Are the Trends of CO_2 Emissions Similar to that of Gas Prices from 1990 to 2020?

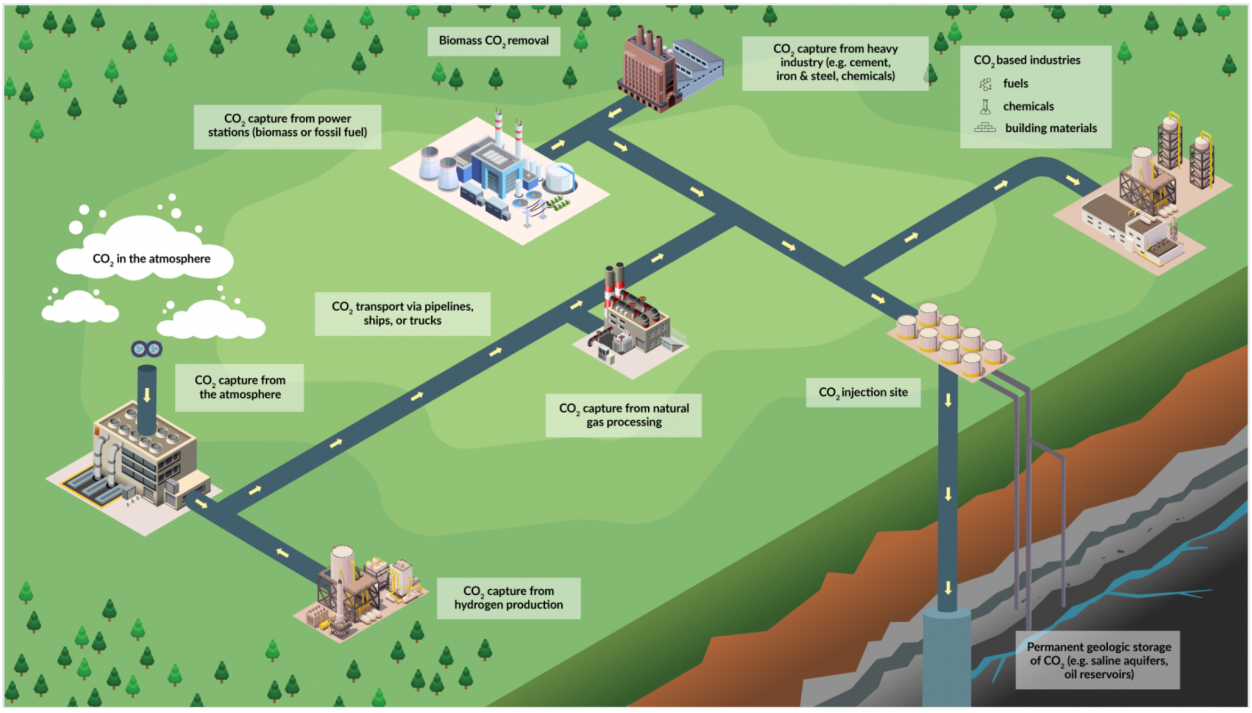
Group Members: Sherry, Grace, and Hiam

November 10th, 2021

Motivation

In 2021, gas prices are on the rise. Additionally, there has been a meeting of **COP26**, the 26th Conference of the Parties to the United Nations Framework Convention on Climate Change. Canada has confirmed that achieving a net-zero economy by 2050 is required to avoid the worst impacts of climate change. This change will impact industries including the gas / oil industry.

Example of future Carbon Dioxide Capture :



Abstract

For this ETL project, we extracted the numbers for CO_2 gas emissions as a result of burning fossil fuels and the gas prices for British Columbia, Alberta, Ontario, and Quebec, inclusively between 1990 and 2020. The data was further manipulated through the use of Pandas to populate clean data sets. Once transformed, the data was loaded onto MongoDB, where we further queried data to create our visualizations.

It was generally observed that the trend of CO_2 gas emissions was closely mirrored to that of the gas prices. Generally, gas prices are affected by the economy, overall supply, consumer demand, and world news. Through timeline analysis, it was confirmed that one can not predict the future prices of gas and emission values.

Project Objective

Hypothesis: Are the trends of Carbon Dioxide (CO_2) Emissions similar to that of Gas Prices from 1990 to 2020 for Quebec? (British Columbia, Alberta, Ontario not analysed at this time)

Questions to answer:

1. What are the average yearly gas prices for each province stated between 1990 and 2020?
2. What are the CO_2 emissions for these years?
3. Do CO_2 emissions have a similar trend compared to gas prices?

