open Government License. The link for this source is https://open.canada.ca/data/en/dataset/261c32ab-4cfd-4f81-9dea-7b64065690dc. The retail sales data is obtained from Statistic Canada with file number 20-10-0008-02. The DOI link for the retail data is https://doi.org/10.25318/2010000801-eng. The working hours in Canada by industry is also obtained from Statistic Canada. The file number for this data set is 141000666 with link :https://open.canada.ca/data/en/dataset/223517c9-3759-45ce-bd83-c2e39ca8ffed?=undefined&wbdisable=true

```
select prname,cvmonth,total_cases,R2019.NAICS,volume_2020,volume_2019,
(volume_2019-volume_2020) as difference,
 ((volume_2019-volume_2020)/volume_2019)*100 as difference_percent
from (
select prname,cvmonth,total_cases from covid_canada_monthly
where prname='Canada'
) as covid
join(
select reference_date,volume as volume_2020,NAICS from retail_sales_data
where year(reference_date)=2020
and geo='Canada'
and NAICS like 'Retail trade%'
and Adjustments='Seasonally adjusted') as R2020
join(
select reference_date,volume as volume_2019,NAICS from retail_sales_data
where year(reference_date)=2019
and geo='Canada'
and NAICS like 'Retail trade%'
and Adjustments='Seasonally adjusted') as R2019
on cvmonth=month(R2019.reference_date) and cvmonth=month(R2020.reference_date);
```

Group query 1 from Jialing Cai

In this query, I aggregate and join the Covid-19 table and retail sales table. For the Covid-19 table, we sum together the original daily value to monthly value. Besides the Covid-19 dataset, the other datasets we have are all monthly values. Then choose the value for the Canada range for Covid-19. Next, join the information for retail sales in 2020 and 2019 in the same month and Canada. The last thing is to get the difference in retail sales volume from 2019 to 2020 in the same month, and also the percentage value. The final result is included in the previous report. The main purpose of this query is to see how Covid-19 influence Canada's economy by comparing the retail sales volume difference from last year to this year. Is the retail sales volume relate to the Covid-19 cases? Which are the months that get the most influence.

```sql
select prname,cvmonth,total_cases,R2019.NAICS,volume_2020,volume_2019,
(volume_2019-volume_2020) as difference,
 ((volume_2019-volume_2020)/volume_2019)*100 as difference_percent
from (
select prname,cvmonth,total_cases from covid_canada_monthly
where prname='Alberta'
) as covid
join(
select reference_date,volume as volume_2020,NAICS from retail_sales_data
where year(reference_date)=2020
and geo='Alberta'
and NAICS like 'Retail trade%'
and Adjustments='Seasonally adjusted') as R2020
join(
select reference_date,volume as volume_2019,NAICS from retail_sales_data
where year(reference_date)=2019
and geo='Alberta'
and NAICS like 'Retail trade%'
and Adjustments='Seasonally adjusted') as R2019
on cvmonth=month(R2020.reference_date)=month(R2019.reference_date);
```

Group query 2 from Jialing Cai

In this query, I join the Covid-19 table and the retail sales table in the Alberta range. The procedure is the same as my group query one. I calculated the difference from last year to this year's retail sales volume in the same month. By comparing the percentage change and difference, it is easy to see the retail sales performance in Alberta. The purpose of this query is to discuss the overall doing for Alberta. How much the retail sales performance changes from last year to this year under the impact of Covid-19.

```sql
select prname,cvmonth,total_cases,retail_volume,work_value,retail_volume/work_value as RW_ratio from
(select prname,cvmonth,total_cases from covid_canada_monthly
       where prname='Canada') as covid
join(
select NAICS,volume as retail_volume,geo,reference_date from retail_sales_data
      where geo='Canada'
      and NAICS like 'Retail trade%'
      and year(reference_date)=2020
      and Adjustments='Seasonally adjusted') as retail
join(
select reference_date, geo, naics, sex, work_value from working_hour_data
       where geo='Canada'
       and naics like '%retail trade%'
       and year(reference_date)=2020
       and sex = 'Both sexes') as working
on cvmonth = month(retail.reference_date) and cvmonth=month(working.reference_date);
```

Group Query 3 from Jialing Cai

For the last query, I join and aggregate three tables. Based on my work from query one. I added the working hour's table in. For this table, I choose the working hours for the retail sales sector. Also, I choose gender as both sexes. These three tables are combined in the same month. For the last column, I divide the retail sales volume by the working hour. The purpose of this query is to see the relationship between working hours and retail sales. when the lock-down is over, everyone goes back to their work. Will the retail sales bounce back immediately to the normal level?

