Overview of Power Generation in Canada

DATA 604: Big Data Management

Final Project Report

University of Calgary

Fall 2020

Group No. 9 PM Section

Shahbaz Masih

Yu Nakamura

Shade Oguntoyinbo

Adebola Omosola

Contents

1) Introduction

In the past few decades, with recorded negative environmental effects, there has generally been a growing need for more generation and utilisation of renewable energy as against the usage of fossil fuels, which over time have produced harmful emissions that have affected the quality of air, water and soil, and have been blamed for global warming. Cleaner renewable energy sources have emerged, and its growing generation and utilisation has resulted in the declining reliance on oil, coal and natural gas for electricity production in several countries across the world[1]

Canada, like most developed countries, is invested in seeking out alternative energy sources which produce little to no pollution and has been focused on a successful transition to cleaner environment for future generations. Canada is the sixth largest energy producer, the fourth largest net exporter, and the eighth largest consumer in the world[1]. Several energy sources have been tapped in the generation and utilisation process. One of such is solar power generation and utilization which are capital intensive and require adequate planning to ensure effective implementation. Thus, it is important to determine which province or territory could potentially generate solar power cost-effectively, such that it efficiently provides electricity for homes and industries in the country. Furthermore, wind technology is developing rapidly and has huge potential as a great renewable energy resource following solar energy[2]. Organisations like Pan-Canadian conducted a wind integration study in Canada to assess the operational and economic implication of wind energy into the Canadian electricity system. In addition, the Canada Energy Regulator report[3] provides a projection into the energy supply and demand projections from 2019 to 2040. It reviews all different sources of energy for electricity among other uses. The projection to 2040 is based on a number of assumptions including technology advancement, climate regulations, energy prices, macroeconomics conditions etc. Coal, as a source of energy, will be phased out over the next few years and we would expect to continue to see increased use of cleaner renewable energy sources – wind, solar and hydro.

Around 30% of the total power generated in Canada is from fuel sources e.g., coal, gas, liquid fuels, uranium, wood etc[4]. Fuel cost is the major cost in power generated from fuel sources and is a major factor in the economics of fuel generated power. One of the ways by which economics of fuel generated power can be looked at is cost of fuel per unit of power produced. Also, efficiency of power generation by fuels can be measured by amount of fuel used per unit of power generated. The evolution of technology over time has caused changes in distribution of power generated from different fuel sources. It has also impacted the contribution of each fuel source to the total power generated by using fuels.

This study is therefore aimed at providing an overview of the power generation in Canada, through the review of various energy sources.

Research Questions:

As this dataset offers an overview of power generation in Canada, it will be valuable to answering a number of questions that are documented below and many more that may arise in the course of this project.

1. From the Economics of Solar power generation data, what is the estimated total capacity of energy that can be generated from solar power projects in various provinces and territories in Canada? Which province/territory could potentially generate the most solar energy? Which array generates the most energy and was is the breakeven price?

2. From wind power data, how much energy was producing from each provincial site, and how effectively operating? Is there any seasonal / monthly relationship between wind energy production, wind speed, and electric selling price index?
3. The dataset for thermally generated power by using different fuel types will be investigated to look into contribution of each fuel type into this type of power generation, key economics factors i.e. total fuel cost and fuel cost to produce one Megawatt Hour of power for different fuel types and power generation process efficiency i.e. total quantity of fuel and quantity of fuel required to produce one Megawatt Hour of power for different fuel type across Canada and different provinces and territories from 2005 to 2019
4. From the supply and disposition data, how has provincial and national power needs been met from 2005 to 2019? Based on unallocated power data provided, what is the trend by provinces- surplus or deficit?
5. Review production and capacity at each energy sector level (solar, wind, thermal and others) by province.
    a. How much energy has been produced and what is total capacity for each energy sector and its components?
    b. Any factors to control total productions at each energy sector level?
    c. Can we confirm any difference between provinces related to above?
    d. Which provinces are the most active or less active producer at each energy sector level?

2) Individual Datasets

2-1. Dataset Description

For this project, the datasets used were sourced from the Open Government License, Canada. The Open Government License encourages and permits the public usage of published information with the condition that the source is adequately acknowledged. The datasets are presented at national and provincial levels and are described in more details below:

Set A

Economics of Solar Power in Canada[5] (Adebola O.)

This dataset has estimates of power generation, breakeven and reference prices for solar-power generation projects at different scales and capacities for several communities in Canada. The datasets originated from the Canada Energy Regulator (formerly known as the National Energy Board). Although this dataset was originally published in 2018, they were modified and updated in June 2020. This dataset has 904932 rows of data comprising of 22 columns made up of location, array types, market pricing, utility scale, seasonal generation in megawatt hours (MWh), installation cost scenarios, breakeven prices and reference prices as well as the total capacity of the projects in Megawatts (MW).

The details and definitions of each of the columns are presented as follows:

1. CGNDB ID – is the community's identification code in NRCan's Canadian Geographic Name Database.
2. Geographical Name – Community name
3. Generic term - Type of Community (City, Town or Village)
4. Province-Territory - The province or territory in which the community is located
5. Latitude - Latitude of the community

6. Longitude - Longitude of the community
7. Nearest NSRDB Latitude - The closest latitude in the National Solar Radiation Database
8. Nearest NSRDB Longitude - The closest longitude in the National Solar Radiation Database
9. Array Types – which represents the type of array modeled
    a. Residential: A solar project that would help power a home (5 kW)
    b. Commercial: A solar project that would help power a business (200 kW)
    c. Community: A solar project that would help power a community (200 kW)
    d. Utility fixed: A solar project whose panels are on fixed mounts and would sell power to the grid (50 MW)
    e. Utility tracker: A solar project whose panels are on tracker mounts and would sell power to the grid (50 MW)
10. Market pricing - represents if it was modeled hourly on:
    a. A Flat rate: A fixed rate, which does not vary by time of day or
    b. Time of use: A changing rate that varies as demand rises and falls at different times of the day
11. Utility-scale: This represents whether provincial tariff costs are applied to utility-scale projects or not. It is categorised as Tariff No Tariff
12. Annual Generation (MWh) - The amount of electricity generated in the project's first year
13. Winter Generation (MWh) - The amount of electricity generated during the winter in the project's first year
14. Spring Generation (MWh) - The amount of electricity generated during the spring in the project's first year
15. Summer Generation (MWh) - The amount of electricity generated during the summer in the project's first year
16. Fall Generation (MWh) - The amount of electricity generated during the fall in the project's first year
17. Installation Cost Scenario – represents the time period of the assumed installation costs. It is categorised as
    a. Current: The cost to install solar projects today
    b. Near-future: The cost to install solar projects in the near future as costs fall a modest amount,
    c. Low-cost future: The cost to install solar projects in a low-cost future, where costs fall even more over the long term
18. Breakeven Price ($/MWh); The solar project's breakeven price in dollars per megawatt-hour
19. Reference Price ($/MWh): The $/MWh cost of purchasing electricity from the grid in the case of residential, commercial, and community projects. The price at which electricity currently sells at for utility-scale projects
20. Breakeven Price (¢/kWh): The solar project's breakeven price in cents per kilowatt-hour
21. Reference Price (¢/kWh): The ¢/kWh cost of purchasing electricity from the grid in the case of residential, commercial, and community projects. The price at which electricity currently sells at for utility-scale projects
22. Total Capacity: The capacity of the project in Megawatts (MW)

Set B

Wind Power Generation[6] (Yu N.)

The Pan-Canadian Wind Integration Study assesses the operational and economic implications of wind energy into the electric system in Canada[6]. The dataset consists of two files: Year 2008 annual wind generation at each wind site (4894 rows) including plant capacity, losses, gross energy and so on as well as hourly wind energy profile at each existing site (116 csv files) from 2008 to 2010.

On top of above dataset, monthly wind speed data[7] and electric power selling price index[8] are also combined for data exploration.

Here is the ER diagram of three datasets for further data exploration.

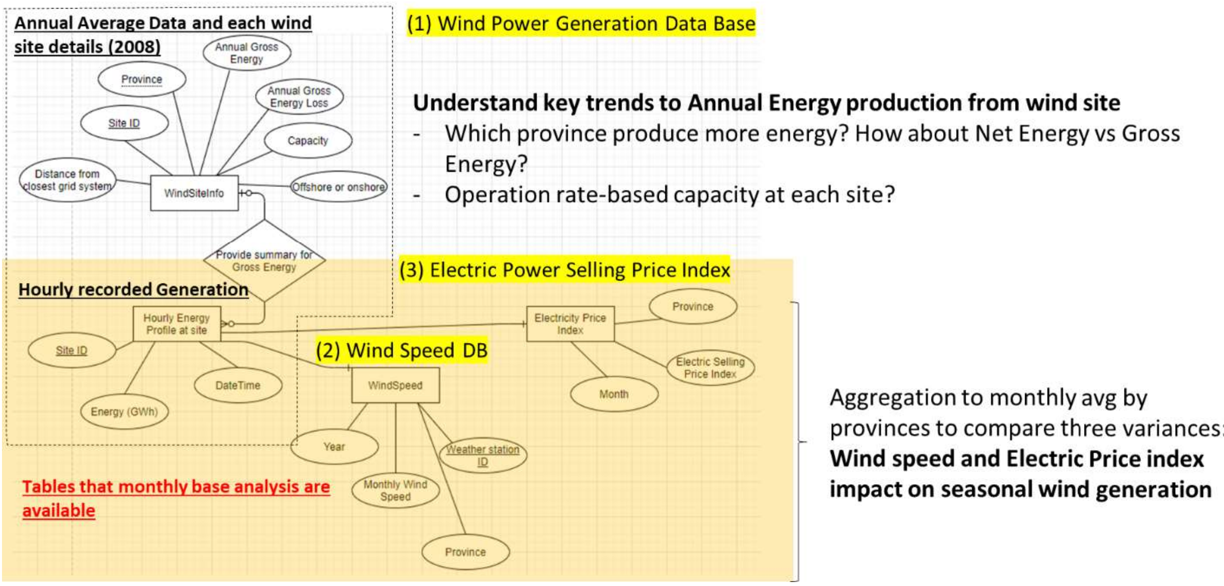

Figure 2.1. ER diagram for wind power generation datasets

Set C (Shahbaz M.)

Electricity from fuels, annual generation by electric utility thermal plants[9]

This data contains information for annual power generated from different fuel sources e.g., gas, coal, liquid fuels, uranium etc. across whole Canada and individual provinces and territories from year 2005 to 2019. The amount of power generated is given in megawatt hours. This data source will be used obtain total power generated from each fuel source in data exploration. The data has 15 columns and 2808 rows.

Electric power generation, annual cost of fuel consumed by electric utility thermal plants[10]

This data contains information for annual cost of fuel consumed for different fuel sources e.g., gas, coal, liquid fuels, uranium etc. for power generation purposes across whole Canada and individual provinces and territories from year 2005 to 2019. The cost of each fuel source is given in thousands of dollars. This data source will be used obtain cost of each fuel source in data exploration for economics analysis. The data has 15 columns and 2657 rows.