To answer these questions, we downloaded CSV files of gas prices, and of CO_2 . Ultimately, we had to map the yearly gas prices with the yearly CO_2 emissions.

Methods

Data Sources Used:

- Gas Prices CSV: Statistics Canada
- CO_2 Emissions CSV: Our World in Data

Final Production Database: MongoDB
Rational: Our data is a time series (Years) which does not require a relational database. MongoDB allows us to retrieve information easily and to add on future data if ever needed by simply adding years. MongoDB is highly optimized compared to a relational database, such as SQL.

Please refer back to ETL-Project in order to get a more in depth explanation of the code and queries done to complete the project.

Code Highlight

EXTRACT

Gas Prices CSV:

We filtered through the provided data on the Statistics Canada website for specific provinces, and a fuel type, as well as the reference time period. The provinces were filtered for; British Columbia, Alberta, Ontario and Quebec, and a timeline of between January 1990 and December 2020. The type of fuel we focused on was Regular unleaded gasoline at self service filling stations.

CO_2 Emissions CSV:

Global data included all countries and many years of CO_2 emissions from the burning of fossil fuels.

Lesson Learned: When we first looked at data for CO_2 emissions, we were looking at the Government of Canada Greenhouse Gas Emissions for all sectors. This was reviewed, and we then found a source (Our World in Data) where it considered only CO_2 emissions as a result of fossil fuel burning.

Step 1: Loading of CSV files and storing it into Dataframe

```
csv_file = "Resources/gas_prices.csv"
orig_gas_data_df = pd.read_csv(csv_file)
```

```
csv2_file = "Resources/gas_emissions.csv"
orig_em_data_df = pd.read_csv(csv2_file)
```

TRANSFORM - Gas Prices CSV

Step 1: Filtered through DataFrame for data needed

```
refined1_gas_data_df = orig_gas_data_df[['REF_DATE', 'GEO', 'VALUE']].copy()
```

Step 2: Cleaning the DataFrame

- *Converting data types where needed and checking for null values*

```
#Converting REF_DATE to datetime
refined1_gas_data_df['REF_DATE'] = pd.to_datetime(refined1_gas_data_df['REF_DATE'])
```

```
#Check result
refined1_gas_data_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3348 entries, 0 to 3347
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   REF_DATE  3348 non-null   datetime64[ns]
 1   GEO       3348 non-null   object
 2   VALUE     3348 non-null   float64
dtypes: datetime64[ns](1), float64(1), object(1)
memory usage: 78.6+ KB
```

- *Split up column to refine data needed*

```
refined1_gas_data_df[['City','Province']]=refined1_gas_data_df['GEO'].str.split(', ', 1,expand=True)
```

Lesson Learned: When splitting a string, an extra space was saved in front of the province name. This had to be corrected when defining the comma, space delimiter. This was observed when we tried to retrieve data from our database, and no results were returned.

- *Extract year value from date column*

```
refined1_gas_data_df['Year'] = pd.DatetimeIndex(refined1_gas_data_df['REF_DATE']).year
```

- *Dropped unnecessary columns and ambiguous data*
- *Renamed columns*
- *Rearrange order of columns*
- *Grouped the data by Year*

	Year	Province	Price
0	1990	Quebec	58.3
1	1990	Quebec	58.9
2	1990	Quebec	59.5
3	1990	Quebec	60.3
4	1990	Quebec	60.4

TRANSFORM - CO_2 Emissions CSV

Step 1: Filter for Canada data between 1990 and 2020 specifically

```
canada_em_data_df = orig_em_data_df.loc[(orig_em_data_df["Entity"]=="Canada")
                                        & (orig_em_data_df["Year"] >= 1990)]
```

Step 2: Cleaning the DataFrame

- *Converted CO_2 emission data type from integer to float*

```
renamed_canada_em_data_df = renamed_canada_em_data_df.astype({"CO2": float})
```

- *Converted CO_2 emission values from Billion Tonnes to Megatonnes*

```
renamed_canada_em_data_df["CO2"]=renamed_canada_em_data_df["CO2"].div(1000000)
```