Group Query - Bowen Li

-- Total work hour and cases in Canada

In the queries below, first the data on Canada are extracted from **covid_canada_monthly** table as **A**. Secondly, we aggregate the work_value to get toal_work_hour in Canada for both sexes group by date and this would be table **B**. And then we join A with B based on same year and month. As a result, we would get the total working hours and covid cases information on 2020.

```
SELECT A.*,
       B.total_work_hour
FROM   (SELECT *
        FROM   covid_canada_monthly
        WHERE  prname = 'Canada') A
       INNER JOIN (SELECT *,
                          Sum(work_value) total_work_hour
                   FROM   working_hour_data
                   WHERE  geo = ( 'Canada' )
                          AND sex = 'Both sexes'
                   GROUP  BY reference_date) B
               ON ( Year(B.reference_date) = A.cvyear
                    AND Month(B.reference_date) = A.cvmonth );
```

-- Total work hour and cases in Alberta

Similar steps as above, except that the area is set as Alberta.

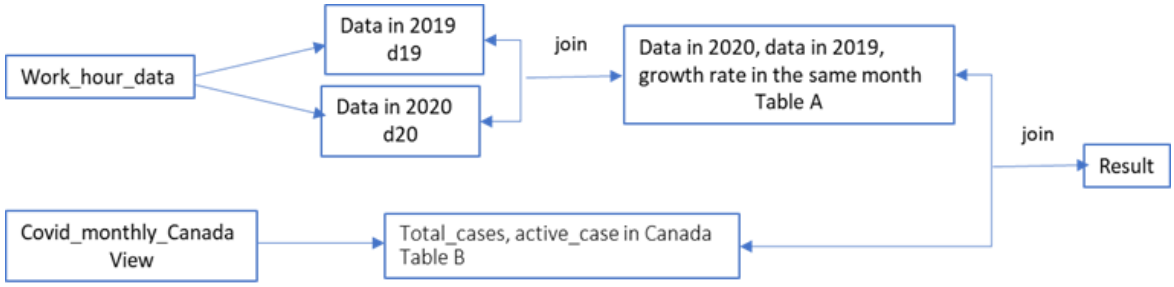
```
SELECT A.*,
       B.total_work_hour
FROM   (SELECT *
        FROM   covid_canada_monthly
        WHERE  prname = 'Alberta') A
       INNER JOIN (SELECT *,
                          Sum(work_value) total_work_hour
                   FROM   working_hour_data
                   WHERE  geo = ( 'Alberta' )
                          AND sex = 'Both sexes'
                   GROUP  BY reference_date,
                             geo) B
               ON ( Year(B.reference_date) = A.cvyear
                    AND Month(B.reference_date) = A.cvmonth
                    AND A.prname = B.geo );
```

-- Total work hour and cases comparison between 2019 & 2020 in Canada

In this query, the information on 2019 and 2020 are extracted separately as **d19** and **d20** and aggregate work_value to get monthly total working hours in 2019 and 2020 at the same time from **work_hour_data**. We join **d19** with **d20** and name it as table **A**.

Select **covid_canada_montly** table that geography is Canada and name it as table **B**. We join table **A** and table **B** base on month.

Select month, working hours infomation in 2019 and 2020, calculate the growth rate from 2019 to 2020 at the same month and covid cases information after joining A and B.




```
SELECT A.monthv,
       A.wh20,
       A.wh19,
       ( ( A.wh20 - A.wh19 ) / A.wh19 ) rate,
       B.active_cases
FROM   (SELECT Month(d20.reference_date) monthv,
               d20.wh20,
               d19.wh19
        FROM   (SELECT *,
                       Avg(work_value) wh20
                FROM   working_hour_data
                WHERE  geo = ( 'Canada' )
                       AND sex = 'Both sexes'
                       AND Year(reference_date) = 2020
                GROUP  BY reference_date) d20
               INNER JOIN (SELECT *,
                                  Avg(work_value) wh19
                           FROM   working_hour_data
                           WHERE  geo = ( 'Canada' )
                                  AND sex = 'Both sexes'
                                  AND Year(reference_date) = 2019
                           GROUP  BY reference_date) d19
                       ON ( Month(d20.reference_date) =
                            Month(d19.reference_date) )) A
       INNER JOIN (SELECT *
                   FROM   covid_canada_monthly
                   WHERE  prname = 'Canada') B
               ON ( A.monthv = B.cvmonth );
```