Electric power generation, annual fuel consumed by electric utility thermal plants[11]

This data contains information for annual fuel consumed from different fuel sources e.g., gas, coal, liquid fuels, uranium etc. for power generation purposes across whole Canada and individual provinces and territories from year 2005 to 2019. The consumption of fuel is given in different units for different sources e.g., liquid sources have been given in kiloliters, solid sources are given in metric tonnes and gaseous fuels have been given in cubic meters. This data source will be used obtain quantity of each fuel source in data exploration for the analysis of efficiency of power generation. The data has 15 columns and 2699 rows.

The relationships within this dataset and main variables/attributes in each data are given in ER (entity relationship) diagram below. The region (Canada, provinces and territories) and the fuel type are two main/primary keys. The power generated by fuels consumes fuels and fuel and power generated has the cost of fuel attached to it which defines the basic relationship between the tables or 3 data types in this dataset as shown in ER diagram below.

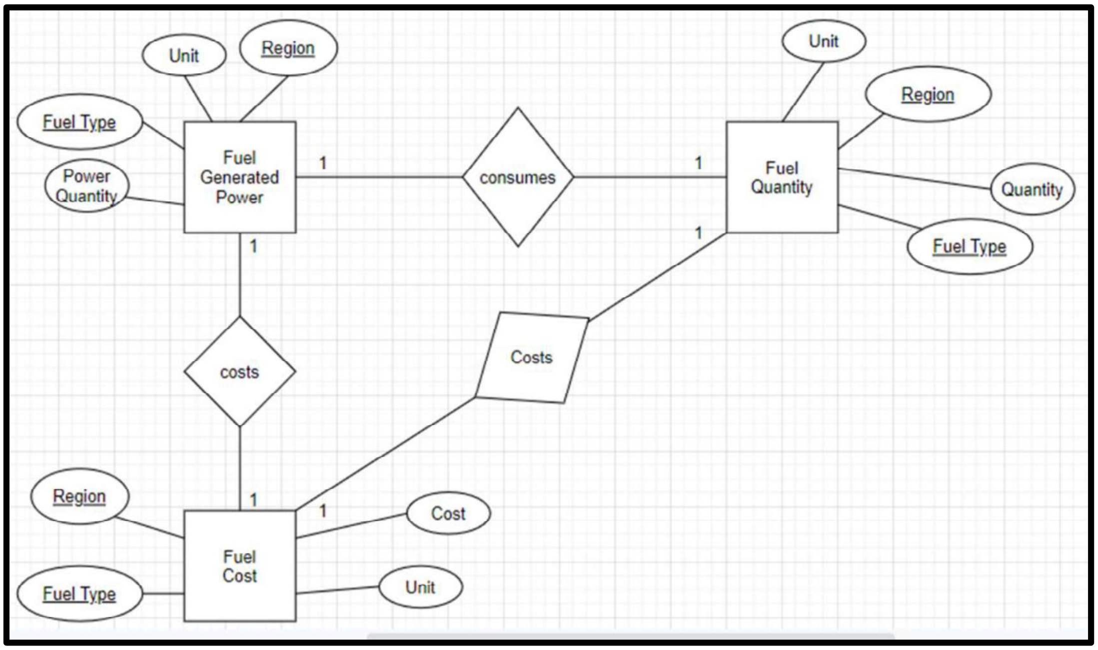

Figure 2.2. ER diagram for fuels generated power

There were no major challenges in dealing with datasets except that regions were changed to the abbreviations (e.g., AB for Alberta) to combine it with team datasets.

Set D (Shade O.)
Electric power, electric utilities and industry, annual supply and disposition[12] (Shade O.)
This dataset has supply and disposition of electric power including generation, imports, exports, sales (to residential, agricultural, mining, manufacturing & other industries) and others. It also details electricity received and delivered to other provinces within Canada. The electricity data shows the electricity quantity in MWh and electricity value in thousand dollars. The dataset has 4938 rows, 16 columns and covers the period 2005 to 2019. A description of the columns in the dataset is below:

1. REF_DATE – shows the year of record from 2005 to 2019
2. GEO - the 7 provinces and 3 territories
3. DGUID - province or territory identification code
4. Electric power, components –
   a) Total generation of electricity - Total utility and industry generation made by hydro, steam (conventional and nuclear), internal combustion, combustion turbine, tidal, wind, solar and other.
   b) Total exports to the United States.
   c) Total imports from the United States.

d) Exports of electricity to the United States, firm - Electricity intended to be available at all times during the period of the agreement of its sale, even under adverse conditions, but subject to force majeure interruptions.
e) Exports of electricity to the United States, secondary - Transactions for short periods and subject to curtailment or cessation of delivery by the supplier or purchaser in accordance with prior agreements or under specific conditions.
f) Interprovincial deliveries of electricity
g) Interprovincial deliveries of electricity, firm - Electricity intended to be available at all times during the period of the agreement of its sale, even under adverse conditions, but subject to force majeure interruptions.
h) Interprovincial deliveries of electricity, secondary - Transactions for short periods and subject to curtailment or cessation of delivery by the supplier or purchaser in accordance with prior agreements or under specific conditions.
i) Interprovincial deliveries of electricity, exchanged - Electricity given in compensation.
j) Total interprovincial receipts of electricity
k) Interprovincial receipts of electricity, purchased
l) Interprovincial receipts of electricity, exchanged - Electricity received in compensation.
m) Total electricity available for use within specific geographic border - This is the sum of generation plus receipts plus imports less deliveries less exports. It is the total supply available for domestic use, and is equivalent to total disposal on a domestic market.
n) Unallocated electricity - calculated field which represents the difference between supply and disposition of electricity.
o) Total sales of electricity to ultimate customers (Residential, Agriculture, Mining, Manufacturing and Other industries)

5. Estimates:
   a) Electricity quantity - Amount of electricity generated, delivered, received, imported or exported
   b) Electricity value - Cost of electricity sold or purchased
6. UOM - unit of measurement i.e., Megawatt hour and Dollars
7. UOM_ID - identification code for unit of measurement
8. SCALAR_FACTOR - scale of value
9. SCALAR_ID - identification code for scalar factor
10. VECTOR - non-unique identifier for geo, electric power components and estimates
11. COORDINATE - non-unique identifier for geo, electric power components and estimates
12. VALUE - amount of electricity quantity and electricity value
13. STATUS - empty field (not relevant to this project)
14. SYMBOL - field is mostly empty (not relevant to this project)
15. DECIMALS - field contains 0 (not relevant to this project)
16. ROW_ID - unique ID for each record

2-2. Individual Data and Queries

As a first step for data exploration, each team member worked on individual dataset to explore the chosen topic. This provided an overview of energy sources in Canada with above datasets.

Set A: Solar Power Generation (Adebola O.)

This dataset provides estimates of solar power generation projects with the aim of determining the generation potential across locations in Canada. It is based on the economics of solar power generation in Canada which provides cost estimates of potential solar power generation projects across all Canadian provinces and territories. It was selected because of its relevance to the analysis of determining which province or territory could potentially generate solar power cost-effectively, such that it efficiently provides electricity for homes and industries. Solar power projects are generally capital-intensive and hence require adequate planning to ensure effective implementation.

Given the size of this dataset, some preliminary clean up and wrangling had to be done outside of the MySQL platform to prepare it for loading and analysis. Some of these include:

- Dropped the columns breakeven and reference prices in c/kwh because it is redundant. It provides the same information as breakeven and reference prices in c/kwh.
- Dropped Longitude and Latitude coordinates as well as the nearest NSRDB longitude and latitude, since they are not relevant to the analysis been carried out
- The CGNDB ID was also dropped because it is not unique enough to serve as the key index. A different key index was generated in replacement so each row of data can be unique.
- The remaining columns were renamed to make the column names more appropriate as follows:

Name in Dataset	**Modified Name after Data Wrangling**
Key – Newly Generated	ID
Geographical Name	G_Name
Generic term - Type of Community	G_Term,
Province-Territory	P_Territory
Array Type	Array_Type
Market Pricing	Market_Pricing
Utility Scale	U_Tariff_Applied,
Annual Generation (MWh)	Annual_G_MWh_First_Yr
Winter Generation (MWh)	Winter_G_MWh_First_Yr
Spring Generation (MWh)	Spring_G_MWh_First_Yr
Summer Generation (MWh)	Summer_G_MWh_First_Yr
Fall Generation (MWh)	Fall_G_MWh_F_Yr
Installation Cost Scenario	Installation_Cost_S
Breakeven Price ($/MWh)	Breakeven_Price_Dol_MWh
Reference Price ($/MWh)	Reference_Price_Dol_MWh
Total Capacity	Total_Capacity_MWs

Table 2.1. Columns from the Solar power dataset

Using this dataset, a number of questions were addressed to provide an overview of solar power projects between different periods in a year as well as across various regions in the Country. These questions were chosen to evaluate the viability of the solar power projects as provided by the capacity, power generation and price estimates. They would help to determine if the projects would be worth investing in for some provinces and territories. These queries promote a better understanding of the dataset and its fits into the broad team project objective of providing an overview of the various energy sources in Canada. Some of these are discussed below:

1. What is the Average Generation (MWh) capacity proposed by Province/Territory and Season?

   This question helps to identify the average solar power generation capacity from the projects per each province and territory during the various seasons of the year. It gives a total overview of the generation capacity in the entire country categorised by locations and seasons.

   As expected, the Summer Season had the highest generation potential of all seasons. As shown in the table below, during the winter season, Quebec had the highest solar power generation, while the highest generation during in the Spring, Summer and Fall season was Nunavut, Saskatchewan and New Brunswick. Saskatchewan had the highest potential solar power generation during the Summer Season despite having a lower number of solar plants (720).

P_Territory	Winter	Spring	Summer	Fall
Alberta	4924.476532231659	11775.487195850088	13727.084615163152	11068.856100642235
British Columbia	4313.1764128570285	10705.018787655988	13177.12012766876	9623.43202755348
Manitoba	5847.277492787272	12563.68566085504	13754.84410581478	11172.24296056727
New Brunswick	7026.851168102762	11602.786769218787	12527.412607425453	11960.780862666836
Newfoundland and Labrador	4128.155067958597	10244.069551625133	11902.688970685573	9460.26199906059
Northwest Territories	4691.154293009922	14349.917130955628	13656.652760576635	7662.3150279436795
Nova Scotia	5857.839543896628	11500.731316390016	12454.318961574498	11724.96436039007
Nunavut	3647.2192698510894	15461.91950069155	12414.816029682755	5644.96321927969
Ontario	6145.176100869006	12312.665866348287	13755.092819394253	11861.249621661913
Prince Edward Island	5297.084691385944	11575.158183717624	12849.027870735565	10856.617489350994
Quebec	6658.624381620456	11773.75096914705	12958.754596301256	11036.087549709744
Saskatchewan	5228.785870879066	12369.676793480594	13817.78206499735	11600.548873521444
Yukon	4030.040457572256	12258.57107930524	12465.467627037138	7031.739564600445

Figure 2.3. Average Generation by season for each province (MWh)

2. What is the total generation capacity by Array Type? and which is the highest or lowest?

   This question helps to identify the total solar power generation capacity from the projects per each array type whether Residential, Commercial, Community, Utility Fixed or Utility Tracker.