- *Dropped unnecessary columns and renamed columns*

	Country	Year	CO2
0	Canada	1990	458.007408
1	Canada	1991	449.699173
2	Canada	1992	463.521037
3	Canada	1993	463.993303
4	Canada	1994	478.267071
5	Canada	1995	490.951380

LOAD

Step 1: Load clean DataFrames as CSV (without the pandas index, with the header) locally

Step 2: Load clean CSV files into MongoDB

- *Gas Prices CSV*

```
# The default port used by MongoDB is 27017  (removes any items already there, then inserts new data)
mng_client = pymongo.MongoClient('localhost', 27017)

# Replaces the mongo db name
mng_db = mng_client['CO2']

# Replace mongo db collection name
collection_name = 'gas_prices'
mng_col = mng_db[collection_name]

# Get csv
csv_file_gas_prices = "Data/clean_gas_prices.csv"
data_gp = pd.read_csv(csv_file_gas_prices)
data_gp_json = json.loads(data_gp.to_json(orient='records'))
mng_col.delete_many({})
mng_col.insert_many(data_gp_json)
```

- *CO_2 Emissions CSV*

```
# Replace mongo db collection name (removes any items already there, then inserts new data)
collection_name = 'emission_data'
mng_col2 = mng_db[collection_name]

# Get csv
csv_file_emission = "Data/clean_emission_data.csv"
data_ed = pd.read_csv(csv_file_emission)
data_ged_json = json.loads(data_ed.to_json(orient='records'))
mng_col2.delete_many({})
mng_col2.insert_many(data_ged_json)
```

QUERIES

Please refer back to Data_to_Mongo in order to see the queries done for each province.

Step 1: Separate out collections by province (ex. Below: Ontario)

```
collection = mng_db.gas_prices
ont_col = pd.DataFrame(list(mng_db.gas_prices.find({'Province':'Ontario'})))
```

Step 2: Get the Emissions Data as a DataFrame

```
em = pd.DataFrame(list(mng_db.emission_data.find()))
```

Step 3: Merge each province to the emissions data by Year (ex. Below: Ontario)

```
ont_col_m = ont_col.merge(em, how='inner', on='Year')
```

Step 4: Clean the merged DataFrame by removing unnecessary columns

Step 5: Append all the province DataFrames together

```
appended_data = ont_col_m.append(bc_col_m)
```

Findings

Using the clean appended data, we plotted the gas prices per province from 1990 to 2020.