Group Query - Adaeze Nwigwe

The tables joined in this query are the COVID 19 monthly data view table and the travellers to canada table. The query reorganizes the travellers table to obtain 2019 travelers values and 2020 travelers value in two separate columns. The percent difference between 2019 and 2020 travelers numbers is then calculated in the query. Since the 2020 travelers data only went up to September 2020, we excluded data from October, November and December 2019 from the analysis. The query then performs an inner join with the COVID-19 global monthly data. Data is inner joined so only rows that are in common between both data sets are shown. Data is joined on the country and on the month column, and then grouped by country and month for the analysis.

```
select month(REF_DATE),
country,
sum(ValueOfYear2019) '2019',
sum(ValueOfYear2020) '2020',
(((sum(ValueOfYear2020)-sum(ValueOfYear2019))/sum(ValueOfYear2019))*100) as percent_change,
sum(total_cases) as total_cases
from (
Select
REF_DATE,
country,
case when year(ref_date)=2019 then VALUE end ValueOfYear2019,
case when year(ref_date)=2020 then VALUE end ValueOfYear2020,
month(ref_date)
from travellers T
where month(REF_DATE)<=9 and year(ref_date)  in (2019,2020)  AND GEO='CANADA'
) a
 INNER JOIN `l02-8`.`covid_global_monthly` C
ON
    month(REF_DATE) = C.cvmonth AND
    COUNTRY = C.countriesandTerritories
    where country in ('China', 'Denmark', 'Finland')
    GROUP BY  COUNTRY, month(REF_DATE);
```

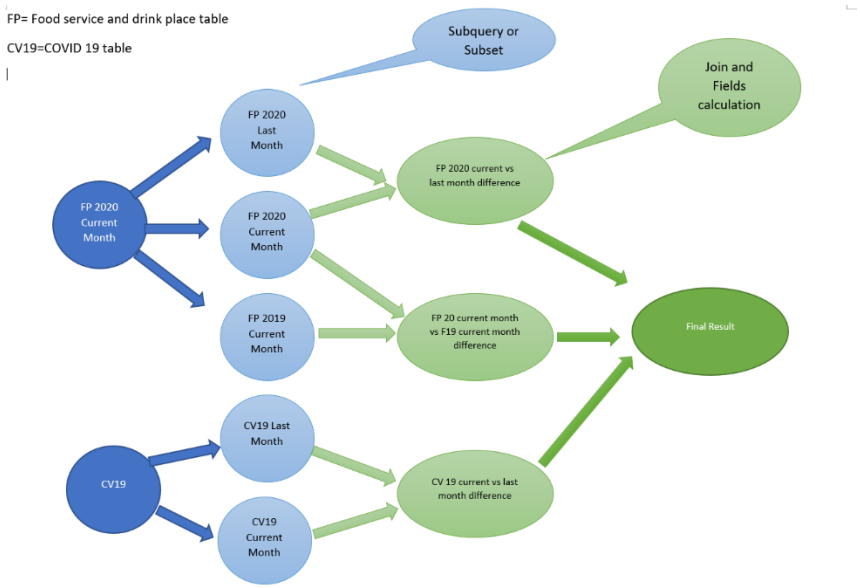
Group Query -- Zhan(Gerry) Lin

The query joins f**ood services and drink place business table** with **COVID 19 monthly Case view** to find out:

1. The business month to month change in 2020
2. The business change in the same month between 2020 and 2019
3. The COVID 19 case month to month change in 2020
4. Compare the change trend between COVID case change and business change

```
SELECT fds1.ref_date                    AS last_month,
       fds2.ref_date                    AS current_month,
       fds1.data_value                  AS last_month_value,
       fds2.data_value                  AS current_month_value,
       Round(( fds2.data_value - fds1.data_value ) / fds1.data_value * 100, 2)
                                        service_change_pct,
       fds3.data_value                  AS current_month_19_value,
       fds2.data_value - fds3.data_value AS current_month_year_change_value,
       Round(( fds2.data_value - fds3.data_value ) / fds3.data_value * 100, 2)
                                        current_month_year_change_pct,
       ccm1.total_cases                 last_month_total_case,
       ccm2.total_cases                 current_month_total_case,
Round(( ccm2.total_cases - ccm1.total_cases ) / ccm1.total_cases * 100, 2) case_change_pct
       FROM   food_drinking_service AS fds1 -- 2020 last month data
       INNER JOIN food_drinking_service AS fds2 -- 2020 current month  data
               ON Year(fds1.ref_date) = Year(fds2.ref_date)  -- join with same year
                  AND ( Month(fds1.ref_date) + 1 ) = Month(fds2.ref_date) -- join current month join last month
                  AND fds1.geo = fds2.geo
                  AND fds1.naics = fds2.naics
                  AND fds1.service_detail = fds2.service_detail
                  AND fds1.seasonal_adjustment = fds2.seasonal_adjustment
       INNER JOIN food_drinking_service AS fds3 -- 2019  current month  data
               ON Year(fds2.ref_date) = ( Year(fds3.ref_date) + 1 )
                  AND Month(fds2.ref_date) = Month(fds3.ref_date)
                  AND fds2.geo = fds3.geo
                  AND fds2.naics = fds3.naics
                  AND fds2.service_detail = fds3.service_detail
                  AND fds2.seasonal_adjustment = fds3.seasonal_adjustment
       INNER JOIN covid_canada_monthly ccm1 -- covid19 last month
               ON fds1.geo = ccm1.prname
                  AND Month(fds1.ref_date) = ccm1.cvmonth
       INNER JOIN covid_canada_monthly ccm2  -- covid19 current month
               ON fds2.geo = ccm2.prname
                  AND Month(fds2.ref_date) = ccm2.cvmonth
WHERE  fds1.geo = 'Alberta'
       AND fds1.ref_date >= '2020-01-00 00:00:00'
       AND fds1.ref_date < '2020-08-00 00:00:00'
       AND fds2.ref_date > '2020-01-00 00:00:00'
       AND fds1.naics = 'Total, food services and drinking places'
       AND fds1.seasonal_adjustment = 'Unadjusted'
ORDER  BY 1
```

Query Logic



References

1. A Wilder-Smith, MD, D O Freedman, MD, Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak, Journal of Travel Medicine, Volume 27, Issue 2, March 2020, taaa020, https://doi.org/10.1093/jtm/taaa020

2. Public Health Infobase – Data on COVID-19 in Canada 2020, Public Health Agency of Canada, Government of Canada.

https://open.canada.ca/data/en/dataset/261c32ab-4cfd-4f81-9dea-7b64065690dc

3. Statistic Canada, 2020. Measuring labour market impacts as COVD-19 restrictions gradually ease.

https://www150.statcan.gc.ca/n1/daily-quotidien/200605/dq200605a-eng.html

4. CBC,Sep 2020. Canadian airlines canceling flights again as hoped-for bounceback in demand fizzles.

https://www.cbc.ca/news/business/airlines-flights-cancelled-covid-1.5729643

5. Statistic Canada, 2020. Retail trade sales by province and territory.

https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2010000801

6. OttawaMatter.com, Sep 2020. Restaurant, bar sales rise in July for third straight month, StatCan reports

https://www.ottawamatters.com/national-business/restaurant-bar-sales-rise-in-july-for-third-straight-onth-statcan-reports-2737328

7. Visual Capitalist - Global COVID-19 Containment: Confirmed Cases Updated Daily August 17 2020, Therese Wood

https://www.visualcapitalist.com/global-covid-containment-confirmed-cases-updated-daily/

8. Daily Number of New Reported Cases of COVID-19 by Country Worldwide 2020, European Centre for Disease Prevention and Control (ECDC)

https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

9. Statistic Canada,2020. Monthly Survey of Food Services and Drinking Places (MSFSDP). https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2419

10. Statistic Canada,2020. Actual hours worked at main job by industry, monthly, seasonally adjusted.

https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410003601

11. Statistics Canada, 2020. Non-resident travellers entering Canada, by country of residence, seasonally adjusted.

https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2410000601&pickMembers%5B0%5D=1.3&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=1990&cubeTimeFrame.endMonth=02&cubeTimeFrame.endYear=1990&referencePeriods=19900101%2C19900201

12. Government of Canada,2020. Travel restriction measures: COVID-19 program delivery. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/operational-bulletins-manuals/service-delivery/coronavirus/travel-restrictions.html