Array_Type	Total_GenCapacity
Residential	1292.7599999989693
Community	25855.200000058998
Commercial	25855.200000058998
Utility Fixed	12927600
Utility Tracker	12927600

Figure 2.4. Total Generation Capacity by Array Types

From the figure above, the highest total generation capacity is a tie between Utility Tracker and Fixed while the lowest is Residential. This shows that the array types with the ability to generate the most solar power for the electricity grid are the Utility Fixed and Utility Tracker fixtures which have the capacity to sell about 50MW to the grid. The Residential on the other hand has the lowest potential based on its peculiar nature of been installed on individual property. The residential fixtures would probably generate enough to supply the property but have very little excess left to sell into the electricity grid. In addition, the breakeven prices for the Residential power projects are the highest despite having the lowest total generation capacity.

3. What is the average breakeven price to set up solar power generation and how many plants exist across Canada and by provinces/territories?

This question helps to identify the average prices at which the solar power projects would breakeven and subsequently begin to make profits for each of the array types and across all the provinces and territories. This provides information on the minimum market price that should be charged make sure losses on the projects are reduced. In addition, it provides the total number of solar power sites in each of the provinces.

P_Territory	Solar_Plants_Count	AVG(Breakeven_Price_Dol_MWh)
Alberta	502	72.37384913740077
British Columbia	1061	95.59768264160232
Manitoba	546	83.12226640996779
New Brunswick	428	95.06257552297511
Newfoundland and Labrador	374	91.19601274725986
Northwest Territories	6	76.69277757311625
Nova Scotia	604	94.98595437753828
Nunavut	8	80.85398800032479
Ontario	1284	75.39620135272156
Prince Edward Island	73	92.15511544316132
Quebec	1195	91.51146983366134
Saskatchewan	720	78.4343590283318
Yukon	3	86.06944426279219

Figure 2.5. Project Count and AVG Breakeven price by Province

Ontario has the highest number of solar plants at 1284, followed by Quebec and British Columbia with 1195 and 1061, respectively. The three territories have the lowest number of plants which is understandable given the lower population density in that region. In terms of breakeven prices, Alberta has the lowest breakeven price at an average of $72.37/MWh followed closely by Ontario at $75.40. On the other hand, British Columbia has the most expensive breakeven price at $95.60 followed closely by New Brunswick at $95.06. This shows that it would potential be cheaper to utilize solar power in Alberta than in British Columbia.

4. Is the average season solar power generated consistent for all provinces and territories in Canada? Given that Spring and Summer can be assumed to be the seasons with the generation, what is the difference in the power generated in these two seasons?

   This question presents a comparison between seasonal solar power generation and shows the computation of the difference between the two most assumed seasons to have the highest solar power generation potential. From the output, the two seasons with the highest solar power generation potential are the Summer and Spring Season.

P_Territory	Winter	Spring	Summer	Fall	diff_sum_spr
Alberta	4924.48	11775.49	13727.08	11068.86	1951.60
British Columbia	4313.18	10705.02	13177.12	9623.43	2472.10
Manitoba	5847.28	12563.69	13754.84	11172.24	1191.16
New Brunswick	7026.85	11602.79	12527.41	11960.78	924.63
Newfoundland and Labrador	4128.16	10244.07	11902.69	9460.26	1658.62
Northwest Territories	4691.15	14349.92	13656.65	7662.32	-693.26
Nova Scotia	5857.84	11500.73	12454.32	11724.96	953.59
Nunavut	3647.22	15461.92	12414.82	5644.96	-3047.10
Ontario	6145.18	12312.67	13755.09	11861.25	1442.43
Prince Edward Island	5297.08	11575.16	12849.03	10856.62	1273.87
Quebec	6658.62	11773.75	12958.75	11036.09	1185.00
Saskatchewan	5228.79	12369.68	13817.78	11600.55	1448.11
Yukon	4030.04	12258.57	12465.47	7031.74	206.90

Figure 2.6. Total Generation by season and province

As expected, Season with the highest solar power generation is the SUMMER for most of the provinces/territories except in Nunavut and Northwest territories where SPRING had the highest generation. This is shown by the result of the difference computation between the summer and spring generation for all the provinces/territories.

Set B: Wind Power Generation (Yu N.)

This dataset was selected for the following reasons:

- Integrated Study issued by Pan Canadian provides the detail information of each wind site. It is well-summarized for the usage to explore fundamental knowledge of wind power generation as of 2008.
- On top of that, there is hourly records at existing plants from 2008 January to 2010 January which provide seasonal trends such as seasonal impacts by the weather and electric selling price index.

Since above datasets have complicated structures, this presents a couple of challenges for data wrangling outside of MySQL environment. See the individual queries report for the details.

- Original data for hourly and monthly records required format changes.

- The more than 150 csv and txt files were merged to one file with R.
- The formats and columns names were re-arranged to the suitable format for the MySQL with R.

With usage of above three data sets (wind site info details, wind production profile, wind speed, electricity selling price index), I addressed a series of questions to capture the position of wind energy in Canada and checked any seasonal impacts to the wind power generation.

1) Which province produced more energy and controlled operating effectively?

   Electricity from wind energy is one of cleanest sources in the world. In 2008, Canadian generation of wind power was 34.1 TWh, and total capacity of the wind sites was 10,773 MW. As of 2008, no wind power plants were recorded in Yukon, North West Territory, and Newfoundland and Labrador.
   Below Table 1 is the review of wind energy in 2008 including total site counts, capacity (MW), Total Wind production (GWh) and Site Operating rate (%) by province.

   Where wind production and operating rate are calculated by:

   - $Wind\ Production\ [GWh] = Gross\ Production\ [GWh] - Electric\ Loss\ [GWh] - Maintenance\ Loss\ [GWh]$
   - $Operating\ Rate\ [\%] = \frac{Total\ Wind\ Production\ [GWh] \times 1000}{(365*24) \times TotalCapacity\ [MW]} \times 100$

GEO	TotalCapacity_MW	TotalProduction_GWh	AvgOperatingRate	TotalSiteNumber
ON	4103	13647.50	0.37	33
QC	2964	9077.64	0.35	28
AB	1439	4549.24	0.36	20
NS	391	1260.75	0.37	10
SK	449	1474.56	0.37	7
PE	201	685.89	0.39	6
NB	484	1476.91	0.36	5
BC	685	1764.28	0.29	5
MB	258	862.44	0.38	2

Figure 2.7. Provincial summary of site capacity, generation, and operating rate

From figure 2.7., Ontario had the most wind energy capacity with 4103 MW of power, followed by Quebec with 2964 MW and Alberta with 1439 MW. Total Production and Capacity is well-correlated which shows strong wind power producer in Canada had bigger capacity.

This capacity by province can be seen from the existing wind site (117 sites) location map below in 2008. Almost all sites in Canada are located in the east coast and US-Canada boundary, while fewer sites were distributed in west coast or inland provinces. (Fig. 2.8.).

Figure 2.8. Existing Plant Location
Referred from https://canwea.ca/wp-content/uploads/2016/10/pcwis-overviewpresentation-web.pdf

On the other hand, inspective of provincial total capacity, operating rate for each province is converged on around 34 to 37%, expect for BC that shows only 29% of operating rate. Equal amount of operating rate between provinces show that the technology level of wind power generation development didn't vary across Canada.

2) Where were existing plants located – offshore or onshore, distance from nearest transmission lines?

Most common plant location in Canada is Onshore with 113 sites vs only 3 located offshore (Figure 2.9.).

55 sites representing 47% of all Canadian sites had transmission lines within 10 km. 27% of sites are located within 10 km to 25km from transmission lines while only 16% of sites are located greater than 100 km from transmission lines (Figure 2.10.). This observation makes sense from the perspective of convenience to convey the power to buyer via closest transmission systems.

Onshore_or_Offshore	TotalSiteNumber	AvgOperatingRate	TotalNetEnergy_GWh
Onshore	113	0.36	34454.66
Offshore - Atlantic	3	0.40	344.55

Figure 2.9. Wind site operation type – onshore or offshore

DistToHighVoltTransmission	TotalSiteNumber	AvgOperatingRate_percent	TotalProduction_GWh
10 km	55	0.36	19117.83
25 km	32	0.36	9487.50
100 km	15	0.36	3250.98
50 km	10	0.35	2416.20
200 km	1	0.34	330.31
150 km	3	0.40	196.38

Figure 2.10. Wind site location based on distance of wind site to nearest existing high voltage (>=230kV) transmission lines

3) Seasonal production trend: Impact of weather

To address this question, I calculated monthly wind speed at each weather station record and monthly production from hourly recorded profile data at each wind site. I used a monthly value of year 2008 for consistency of the wind power generation summary data described in above.

In terms of monthly production, overall, highest production is confirmed in January, while less production tends to be seen during summer. Manitoba has a unique trend showing least production from December not summer.

This production pattern captures the seasonal weather trend that wind speed tends to increase during winter and decrease during summer.

month(Date)	GEO	TotalProduction_GWh	WindSpeed_mph
1	QC	1041.57	15.99
10	QC	1037.14	14.01
11	QC	968.71	16.05
3	QC	944.66	17.38
2	QC	848.54	15.52
12	QC	834.56	16.88
4	QC	774.17	16.59
7	QC	765.13	12.99
9	QC	721.88	13.31
8	QC	674.12	12.79
6	QC	624.45	13.99
5	QC	617.85	14.55

Figure 2.11. Monthly Production and wind speed of Quebec

Figure 2.12. summarizes the provincial total wind power production and wind speed. There is no clear trend of the weather impact due to geography to wind power generation.

GEO	JanProduction_GWh	WindSpeed_mph
ON	1784.27	15.01
QC	1041.57	15.99
AB	599.28	12.43
SK	184.48	15.79
BC	181.89	10.35
NB	163.96	16.32
NS	131.32	19.74
MB	98.31	15.80
PE	74.87	18.75

Figure 2.12. Provincial weather effect to the wind power production in January 2008

4) Seasonal production trends: impact of the electric price index

Combined Monthly Electric Selling price index to the monthly production data, I would like to discuss both two relationships in this section.

Here is the example of monthly comparison table for production vs. Price Index. In Ontario, lowest price index is 74.9 CAD in May, and highest one is 105.3 CAD in June and July. Price Index varies based on season, but it's hard to tell my expected relationship which wind energy produced more in the month of high selling price index. This was confirmed in the data of other provinces as well.

month(Date)	GEO	Energy_GWh	PriceIndex
1	ON	1784.27	80.9
11	ON	1472.38	95
10	ON	1461.67	87.4
12	ON	1414.99	90.2
3	ON	1405.31	98.9
4	ON	1276.84	91.1
5	ON	1262.59	74.9
6	ON	1125.63	105.3
2	ON	1112.30	94.6
7	ON	973.52	105.3
8	ON	769.24	90.8
9	ON	754.41	93.5

Figure 2.13. Monthly Comparison between Price Index and Production

On the other hands, looking at the provincial comparison between price index and total production, we definitely see that provinces with high selling price index such as Ontario and Quebec produced more wind energy annually. But it is still not clear one to one trend since Alberta who has the lowest electric prince index still had produced a lot of wind power energy.