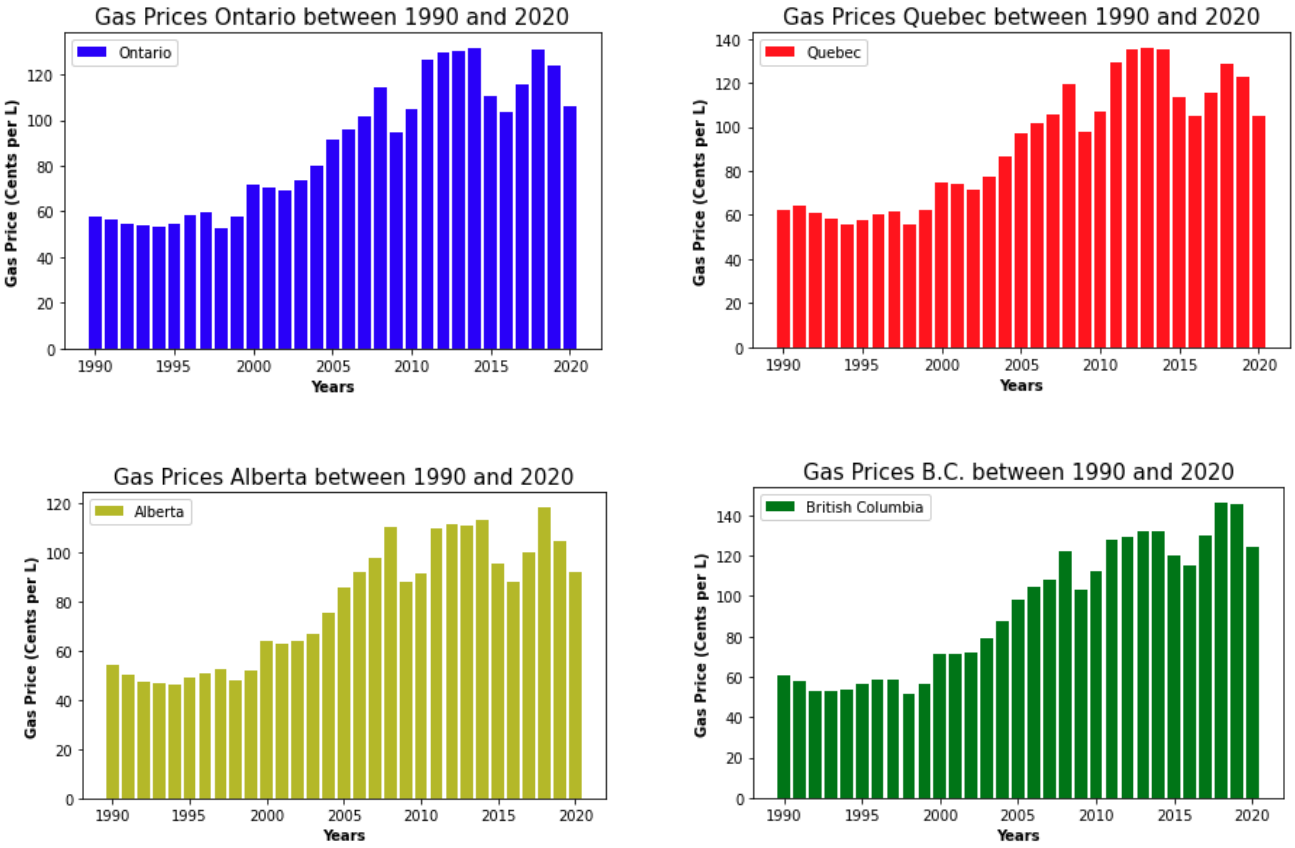


Figure 1.0: Gas Prices in Cents per L for Ontario, Quebec, Alberta, and British Columbia between 1990 and 2020.

Using the clean appended data, we also plotted the CO_2 emissions from 1990 to 2020.