GEO	TotalProduction_GWh	PriceIndex
ON	14813.14	92.33
QC	9852.77	93.15
AB	4937.79	61.06
SK	1600.38	79.60
BC	1914.99	64.63
NB	1603.07	93.82
NS	1368.35	73.60
MB	935.97	79.80

Figure 2.14. Provincial price index effect to wind power production

Set C: Thermal Plant Power Generation (Shahbaz M.)

The data has been explained with ER diagram in Data Description section with ER diagram. From the available data, it can be concluded that queries can be woven around quantity of power generated by using different fuels with associated cost of fuel and fuel quantity to get an overview of power generated from fuels by using thermal plants. Fuel type was chosen to be investigated individually and location was explored in group queries. The queries and their results discussion are given below.

1) What is total power generated by using each fuel type by thermal plants since 2005 to 2019 across Canada, provinces and territories?

The results of this query have been shown in the table below. The results have been sorted by Region and Fuel type. Different regions have shown different fuel types to be their main source of fuel generated power based on quantity of power generated in Megawatt hours. For example, in Alberta, Manitoba and Nova Scotia coal has been the main fuel for fuel generated power based on quantity of power generated. While in Canada, Quebec and Ontario, Uranium has been the main source of fuel generated power based on quantity of fuel generated power.

Region	Fuel_Type	Total_Power_By_FuelType	Unit
Alberta	Total liquids	161569	Megawatt hours
Alberta	Total petroleum products	161056	Megawatt hours
Alberta	Diesel	161016	Megawatt hours
Alberta	Petroleum coke	96491	Megawatt hours
Alberta	Propane	513	Megawatt hours
Alberta	Light fuel oil	40	Megawatt hours
Alberta	Other liquid fuels	0	Megawatt hours
Alberta	Other solid fuels	0	Megawatt hours
British Columbia	Total gas	27990246	Megawatt hours
British Columbia	Natural gas	21473172	Megawatt hours
British Columbia	Total solids	6645120	Megawatt hours
British Columbia	Wood	6644927	Megawatt hours
British Columbia	Other gaseous fuels	5990868	Megawatt hours
British Columbia	Steam from waste heat	2991066	Megawatt hours
British Columbia	Total petroleum products	920683	Megawatt hours
British Columbia	Diesel	920683	Megawatt hours

Figure 2.15. Quantity of power generated based on fuel type

2) What is the total fuel cost for each fuel type used for power generation since 2005 to 2019 across Canada provinces and territories?

   The results of this query have been shown in the table below. The results have been sorted by Region and Fuel type. Different regions have shown different fuel types to be their highest cost fuels to generate thermal power. For example, in Alberta, Canada and BC, gas has cost the most among all fuel types from 2005 to 2019 while in Nova Scotia, coal as cost the most among all fuel types based on total money spend on fuels generated power.

Region	Fuel_Type	Total_Cost_By_FuelType	Unit
British Columbia	Total gas	1330868000	Dollars
British Columbia	Natural gas	1157546000	Dollars
British Columbia	Total liquids	253274000	Dollars
British Columbia	Total petroleum products	253274000	Dollars
British Columbia	Diesel	253274000	Dollars
British Columbia	Other gaseous fuels	167824000	Dollars
British Columbia	Total solids	96255000	Dollars
British Columbia	Wood	96255000	Dollars
British Columbia	Methane	5498000	Dollars
British Columbia	Other liquid fuels	0	Dollars
British Columbia	Other solid fuels	0	Dollars
Canada	Total gas	23484894000	Dollars
Canada	Natural gas	23275291000	Dollars
Canada	Total solids	19737526000	Dollars
Canada	Total coal	17975517000	Dollars

Figure 2.16. Cost of fuel to generate power based on fuel type

3) How much fuel was consumed for each fuel type by thermal plants since 2005 to 2019 across Canada, provinces and territories?
The results of this query have been shown in the table below. The results have been sorted by Region and Fuel type. Different regions have shown different fuel types to be their most consumed fuels to generate thermal power based on quantity of fuel used. For example, in Newfoundland and Labrador, liquids were ranked as the most consumed fuel type but in Northwest Territories petroleum products were the most fuel consumed to generate thermal power based on quantity of fuel used.

Region	Fuel_Type	Total_Fuel_Consumed	Unit
New Brunswick	Other solid fuels	0	Metric tonnes
Newfoundland and Labrador	Total liquids	4971849	Kilolitres
Newfoundland and Labrador	Total petroleum products	4971849	Kilolitres
Newfoundland and Labrador	Total heavy fuel oil	4585306	Kilolitres
Newfoundland and Labrador	Imported heavy fuel oil	3720259	Kilolitres
Newfoundland and Labrador	Canadian heavy fuel oil	865047	Kilolitres
Newfoundland and Labrador	Light fuel oil	336944	Kilolitres
Newfoundland and Labrador	Diesel	50345	Kilolitres
Newfoundland and Labrador	Total solids	0	Metric tonnes
Newfoundland and Labrador	Wood	0	Metric tonnes
Northwest Territories	Total petroleum products	359881	Kilolitres
Northwest Territories	Diesel	359881	Kilolitres
Northwest Territories	Total liquids	359881	Kilolitres
Northwest Territories	Natural gas	120397	Cubic metres
Northwest Territories	Total gas	120397	Cubic metres

Figure 2.17. Quantity of fuel consumed to generate power based on fuel type

4) What is the fuel cost per unit power generated by thermal plants since 2005 to 2019 across Canada and provinces for each fuel type? (which is a good economic indicator for power generation)

Fuel cost data and quantity of power generated data were joined together to answer this query and results have been shown in the table below.

REGION	Fuel_Type	Total_Cost_By_FuelType	Total_Power_By_FuelType	FuelCost_Per_Unit Power_Dollars_Per_MWhour
New Brunswick	Canadian bituminous coal	69352000	1331174	52.1
Nova Scotia	Canadian bituminous coal	424049000	16306764	26
Alberta	Canadian bituminous coal	171849000	7732060	22.23
Canada	Canadian bituminous coal	665249000	25369998	26.22
Ontario	Canadian heavy fuel oil	164892000	1005833	163.94
Prince Edward Island	Canadian heavy fuel oil	8597000	29489	291.53
Saskatchewan	Canadian heavy fuel oil	876000	11488	76.25
Quebec	Canadian heavy fuel oil	397873000	3343031	119.02
Nova Scotia	Canadian heavy fuel oil	239025000	3660461	65.3
New Brunswick	Canadian heavy fuel oil	1419615000	19484855	72.86
Canada	Canadian heavy fuel oil	2482062000	30855370	80.44
Newfoundland and Labrador	Canadian heavy fuel oil	251185000	3320213	75.65
Alberta	Canadian subbituminous coal	6604550000	580233741	11.38
Canada	Canadian subbituminous coal	6604550000	580233741	11.38
Canada	Coke oven gas	0	0	NULL

Figure 2.18. Fuel cost per unit of power generated based on fuel type

The results show that different fuel types have different cost of fuel per unit of power generated in different regions. For example, fuel coast for power generated by using Canadian bituminous coal per unit of power ranged from 22.23 to 52.1 dollars per Megawatt hours of power generated.

5) Where in Canada is each fuel cost cheapest/lowest per unit of power produced?
Fuel cost data and quantity of power generated data were joined together to answer this query and filtered for the lowest cost of fuel for each fuel type. The results have been shown in the table below.

Fuel_Type	REGION	Lowest FuelCost_Per_UnitPower_Dollars_Per_Mwhour
Canadian bituminous coal	Alberta	22.23
Canadian heavy fuel oil	New Brunswick	65.3
Canadian subbituminous coal	Alberta	11.38
Diesel	Alberta	112.85
Imported bituminous coal	New Brunswick	28.59
Imported heavy fuel oil	New Brunswick	69.52
Imported subbituminous coal	Manitoba	23.31
Light fuel oil	Alberta	50
Lignite	Ontario	16.89
Methane	Alberta	0
Natural gas	Alberta	18.84
Other gaseous fuels	Alberta	2.87
Other liquid fuels	Alberta	NULL
Other solid fuels	Alberta	0
Petroleum coke	Alberta	7.59

Figure 2.19. The lowest fuel cost per unit of power generated based on fuel type

The results show that different fuel types have lowest costs of fuel per unit of power generated in different regions. For example, the lowest fuel coast for power generated by using Canadian bituminous coal per unit of power was 22.23 dollars per MW Hour in Alberta and the lowest fuel coast for power generated by using Canadian heavy fuel oil per unit of power was 65.3 dollars per MW Hour in New Brunswick.

6) What is the total amount of fuel used per unit of power generated by thermal plants since 2005 to 2019 across Canada, provinces and territories for each fuel type? (which is a good indicator of how efficiently fuel was utilized for power generation and is related to economics of power generation)

   To answer this query, quantity of power generated table and fuel consumption table were joined together. The results are shown in the table below. The results indicated that amount to fuel consumed to generate unit power differed from one region to another for the same fuel type. For example, total coal consumed to generate one unit of power in Alberta was 0.59 units but in Manitoba it was 0.60 units.

REGION	Fuel_Type	Total_Fuel_Consumed_ByFuelType	Total_Power_By_FuelType	FuelConsumed_Per_UnitPower_UnitsFuel_Per_MWhour
Alberta	Total solids	357039922	594439112	0.60
Alberta	Total coal	348217540	587965801	0.59
Alberta	Total subbituminous coal	344198961	580233741	0.59
Alberta	Canadian subbituminous coal	344198961	580233741	0.59
Alberta	Total gas	62208014	210020448	0.30
Alberta	Natural gas	54918494	208861602	0.26
Alberta	Canadian bituminous coal	4018579	7732060	0.52
Alberta	Total bituminous coal	4018579	7732060	0.52
Alberta	Wood	8794218	6376820	1.38
Alberta	Other gaseous fuels	7111578	709839	10.02
Alberta	Methane	177942	449007	0.40
Alberta	Total liquids	55465	161569	0.34
Alberta	Total petroleum products	55163	161056	0.34
Alberta	Diesel	55153	161016	0.34
Alberta	Petroleum coke	28164	96491	0.29

Figure 2.20. Fuel consumed per unit of power generated

7) Where in Canada is the least amount of fuel used per unit of power generated for each fuel type? (which is an indication of power generation process efficiency and is attached to economics of power generation as well)

   To answer this query, the same tables were used as in query 6 and results in query 6 were filtered for least amount of fuel consumed for each fuel type to generate one unit of power. The results have shown in table below. The results showed that most of the fuel types were most efficiently used (least amount of fuel used to generate unit power) in Alberta.