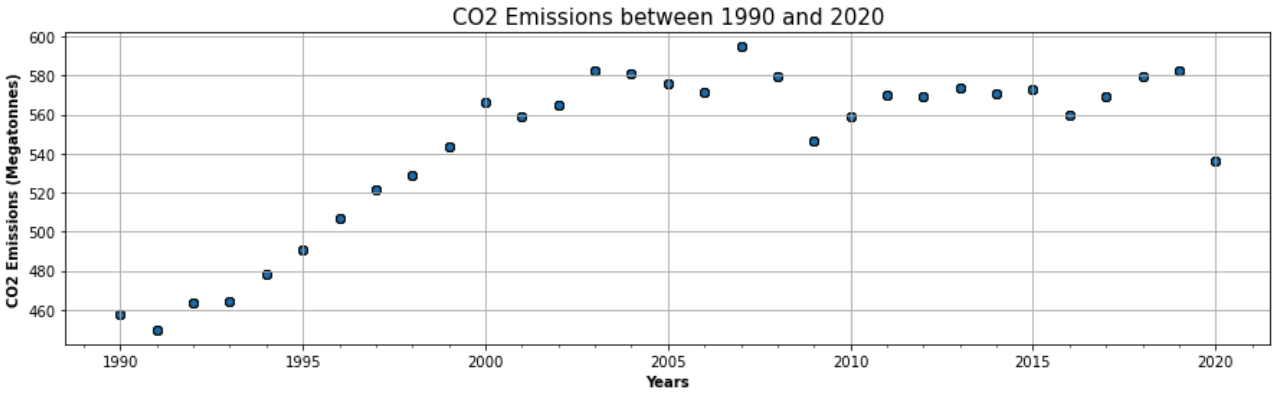


Figure 1.1: CO_2 Emissions between the years 1990 and 2020 for Canada

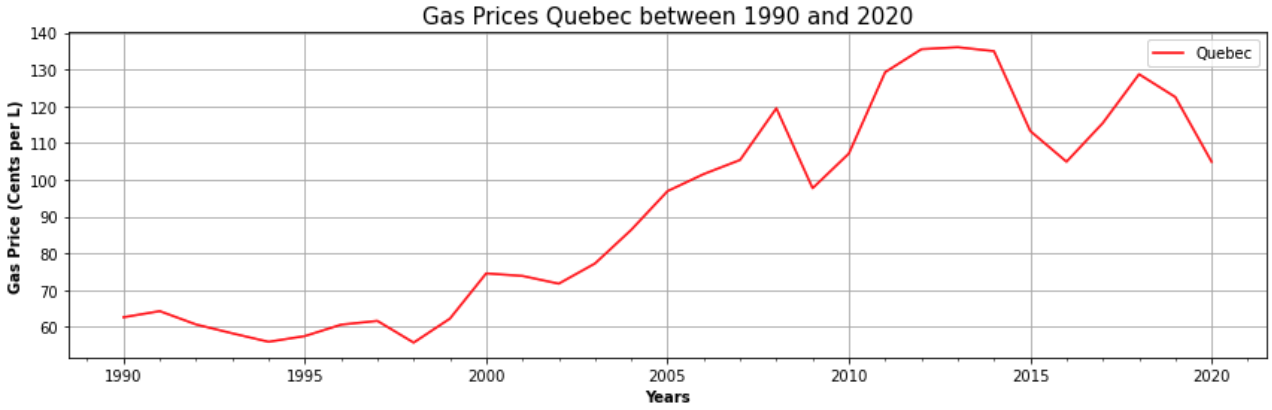


Figure 1.2: Gas Prices Trend for the Province of Quebec from 1990 to 2020

1990-2008 There is a general upward trend of emission and gas prices. The CO_2 emissions increased more rapidly than the gas prices. This trend makes sense, as yearly gas prices would not exponentially increase, but increases in the demand of fossil fuels, would spike a large increase in CO_2 emissions.

1998 Gas prices dipped (prices historically low) There was an adjustment for inflation. Emissions are still trending upward (Generally, more people purchased fuel causing the increase in emissions.)[(1)]

2000-2003 There is a small downward trend of emission and gas prices. During this period, West Texas Intermediate (WTI) prices were stable, generally trading in the area of $30 per barrel. Contributing factors included weak economic growth and oil demand following the events of September 11, 2001.[(2)]

2007-2008 Upward trend in both gas prices and emissions. This period was characterized by extreme price volatility. Major contributing factors included ongoing significant growth in financial investment in oil; a falling U.S. dollar; growing demand from emerging economies; geopolitical

instability; rising finding and development costs; and, slow non-OPEC (Organization of the Petroleum Exporting Countries) production growth.[2]

2008-2009 This sudden decrease of gas prices after the increase from 2007-2008 may have been affected by the global recession that was characterized by broad-based wealth destruction with the collapse of the U.S. housing market, the failure of a number of major financial institutions and the decline in stock markets. Likewise, more people were being economical with fuel consumption, which justifies the decrease in CO_2 emissions.[2]

2009-2010 Here, there is an increase in gas prices, as the global economy slowly started to improve, and CO_2 emissions similarly started to increase as well.[2]

2014-2015 Booming U.S. shale oil production played a significant role in the oil price plunge from mid-2014 to early 2016. The initial drop in oil prices from mid-2014 to early 2015 was primarily driven by supply factors, including booming U.S. oil production, receding geopolitical concerns, and shifting OPEC policies. This partly explains why the oil price plunge failed to provide a subsequent boost to global activity. Generally, the emissions looked to increase slightly, as all of the fuel created would have been purchased.[3]

2015-2016 There was a drop in gas prices and a drop in emissions. However, deteriorating demand prospects played a role , particularly from mid-2015 to early 2016. The dollar was strong. Inventories were huge. The economy was weak. Oil production was growing. Global demand for oil was decreasing. The economies of Europe and developing countries were weakening. Vehicles are becoming more fuel-efficient.[4] The world economy is growing, suggesting a turning point in clean energy development. A hoped for decoupling of economic growth and increased carbon emissions.[5]

2016-2018 Increase in gas prices occured as the gas economy leveled as a result of the previous years. Due to the high demand of gasoline, CO_2 emissions were also on a rise.[6]

2020 Dip in oil prices, and emissions went down. This is the year of the pandemic, prices of gas went down - a lot of people were 'locked down' world-wide. Thus, it makes sense that fewer emissions would follow.[7]

Please refer back to clean_CO2_project_data to see the specific values.

Conclusions

With the above timeline analysis, **one can not predict the future prices of gas and emission values**.[8][9] Generally, as stated above, gas prices were affected by the economy, overall supply, consumer demand, and world news. Currently, it looks like emissions values are similar to the gas prices (use / generation of oil). During 1998 and 2014-2015, there was not a similar trend to gas price and emission value. Gas prices plunged, but emissions went up.