Fuel_Type	REGION	Least FuelConsumed_Per_UnitPower_UnitsFuel_Per_Mwhour
Canadian bituminous coal	Alberta	0.41
Canadian heavy fuel oil	New Brunswick	0.25
Canadian subbituminous coal	Alberta	0.59
Diesel	Alberta	0.17
Imported bituminous coal	New Brunswick	0.37
Imported heavy fuel oil	New Brunswick	0.24
Imported subbituminous coal	Manitoba	0.38
Light fuel oil	Alberta	0.25
Lignite	Ontario	0.75
Methane	Alberta	0
Natural gas	Alberta	0.22
Other gaseous fuels	Alberta	0.35
Other liquid fuels	Alberta	NULL
Other solid fuels	Alberta	0.41
Petroleum coke	Alberta	0.29
Propane	Alberta	0.59
Total bituminous coal	Alberta	0.37
Total coal	Alberta	0.38
Total gas	Alberta	0.22

Figure 2.21. Least amount of fuel used to generate unit power for each fuel type

8) Did the most efficient use of fuel to produce power lead to the least cost of fuel to generate power or there are some other factors involved e.g., cost of fuel per unit can vary from region to region?

In other words, this query can be stated as "are the regions where least fuel was consumed to generate the unit power the same as the regions of the least fuel cost to produce unit power?" To answer this query, results in queries 6 and 7 were joined together. The results are given in the table below. The results show that efficient use of fuel did lead to least cost of fuel consumed and the impact of other parameters like variation of cost from one region to the next and process efficiencies was negligible. Alberta had the most efficient use of fuel as shown in query 6 for numerous fuel types. The table below shows Alberta for a lot of categories of fuels and asserts the derived conclusion.

Fuel_Type	REGION	FuelConsumed_Per_UnitPower_UnitsFuel_Per_Mwhour	FuelCost_Per_UnitPower_Dollars_Per_Mwhour
Canadian bituminous coal	Alberta	0.41	22.23
Canadian heavy fuel oil	New Brunswick	0.25	65.3
Canadian subbituminous coal	Alberta	0.59	11.38
Diesel	Alberta	0.17	112.85
Imported bituminous coal	New Brunswick	0.37	28.59
Imported heavy fuel oil	New Brunswick	0.24	69.52
Imported subbituminous coal	Manitoba	0.38	23.31
Light fuel oil	Alberta	0.25	50
Lignite	Ontario	0.75	16.89
Methane	Alberta	0	0
Natural gas	Alberta	0.22	18.84
Other gaseous fuels	Alberta	0.35	2.87
Other liquid fuels	Alberta	NULL	NULL
Other solid fuels	Alberta	0.41	0
Petroleum coke	Alberta	0.29	7.59
Propane	Alberta	0.59	376.22
Total bituminous coal	Alberta	0.37	22.23
Total coal	Alberta	0.38	11.53
Total gas	Alberta	0.22	18.84

Figure 2.22. Correlation between fuel consumption & cost of fuel per unit of power generated

Set D: Electricity Data - Power, Utilities, Annual supply, and Deposition (Shade O.)

I selected this dataset to educate me on the different sources of electricity in Canada by province and territories. Also, because it had total generation details which would fit nicely with the team's datasets.

This dataset was easy to work with for the most part. The major challenge was separating the electricity value from electricity quantity as this was lumped together in a column. Reshaping the table rendered some columns redundant and they were deleted.

This dataset was fairly clean and ready for use. Cleaning and preparing the dataset occurred in excel and MySQL as detailed below:

Excel:

- In reviewing the data, none of the 16 columns in the dataset had unique identifier field. A new field called "ROW_ID" was created from a combination of records from four different columns (YEAR-PROVINCE-ELECTRIC_POWER_COMPONENTS-ESTIMATES). It was necessary to create this column to achieve uniqueness of each record.
- Another field called YR_GEO_COMP was created to serve as the field on which a join will be performed.
- Deleted column "TERMINATED"

MySQL:

- Created load and create sql files for loading the csv file into MySQL workbench
- Dropped five (5) redundant columns (STATUS, SYMBOL, DECIMALS, VECTOR and COORDINATE) the ss_pow_gen table so that the rows about estimates became columns (Electricity_qty_MWh and Electricity_value_$000).
- Reshaped the table so that the rows about 'estimates' become columns (Electricity_qty_MWh and Electricity_value_$000). This also renders the following columns redundant: UOM, UOM_ID, SCALAR_FACTOR and SCALAR_ID. The unit of measurement and scalar factor are factored into the new name for the fields. A view was created from this reshaped table

REF_DATE	GEO	Electric_power_componei	YR_GEO_COMP	ROW_ID	Electricity_qty_MWh	Electricity_value_$000
2005	Alberta	Total sales of electricit...	2005-Alberta-Total sales ...	2005-Alberta-Total sale...	50974521	4297659
2005	British Columbia	Total sales of electricit...	2005-British Columbia-To...	2005-British Columbia-T...	49995707	2701511
2005	Canada	Total sales of electricit...	2005-Canada-Total sales...	2005-Canada-Total sale...	487699192	34377200
2005	Manitoba	Total sales of electricit...	2005-Manitoba-Total sal...	2005-Manitoba-Total sal...	19928571	968508
2005	New Brunswick	Total sales of electricit...	2005-New Brunswick-Tot...	2005-New Brunswick-To...	14246567	1098337
2005	Newfoundland and Labrador	Total sales of electricit...	2005-Newfoundland and ...	2005-Newfoundland an...	9269907	578681
2005	Northwest Territories	Total sales of electricit...	2005-Northwest Territori...	2005-Northwest Territor...	316626	77491
2005	Nova Scotia	Total sales of electricit...	2005-Nova Scotia-Total s...	2005-Nova Scotia-Total ...	11502591	952408
2005	Nunavut	Total sales of electricit...	2005-Nunavut-Total sale...	2005-Nunavut-Total sal...	134897	60539
2005	Ontario	Total sales of electricit...	2005-Ontario-Total sales...	2005-Ontario-Total sale...	139402992	13005470
2005	Prince Edward Island	Total sales of electricit...	2005-Prince Edward Isla...	2005-Prince Edward Isla...	1093986	127838
2005	Quebec	Total sales of electricit...	2005-Quebec-Total sales...	2005-Quebec-Total sale...	173444783	9244440

Figure 2.23. Reshaped table showing all the electricity quantity and values as separate columns

From the wrangle data, the following queries were created to address the research question:

1. Show cumulative unallocated electricity by province from 2005 - 2019:

   Every province and territory track unallocated electricity which could be a surplus or deficit. Where this is a deficit, it means the province or territory was unable to meet their electricity need and had to import from the United States or receive from other provinces. The first query shows cumulative unallocated electricity from 2005 - 2019 while the second one shows the same variable from 2013 to 2019.

   The cumulative unallocated electricity for all provinces for the past 15 years is surplus while in the last 6 years, 3 provinces (AB, NB & PEI) have been in deficit positions (-3497430MWh, -824838MWh & 184464MWh respectively). Again, this is useful for planning.

GEO	Electric_power_components	Electricity_qty_MWh
Nunavut	Unallocated electricity	38430
Prince Edward Island	Unallocated electricity	378496
Northwest Territories	Unallocated electricity	528634
Yukon	Unallocated electricity	585908
New Brunswick	Unallocated electricity	3356278
Nova Scotia	Unallocated electricity	7201845
British Columbia	Unallocated electricity	10450342
Newfoundland and Labrador	Unallocated electricity	19043364
Saskatchewan	Unallocated electricity	29413526
Manitoba	Unallocated electricity	41079426
Alberta	Unallocated electricity	91307060
Ontario	Unallocated electricity	139821433
Quebec	Unallocated electricity	177616223

Figure 2.24. Cumulative unallocated electricity by province/territory from 2005 – 2019

GEO	Electric_power_components	Electricity_qty_MWh
Prince Edward Island	Unallocated electricity	-184464
New Brunswick	Unallocated electricity	-824838
Alberta	Unallocated electricity	-3497430

Figure 2.25. Cumulative unallocated electricity by province/territory from 2013 – 2019

2. Total Electricity Generation by Provinces per year from 2005 - 2019

REF_DATE	GEO	Electric_power_components	VALUE
2005	Alberta	Total generation of electricity	63738176
2005	British Columbia	Total generation of electricity	67773670
2005	Manitoba	Total generation of electricity	37049409
2005	New Brunswick	Total generation of electricity	21063054
2005	Newfoundland and Labrador	Total generation of electricity	42136351
2005	Northwest Territories	Total generation of electricity	533580
2005	Nova Scotia	Total generation of electricity	12476923
2005	Nunavut	Total generation of electricity	141645
2005	Ontario	Total generation of electricity	158750451
2005	Prince Edward Island	Total generation of electricity	46417
2005	Quebec	Total generation of electricity	180296084
2005	Saskatchewan	Total generation of electricity	20020456

Figure 2.26. Total electricity generated by Provinces

The result of this query outputs the total electricity generation by provinces and territories. National sum has been excluded in this query. This will be most useful for our group exploration as we explore the renewable energy sources (wind, solar etc.). We will look at what percentage of total generation come from these renewable sources and how this has changed over the years. We will also see how other sources of energy especially fossil fuels continue to reduce over time. The changes by province over the years has not been significant.

3. Total sales of electricity to ultimate customers by province/territory in 2019

REF_DATE	GEO	Electric_power_components	Electricity_qty_MWh	Electricity_value_$000
2019	Nunavut	Total sales of electricity to ultimate customers	177259	130741
2019	Northwest Territories	Total sales of electricity to ultimate customers	304035	120366
2019	Yukon	Total sales of electricity to ultimate customers	423121	62533
2019	Prince Edward Island	Total sales of electricity to ultimate customers	1420448	224106
2019	Newfoundland and Labrador	Total sales of electricity to ultimate customers	9139357	814159
2019	Nova Scotia	Total sales of electricity to ultimate customers	10813255	1435199
2019	New Brunswick	Total sales of electricity to ultimate customers	12824057	1462900
2019	Manitoba	Total sales of electricity to ultimate customers	22019553	1622663
2019	Saskatchewan	Total sales of electricity to ultimate customers	23207739	2683059
2019	Alberta	Total sales of electricity to ultimate customers	52572459	4595274
2019	British Columbia	Total sales of electricity to ultimate customers	56846840	5272466
2019	Ontario	Total sales of electricity to ultimate customers	131112710	17060388
2019	Quebec	Total sales of electricity to ultimate customers	174295138	12460416

Figure 2.27. Total Sales of Electricity by province in 2019

The result of this query will show the total sales of electricity to ultimate customers (residential users, agriculture, mining, manufacturing and other industries) in 2019. Looking at the above, while Quebec sold the highest quantity of electricity (174,259.138MWh), it didn’t earn the highest revenue ($12.5b) from 2019 sales. Ontario was second highest seller of electricity (131,112.71MWh) but earned the highest revenue ($17.1b) from the sales. This suggest that electricity (MWh) was cheaper in Quebec than Ontario in 2019. It is useful for making projections into sales pattern for the future. Exploring sales by each buying customer is an opportunity to be explored in the future.

4. Province with highest cumulative value of total export for the last 10 years

GEO	Electric_power_components	Electricity_value_$000
Quebec	Total exports of electricity to the United States	11578753

Figure 2.28. Province with highest export to the US in the last 10 years

This query outputs province with highest cumulative export in the last 10 years. The first Again, this query can be modified to show cumulative imports. In this case, cumulatively, Quebec has been the biggest exporter of electricity to the United States, exporting electricity worth $11,578,753,000.

3) Data Exploration – Group Queries

3-1. Aggregations, Preparations for Group Data Exploration

3-1-1. Solar Dataset (Adebola O.)

- Our datasets were so different with distinct features, so we had to come up with a way to join the dataset together using a common key. Thus, a new column was created 'GEO' to serve as a primary key, for the joins to be created with other datasets in the team.
- Also, to make the dataset comparable and create an assessment of energy sources for the country, aggregation had to be done using sum, average and percentages to manage the magnitude of difference and ease comparisons across locations and energy sources. This meant that some of the features of the individual dataset was lost.
- Our datasets were so different with distinct features which made it difficult to compare all of them. I had to subdue some of the features of my dataset, and chose only few columns which has certain similarities to other datasets. Some data which had no obvious relationship with other datasets within the team were dropped.
- Datasets had different scales and measurements which prompted the need for conversions and comparisons to be done using percentages and averages in some instances. Quite a number of the queries done on my individual dataset required some aggregation as well due to the size to the dataset. This made it easier to analyse and interpret the output.
- Using views in MySQL made it easier to select the appropriate datasets or columns for comparison with other datasets because it provided an overview of seasonal solar power generation, total capacity, and prices across the various locations.

3-1-2. Wind Dataset (Yu N.)

For Wind generation dataset, I integrated three separate databases into one view by using GEO (province and territory) as a common key to provide the overview of wind power generation across Canada.

- 2008 Annual Wind Power Site data
- Aggregated monthly wind production data. Details are described in my individual report.
- Electric Price Index Data

The new view created includes a total of 16 columns listed below.

- GEO, Total Capacity*, Wind Total Production*, Operating Rate, Total Site Counts, Price Index, Total Spring Generation*, Total Summer Generation*, Total Fall Generation* and Total Winter Generation*
- Two different units were adopted for above columns with * to be able for the fair comparison with other member's datasets. One unit is MW or GWh which aggregated with summation of annual or monthly production data for each province. Other unit is the provincial percentage which computed by following the equation of 'Total Production percentage = Each Province Production / All Canadian Production x 100'.

3-1-3. Thermal Plants Dataset (Shahbaz M.)

The data has been explained with ER diagram in Data Description section with ER diagram. From the available data, it can be concluded that queries can be woven around quantity of power generated by using different fuels with associated cost of fuel and fuel quantity to get an overview of power generated from fuels by using thermal plants. Fuel type was chosen to be investigated individually and location was explored in group queries. The locations names were changed to the abbreviations as shown in figure below by using queries to merge with group data set. No other changes were required to make group queries worked. The creation of the views and the procedure of getting answers have been explained in the group queries attached. There no missing values in the data set therefore no cleaning/wrangling was required.

```sql
-- changed GEO locations to match with other data sets
UPDATE power_generated
SET GEO = REPLACE(GEO, 'Canada', 'CA');
UPDATE power_generated
SET GEO = REPLACE(GEO, 'Alberta', 'AB');
UPDATE power_generated
SET GEO = REPLACE(GEO, 'British Columbia', 'BC');
UPDATE power_generated
SET GEO = REPLACE(GEO, 'Manitoba', 'MB');
UPDATE power_generated
SET GEO = REPLACE(GEO, 'New Brunswick', 'NB');
```

Figure 3.1. Change of locations names to abbreviations

3-1-4. Electricity Dataset (Shade O.)

Firstly, we needed a common column to join the individual datasets. I created a column named GEO as agreed with the team. This column was the codes for the provinces, territories and country.

From the electricity dataset, the team needed total electricity in the following details:

- Total electricity generated in 2008 by province/territory for comparison with the wind power data (value & percentages)

- Total electricity generated in 2018 by province/territory for comparison with the solar power data (value & percentages)
- Total electricity generated from 2005 - 2019 by province/territory for comparison with the thermal plants power generation (value & percentages)

Understanding the need of the different team member informed the manipulation of the dataset. A view was created with the following columns: average electricity generated from 2005 – 2015 by province/territory, total electricity generated by province/territory in 2008 and total electricity generated by province/territory in 2018. GEO would be the column on which the join with the different datasets will be performed.

Each of the column in the view was created by aggregating and left-joining together on GEO.

GEO	Avg_Gen_Electricity	Total_2008_Gen_Electricity	Total_2018_Gen_Electricity
AB	72385558	66086715	75884142
BC	68342760	66072066	69924006
CA	621674010	618047146	640326756
MB	35072179	35523621	31709590
NB	14497116	14278669	13544446
NL	41859021	43214504	43633614
NT	688185	661160	758860
NS	11100257	12215691	10176065
NU	154362	182199	194366

Figure 3.2. gq_ss_pow_gen view created from electricity data

3-1-5. Group wrangling

The datasets were aggregated by province (sum and averages). These views can be extended in future by left-joining aggregated data from other energy sources by provinces/territories to further explore this subject. The team created 2 views from the aggregation and joins of the individual datasets:

- gq_alldata_merged_gwh: This view had 31 columns and 14 rows detailing GEO, Solar power generation (by season, breakeven price, site counts), wind power generation (by season, price index, site counts, capacity), electricity generation and generation from fuels (gas, coal & uranium). Below is a summary of the additional cleaning to create this view:
    - left join of the individual datasets
    - calculating percentages for the respective columns
    - rounding off percentages to 2 places for consistency

```sql
Create View gq_AllData_Merged_GWh AS
select electricity.GEO, solar_TotalCapacity_MW, Solar_TotalAnnualGeneration_MWh/1000 AS Solar_TotalAnnualGeneration_GWh,
Solar_TotalWinterGeneration_MWh/1000 AS Solar_TotalWinterGeneration_GWh,
Solar_TotalSpringGeneration_MWh/1000 AS Solar_TotalSpringGeneration_GWh,
Solar_TotalSummerGeneration_MWh/1000 AS Solar_TotalSummerGeneration_GWh,
Solar_TotalFallGeneration_MWh/1000 AS Solar_TotalFallGeneration_GWh,
Solar_TotalBreakevenPrice_MWhCAD, Solar_TotalReferencePrice_MWhCAD,
Solar_AVGBreakevenPrice_MWhCAD, Solar_AVGReferencePrice_MWhCAD, Solar_TotalSiteCounts, Solar_SiteCounts_Percent,
Wind_TotalCapacity_MW, ROUND(Wind_TotalProduction_GWh,4) AS Wind_TotalProduction_GWh,
ROUND(Wind_OperatingRate,4) AS Wind_OperatingRate, Wind_TotalSiteCounts,
ROUND(Wind_SpringGeneration_GWh,4) AS Wind_SpringGeneration_GWh,
ROUND(Wind_SummerGeneration_GWh,4) AS Wind_SummerGeneration_GWh,
ROUND(Wind_FallGeneration_GWh,4) AS Wind_FallGeneration_GWh, ROUND(Wind_WinterGeneration_GWh,4) AS Wind_WinterGeneration_GWh,
ROUND(PriceIndex,4) AS PriceIndex, Electricity_TotalGen_2008_MWh/1000 AS Electricity_TotalGen_2008_GWh,
Electricity_TotalGen_2018_MWh/1000 AS Electricity_TotalGen_2018_GWh,
Electricity_TotalUsedByProducers_2008_MWh/1000 AS Electricity_TotalUsedByProducers_2008_GWh,
Electricity_TotalUsedByProducers_2018_MWh/1000 AS Electricity_TotalUsedByProducers_2018_GWh,
fuel.Fuel_TotalProduction_MWh_2008/1000 AS Fuel_TotalProduction_GWh_2008,
fuel.Fuel_TotalProduction_MWh_2018/1000 AS Fuel_TotalProduction_GWh_2018,
updated_fuel.Gas_Fuel_TotalProduction_MWh_2018/1000 AS Gas_Fuel_TotalProduction_GWh_2018,
updated_fuel.Coal_Fuel_TotalProduction_MWh_2018/1000 AS Coal_Fuel_TotalProduction_GWh_2018,
updated_fuel.Uranium_Fuel_TotalProduction_MWh_2018/1000 AS Uranium_Fuel_TotalProduction_GWh_2018
from so_ss_power_gen as electricity
left join
solar_power_all as solar
on electricity.GEO = solar.GEO
left join
gq_winddb_bygeo as wind
on electricity.GEO = wind.GEO
left join
sm_updated_fuel_power_Overview as fuel
on electricity.GEO = fuel.GEO
left join
so_elecusedbyprdcers as elecUsedByPrdcers
on fuel.GEO = elecUsedByPrdcers.GEO
left join
sm_updated_fuel_power_Overview as updated_fuel
on updated_fuel.GEO = elecUsedByPrdcers.GEO;
```

Figure 3.3. Query for creating the merged GWh data

- gq_alldata_merged_percent: This view had 31 columns and 14 rows detailing similar information as above view but with percentages for ease of comparison. Below is a summary of the additional cleaning to create this view:
    - conversion of MWh to GWh
    - left join of the individual datasets
    - calculating averages for the respective columns
    - rounding off values to 4 places for consistency

```sql
-- Create View for the province comparison for all energy sectors with unit = percent
DROP VIEW IF EXISTS gq_AllData_Merged_percent;
Create View gq_AllData_Merged_percent AS
select fuel.GEO, Solar_TotalCapacity_Percent, Solar_TotalAnnualGeneration_Percent,
Solar_TotalWinterGeneration_Percent, Solar_TotalSpringGeneration_Percent,
Solar_TotalSummerGeneration_Percent, Solar_TotalFallGeneration_Percent,
Solar_TotalBreakevenPrice_MWhCAD, Solar_TotalReferencePrice_MWhCAD,
Solar_AVGBreakevenPrice_MWhCAD, Solar_AVGReferencePrice_MWhCAD, Solar_TotalSiteCounts, Solar_SiteCounts_Percent,
ROUND(Wind_TotalCapacity_percent,2) AS Wind_TotalCapacity_percent,
ROUND(Wind_TotalProduction_percent,2) AS Wind_TotalProduction_percent,
ROUND(Wind_OperatingRate,2) AS Wind_OperatingRate, Wind_TotalSiteCounts,
ROUND(Wind_SpringGeneration_percent,2) AS Wind_SpringGeneration_percent,
ROUND(Wind_SummerGeneration_percent,2) AS Wind_SummerGeneration_percent,
ROUND(Wind_FallGeneration_percent,2) AS Wind_FallGeneration_percent,
ROUND(Wind_WinterGeneration_percent, 2) AS Wind_WinterGeneration_percent, ROUND(PriceIndex,2) AS PriceIndex,
Electricity_TotalGen_2008_percent, Electricity_TotalGen_2018_percent,
Electricity_TotalUsedByProducers_2008_percent, Electricity_TotalUsedByProducers_2018_percent,
fuel.Fuel_TotalProduction_percent_2008, fuel.Fuel_TotalProduction_percent_2018, updated_fuel.Fuel_Gas_TotalProduction_2018_percent,
updated_fuel.Fuel_Coal_TotalProduction_2018_percent, updated_fuel.Fuel_Uranium_TotalProduction_2018_percent
from sm_updated_power_pct as fuel
left join
solar_power_all as solar
on fuel.GEO = solar.GEO
left join
gq_winddb_bygeo as wind
on fuel.GEO = wind.GEO
left join
so_ss_power_gen as electricity
on fuel.GEO = electricity.GEO
left join
so_elecusedbyprdcers as elecUsedByPrdcers
on fuel.GEO = elecUsedByPrdcers.GEO
left join
sm_updated_fuel_power_Overview as updated_fuel
on updated_fuel.GEO = elecUsedByPrdcers.GEO;
```

Figure 3.4. Query for creating the merged percent data

3-2. Group Data Exploration

With merged tables of all individual datasets via all preparation work described in 3-1, we reviewed provincial trends of power generation at various energy sectors: solar, wind, and thermal (gas, coal, uranium)

1) What seasonal trend of generation can be seen in solar and wind power?