Tellingly, during 2015-2016 one can see a downward trend on gas prices as the economy was becoming more fuel-efficient (note: dollar was strong and economy was weak). Although, after 2015-2016, the demand for oil increased again.

Future Considerations

We hope for a permanent decoupling of economic growth and carbon emissions by 2050. However, some countries are trying to reach Zero emissions by 2030. Perhaps we could analyze trends of this in the future.

References

(1) https://www.latimes.com/archives/la-xpm-1998-dec-01-mn-49558-story.html
(2) https://www.cer-rec.gc.ca/en/data-analysis/energy-markets/archive/canadian-energy-pricing-trends-2011/canadian-energy-pricing-trends-2000-2010-energy-facts.pdf
(3) https://blogs.worldbank.org/developmenttalk/what-triggered-oil-price-plunge-2014-2016-and-why-it-failed-deliver-economic-impetus-eight-charts
(4) https://www.investopedia.com/articles/investing/102215/4-reasons-why-price-crude-oil-dropped.asp
(5) https://insideclimatenews.org/news/07122015/global-carbon-emissions-rising-decades-decline-2015-study-climate-change-paris/
(6) https://www.cer-rec.gc.ca/en/data-analysis/energy-markets/market-snapshots/2018/market-snapshot-canadian-gasoline-prices-rise-highest-level-in-over-3-years.html
(7) https://www.ctvnews.ca/autos/gasoline-sales-plunge-to-lowest-level-in-20-years-during-first-year-of-pandemic-1.5611989
(8) https://www.canadianenergycentre.ca/oil-and-gas-to-lead-us-energy-consumption-to-2050-says-new-forecast/
(9) https://www.cnbc.com/2021/04/15/oil-could-plummet-to-10-by-2050-if-paris-climate-goals-are-achieved.html

Appendix

Year	Province	Gas Price (CAD Cents per L)	CO2 Emissions (Megatonnes)	Gas Price Difference (CAD Cents per L)	CO2 Emissions Difference (Megatonnes)
1990	Quebec	62.61666667	458.007408		
1991	Quebec	64.27916667	449.699173	1.6625	-8.308235
1992	Quebec	60.65833333	463.521037	-3.620833333	13.821864
1993	Quebec	58.24166667	463.993303	-2.416666667	0.472266
1994	Quebec	55.94166667	478.267071	-2.3	14.273768
1995	Quebec	57.49583333	490.95138	1.554166667	12.684309
1996	Quebec	60.60416667	506.99962	3.108333333	16.04824
1997	Quebec	61.62916667	521.278446	1.025	14.278826
1998	Quebec	55.71666667	528.987104	-5.9125	7.708658
1999	Quebec	62.25	543.729999	6.533333333	14.742895
2000	Quebec	74.55	566.552294	12.3	22.822295
2001	Quebec	73.85	559.042165	-0.7	-7.510129
2002	Quebec	71.74583333	564.669174	-2.104166667	5.627009
2003	Quebec	77.24166667	582.343586	5.495833333	17.674412
2004	Quebec	86.39583333	580.758769	9.154166667	-1.584817
2005	Quebec	96.91666667	575.852411	10.52083333	-4.906358
2006	Quebec	101.5666667	571.409763	4.65	-4.442648
2007	Quebec	105.3875	594.665366	3.820833333	23.255603
2008	Quebec	119.4458333	579.367139	14.05833333	-15.298227
2009	Quebec	97.70416667	546.610441	-21.74166667	-32.756698
2010	Quebec	107.15	558.804007	9.445833333	12.193566
2011	Quebec	129.2583333	569.870198	22.10833333	11.066191
2012	Quebec	135.5125	569.262731	6.254166667	-0.607467
2013	Quebec	136.05	573.698737	0.5375	4.436006
2014	Quebec	134.9833333	570.563417	-1.066666667	-3.13532
2015	Quebec	113.2666667	573.061376	-21.71666667	2.497959
2016	Quebec	104.9125	559.566962	-8.354166667	-13.494414
2017	Quebec	115.425	569.360436	10.5125	9.793474
2018	Quebec	128.7083333	579.470372	13.28333333	10.109936
2019	Quebec	122.5041667	582.389608	-6.204166667	2.919236
2020	Quebec	104.9083333	535.82299	-17.59583333	-46.566618