As mentioned in introduction, solar and wind power technology is top two well-developed (developing) renewable energy in the world which domain 29% of renewable energy in Canada (24% from Solar, 5% from wind power) next to hydropower[13].

Due to an integration of solar and wind power individual datasets that seasonal production data is available, it enables us to provide a unique trend of generation.
Figure 3.6. shows comparison of solar, and wind power generation during winter and summer. Higher generation was confirmed during summer from solar, while higher wind generation was seen during winter.

On top of that, wind power generation in fall is actually higher than the one during winter in almost provinces except for Alberta and Ontario, compared to all consistent trends that summer solar production is higher than fall in all provinces (Figure 3.8.).

```
select GEO,
(case when Solar_TotalWinterGeneration_GWh < Solar_TotalSummerGeneration_GWh
    then 'Solar: Summer > Winter' else NULL end) AS Solar,
(case when Wind_WinterGeneration_GWh > Wind_SummerGeneration_GWh
    then 'Wind: Summer < Winter' else 'No sites located' end) AS Wind
from gq_renewable_season;
```

Figure 3.5. Query showing seasonal generation of Solar and Wind: Summer vs Winter

GEO	Solar	Wind
AB	Solar: Summer > Winter	Wind: Summer < Winter
BC	Solar: Summer > Winter	Wind: Summer < Winter
MB	Solar: Summer > Winter	Wind: Summer < Winter
NB	Solar: Summer > Winter	Wind: Summer < Winter
NS	Solar: Summer > Winter	Wind: Summer < Winter
ON	Solar: Summer > Winter	Wind: Summer < Winter
PE	Solar: Summer > Winter	Wind: Summer < Winter
QC	Solar: Summer > Winter	Wind: Summer < Winter
SK	Solar: Summer > Winter	Wind: Summer < Winter
NL	Solar: Summer > Winter	No sites located
NT	Solar: Summer > Winter	No sites located
NU	Solar: Summer > Winter	No sites located
YT	Solar: Summer > Winter	No sites located

Figure 3.6. Seasonal Generation of Solar and Wind: Summer vs Winter

```sql
select GEO,
(case when Solar_TotalFallGeneration_GWh < Solar_TotalSummerGeneration_GWh
    then 'Solar: Summer > Fall' else NULL end) AS Solar,
(case when Wind_WinterGeneration_GWh < Wind_FallGeneration_GWh
    then 'Wind: Fall > Winter'
when Wind_WinterGeneration_GWh > Wind_FallGeneration_GWh then 'Wind: Winter > Fall'
when GEO in ('NL', 'NT', 'NU', 'YT') then 'No sites located' else NULL end) AS Wind
from gq_renewable_season;
```

Figure 3.7. Query showing seasonal generation of Solar and Wind: comparison to Fall

GEO	Solar	Wind
AB	Solar: Summer > Fall	Wind: Winter > Fall
BC	Solar: Summer > Fall	Wind: Fall > Winter
MB	Solar: Summer > Fall	Wind: Fall > Winter
NB	Solar: Summer > Fall	Wind: Fall > Winter
NS	Solar: Summer > Fall	Wind: Fall > Winter
ON	Solar: Summer > Fall	Wind: Winter > Fall
PE	Solar: Summer > Fall	Wind: Fall > Winter
QC	Solar: Summer > Fall	Wind: Fall > Winter
SK	Solar: Summer > Fall	Wind: Fall > Winter
NL	Solar: Summer > Fall	No sites located
NT	Solar: Summer > Fall	No sites located
NU	Solar: Summer > Fall	No sites located
YT	Solar: Summer > Fall	No sites located

Figure 3.8. Seasonal Generation of Solar and Wind: comparison to Fall

This observation alludes that weather impact on each energy production would work differently. Since sunlight duration is most dominated factors to control solar production, higher production during summer is reasonable to explain. On the other hand, wind speed is an important factor to wind power production that tend to have opposite trend against sunlight, that is moderate during summer and windy from Nov to March. Lowest wind production in summer as well as the difference production level of winter vs. fall in provinces are explained by wind speed seasonal trends.

2) What is total capacity and site counts for solar and wind power generation across Canada?

   With usage of same integrated table described in group question 1, from figure 3.10. showing total capacity and site counts, large amount of wind and solar plants are located in Ontario and Quebec for both wind and solar power.
   We can see solar power is widely developed across Canada compared to wind power as well. There is certain amount of solar power sites/projects in the provinces which had no wind sites such as Yukon, North West Territory, and Newfoundland and Labrador.
   Other insights from figure 3.12. is that British Columbia who hold the larger solar capacity next to Ontario and Quebec, had only 5 wind sites, while more wind sites are developed in Alberta who had less solar site number compared to others.

```sql
select GEO, solar_TotalCapacity_MW AS Solar, Wind_TotalCapacity_MW AS Wind
from gq_alldata_merged_gwh
where GEO != 'CA'
order by solar_TotalCapacity_MW desc;
```

Figure 3.9. Query showing Solar and Wind Site Capacity (MW)

GEO	Solar	Wind
ON	7347030	4103
QC	4254300	2964
BC	3237020	685
NS	2724770	391
NB	2388090	484
SK	2319550	449
AB	1467000	1439
MB	1272200	258
NL	758752	NULL
PE	119044	201
NU	9619.68	NULL
NT	7214.76	NULL
YT	3607.38	NULL

Figure 3.10. Solar and Wind Site Capacity (MW)

```sql
select GEO, Solar_TotalSiteCounts AS Solar, Wind_TotalSiteCounts AS Wind
from gq_alldata_merged_gwh
where GEO != 'CA'
order by Solar_TotalSiteCounts desc;
```

Figure 3.11. Query showing Solar and Wind Site Counts

GEO	Solar	Wind
ON	1284	33
QC	1195	28
BC	1061	5
SK	720	7
NS	604	10
MB	546	2
AB	502	20
NB	428	5
NL	374	NULL
PE	73	6
NU	8	NULL
NT	6	NULL
YT	3	NULL

Figure 3.12. Solar and Wind Site Counts

3) How is the provincial contribution of electricity generation from each energy sources? Any notable distribution?

Our original plan for this query was to create comparison table with total production (GWh) by provinces. However, we noticed a couple of difficulties.

- Year coverage of each datasets are different
- Summation of productions from solar power dataset is wrong since there is multiple scenarios or types such as different tariff systems are applied, or different array system is adopted etc. at single project which is unable to identify as one representative value in a simple way.

Those became an obstacle to implement an apple-to-apple comparison over various energy sources with same methodology (Figure 3.13.).

GEO	Electricity_TotalGen_2008_GWh	Wind_TotalProduction_GWh	Solar_TotalAnnualGeneration_GWh	Uranium_Fuel_TotalProduction_GWh_2018	Coal_Fuel_TotalProduction_GWh_2018	Gas_Fuel_TotalProduction_GWh_2018	Fuel_TotalProduction_GWh_2008
QC	193683.4490	9077.6394	6304517.12	0.0000	NULL	72.6630	1104.3910
ON	163842.4800	13647.5007	11310320.64	90155.2360	0.0000	9941.0600	33813.4120
AB	66086.7150	4549.2420	2126250.88	NULL	29384.7710	21653.2700	51339.3710
BC	66072.0660	1764.2791	4275940.608	NULL	NULL	1218.6810	3392.7240
NL	43214.5040	NULL	947054.016	NULL	NULL	NULL	1129.6780
MB	35523.6210	862.4375	1925770.24	NULL	5.2610	9.7310	423.3220
SK	21766.5460	1474.5556	3485135.872	NULL	10346.2280	9017.4130	16245.2480
NB	14278.6690	1476.9097	3596545.792	4874.3600	2328.6690	977.7410	8521.3570
NS	12215.6910	1260.7522	3953242.112	NULL	4975.1790	1424.4890	10752.4560
NT	661.1600	NULL	10170.733	NULL	NULL	66.0570	86.8800
YT	372.6150	NULL	4509.015	NULL	NULL	30.1300	23.8850
NU	182.1990	NULL	12488.761	NULL	NULL	NULL	182.1990
PE	147.4310	685.8925	168722.928	NULL	NULL	NULL	5.7650

Figure 3.13. Total Production (GWh) from various energy sources

To solve this problem, we decided to adopt a provincial percentage of each production i.e., production in each province / Canadian production to provide the provincial contribution of electricity generation from each energy sources (Figure 3.15.).

From figure 3.15., the fuel production in the Ontario and Quebec is relatively less while their contribution from solar and wind is quite larger. As well as Alberta who 52% of energy is from fuel, produced relatively less contribution from both wind and solar energies.

Notable thing is that Quebec generates 33.26% of electricity generated in Canada, only 1.35% of power generated from fuel sources.

```
select GEO, Electricity_TotalGen_2018_percent AS TotalElectricity,
    Wind_TotalProduction_percent As Wind,
    Solar_TotalAnnualGeneration_Percent As Solar,
    Fuel_TotalProduction_percent_2018 As Fuel,
    Electricity_TotalUsedByProducers_2018_percent As Electricity_UsedByProducer
from gq_alldata_merged_percent
where GEO != 'CA'
order by Electricity_TotalGen_2018_percent DESC;
```

Figure 3.14. Query showing Provincial contribution of electricity generation from energy sources (%)

GEO	TotalElectricity	Wind	Solar	Fuel	Electricity_UsedByProducer
QC	33.26	26.61	16.54	1.35	28.74
ON	24.41	40.01	29.67	10.73	13.69
AB	11.85	13.34	5.58	52.00	31.02
BC	10.92	5.17	11.22	1.69	21.12
NL	6.81	NULL	2.48	1.39	1.08
MB	4.95	2.53	5.05	0.03	0.30
SK	3.77	4.32	9.14	19.57	0.63
NB	2.12	4.33	9.43	4.83	2.68
NS	1.59	3.70	10.37	7.97	0.28
NT	0.12	NULL	0.03	0.17	0.45
PE	0.10	2.01	0.44	0.00	0.01
YT	0.07	NULL	0.01	0.06	0.00
NU	0.03	NULL	0.03	0.20	0.01

Figure 3.15. Provincial contribution of electricity generation from energy sources (%)

4) In province, which energy industry (wind power, coal, gas and uranium) is the most oriented production sources and any difference over provinces?

   To address this question, we computed a percentage with different way from group question 3. What we would like to know from this question 4 is energy source distribution % in each province with a below equation:

$$Production\ from\ each\ energy(\%) = \frac{Total\ Production(GWh)}{Total\ Electricity\ in\ each\ Province\ (GWh)} \times 100$$

   Solar power generation is excluded because its total production calculated from generation data doesn't represent the provincial total production as explained in group question 3.

   From fig 3.17., 3 to 12% of energy is generated from wind power and higher dependency of gas and coal energies was confirmed in Alberta, Saskatchewan and Nova Scotia. Nuclear energy (Uranium) is developed only in Ontario and New Brunswick, that 58% of energy in Ontario was contributed from nuclear energy sources.

```
select GEO, ROUND(Wind_TotalProduction_GWh/Electricity_TotalGen_2018_GWh*100,2) AS Wind,
    ROUND(Coal_Fuel_TotalProduction_GWh_2018/Electricity_TotalGen_2018_GWh*100,2) AS Coal,
    ROUND(Gas_Fuel_TotalProduction_GWh_2018/Electricity_TotalGen_2018_GWh*100, 2) As Gas,
    ROUND(Uranium_Fuel_TotalProduction_GWh_2018/Electricity_TotalGen_2018_GWh*100, 2) As Uranium
from gq_alldata_merged_gwh
where GEO != 'CA' and GEO !='PE' and GEO !='NL' and GEO != 'NU'
order by Electricity_TotalGen_2018_GWh desc;
```

Figure 3.16. Query showing Wind, Coal, Gas, and Uranium production % of total electricity in Province

GEO	Wind	Coal	Gas	Uranium
QC	4.26	NULL	0.03	0.00
ON	8.73	0.00	6.36	57.69
AB	5.99	38.72	28.53	NULL
BC	2.52	NULL	1.74	NULL
MB	2.72	0.02	0.03	NULL
SK	6.11	42.87	37.36	NULL
NB	10.90	17.19	7.22	35.99
NS	12.39	48.89	14.00	NULL
NT	NULL	NULL	8.70	NULL
YT	NULL	NULL	6.31	NULL

Figure 3.17. Wind, Coal, Gas, and Uranium production % of total electricity in Province

5) What is the contribution of power generated by fuels to the total power across Canada, provinces and territories in 2008 and 2018? Is there any trend?

   To answer this query, dataset for fuel generated power (Set C) and data set for total power generated across Canada (Set D) were joined together. The results in table below show that contribution of fuel generated power has reduced from 2008 to 2018 for Canada, all provinces and territories except in Saskatchewan, Northwest territories and Yukon. In Saskatchewan, the contribution of fuels generated power has increased from 74.63% in 2008 to 80.23% in 2018. Also, in Ontario, the contribution of fuels generated has reduced significantly from around 21% in 2008 to around 7% in 2018.

GEO	Total_2008_Gen_Electricity	Total_2018_Gen_Electricity	Fuel_TotalProduction_MWh_2008	Fuel_TotalProduction_MWh_2018	Fuel_TotalProduction_MWh_2008_Percent	Fuel_TotalProduction_MWh_2018_Percent
AB	66086715	75884142	51339371	51443160	77.68	67.79
BC	66072066	69924006	3392724	1671864	5.13	2.39
CA	618047146	640326756	127020688	98932826	20.55	15.45
MB	35523621	31709590	423322	30018	1.19	0.09
NB	14278669	13544446	8521357	4782588	59.68	35.31
NL	43214504	43633614	1129678	1373237	2.61	3.15
NT	661160	758860	86880	166527	13.14	21.94
NS	12215691	10176065	10752456	7888771	88.02	77.52
NU	182199	194366	182199	193660	100	99.64
ON	163842480	156273287	33813412	10616614	20.64	6.79
PE	147431	648300	5765	3048	3.91	0.47
QC	193683449	212966773	1104391	1339938	0.57	0.63
SK	21766546	24135457	16245248	19364061	74.63	80.23
YT	372615	477850	23885	59340	6.41	12.42

Figure 3.18. Contribution of fuels generated from 2008 to 2018

6) What is the contribution of power generated by gas, coal and uranium in 2018 to the total power supply in Canada and across all provinces and territories?

To answer this query, dataset for fuel generated power (Set C) and data set for total power generated across Canada (Set D) were joined together. The results in table below show the query results. In Canada, contribution of power generated by using gas as fuel was 6.94%, by using coal was 7.35% and by using uranium was 14.84% in 2018. Uranium was being used as fuel in only two provinces i.e., Ontario and New Brunswick in 2018 and contributed 57.69 % and 35.99% to both provinces total power respectively in 2018.

GEO	Gas_Fuel_TotalProduction_MWh_2018	Fuel_Gas_TotalProduction_2018_percent	Coal_Fuel_TotalProduction_MWh_2018	Fuel_Coal_TotalProduction_2018_percent	Uranium_Fuel_TotalProduction_MWh_2018	Fuel_Uranium_TotalProduction_2018_percent
AB	21653270	28.53	29384771	38.72	NULL	NULL
BC	1218681	1.74	NULL	NULL	NULL	NULL
CA	44411235	6.94	47040108	7.35	95029596	14.84
MB	9731	0.03	5261	0.02	NULL	NULL
NB	977741	7.22	2328669	17.19	4874360	35.99
NL	NULL	NULL	NULL	NULL	NULL	NULL
NT	66057	8.7	NULL	NULL	NULL	NULL
NS	1424489	14	4975179	48.89	NULL	NULL
NU	NULL	NULL	NULL	NULL	NULL	NULL
ON	9941060	6.36	0	0	90155236	57.69
PE	NULL	NULL	NULL	NULL	NULL	NULL
QC	72663	0.03	NULL	NULL	0	0
SK	9017413	37.36	10346228	42.87	NULL	NULL
YT	30130	6.31	NULL	NULL	NULL	NULL

Figure 3.19. Contribution of power generated by gas, coal and uranium in 2018

4) Discussion

Adebola O.

- The dataset on solar power generation were estimates produced to assess the potentials of solar power in various communities across Canada. It was really detailed and categorized with location details (longitude and latitude). It would have been great to be able to compare each of these locations with other energy sources to determine if certain areas have peculiar properties that make that 'hotspots' for energy generation in the country. However, these details had to be subdued while writing the individual and group queries.
- In the course of analysing and creating queries, there was a need for lots of transformations and conversions due to the size to the solar power dataset. These transformations made it easier to analyse and interpret the output. From this, I learnt that the dataset downloaded from a database directly might not make much sense and might produce better queries after some form of wrangling and aggregations are done.
- If I had the chance to work with this dataset or a similar one, it would be interesting to compile additional datasets and explore across more locations and other countries to create a case study of the Canadian energy generation potential compared with others. This would also include a comparison of the generation and utilisation trends/patterns across the countries.
- Also, it would be interesting to work with actual solar power production and utilisation data complied after a few years of implementation of the solar power projects. This would create

an opportunity to compare estimates to the actual numbers and also provide a time series analysis for these datasets.

Yu N.

- Due to a usage of different data sources for the project, our common key over all datasets is only GEO (province and territory). With GEO as foreign key, we succeeded in merging all datasets to provide the province overview from various energy sources including solar, wind, fuel, and electricity. However, aggregation of all sites/project information to the province level leads a significant reduction of data resolution. For example, even though we had more than 1000 project or plants information for each province which number varies from low to high, we could only discuss about overall average with aggregated data by province.
- To avoid losing the detail information of original dataset, I would like to suggest a usage of spatial key (longitude and latitude) for merging datasets as a potential improvements of data exploration. Since our dataset of solar and wind power (Set A and B) have longitude and latitude of project or site locations, there is an opportunity to compare data with a deeper resolution such as city or distinct area. To make spatial join, I think we would create a category which stores district number or code based on longitude/latitude locations or follow each city's longitude/latitude boundary to assign city names. This enables us more detailed lookback on solar-wind power comparison across Canada.
- Other things I noticed was how difficult it was to make a consistent comparison of individual data (in my case, wind power dataset) to other datasets. For example, wind power generation data is collected from 2008. Hourly production data is available at each site from 2008 to 2010 but it is not measured up to today. This dataset is not great to answer common questions such as what the recent trend in renewable energy is. Because of lack of period coverage, we had to use our insight from integrated datasets by compromise. A good example of this is bringing up the idea to use provincial contribution % instead of actual production number assuming the provincial contribution has not changed significantly in the past 10 years.
- However, even we had noticeable flaws of our datasets, I think we still got a chunk of learning and mature discussion through our group queries described above.

Shahbaz M.

- Fuels generated power has been a major contribution to power generation in Canada as indicated by data from 2005 to 2019. Alberta is the most fuel efficient (least fuel consumption per unit of power generated) for many fuel categories e.g., gas, coal etc. The contribution of fuel generated power has reduced from 2008 to 2018 for Canada, all provinces and territories except in Saskatchewan, Northwest territories and Yukon. In Canada, contribution of power gas generated was 6.94%, by using coal was 7.35% and by using uranium was 14.84% in 2018. Uranium was being used as fuel in only two provinces i.e., Ontario and New Brunswick in 2018 and contributed 57.69 % and 35.99% to both provinces total power respectively in 2018.
- I would have liked to compare each fuel type contribution over the years to Canadian power generation industry if given more time. More data to give more insight into economics of fuels generated power would have been very useful to compare to with breakeven points of Wind power and solar power in the other datasets. Therefore, I would

have searched for more data. Also, I would like to compare it with other power generation sources in Canada e.g., hydro, biomass, tidal etc. therefore I will gather more data for other power sources as well for their economics and contribution. Comparing Canadian fuel generation power industry with other countries will be another avenue to explore.

Shade O.

- In 2019, Canada imported from the United States, a quarter of the electricity exported to the United States. Quebec is the largest producer and exporter of electricity to the United States. From 2013 – 2019, Alberta, New Brunswick and Prince Edward Island have all had a deficit in unallocated electricity.
- I was able to leverage some tools and functions that we couldn't really elaborate on in the course of the semester. Aggregations, view creation, joins, reshaping tables, utilizing functions such as average, round etc. were used during this project. It also provided an opportunity to create tables from scratch, clean data, query the data to obtain results and interpret results obtained in a meaningful way.
- Doing this project again, I would obtain more datasets on different energy sources to provide a more holistic review of all energy sources in Canada. I would also initiate team discussion on dataset sooner as this delayed our progress. While visualization was not a requirement for this course, I would have leveraged that more to provide pictorial representation of my data.
- In future, I would love to explore other sources of energy and their contribution to electricity generation. Also, studying the electricity generation trend over time via the different energy sources would be great. This will provide a clear picture of Canada's transition from fossil fuels to cleaner renewable energy. It will be interesting to see how the transition to cleaner energy is impacting the greenhouse gases emission (if any).

5) Conclusion

We successfully integrated all datasets from individual member's work for group data exploration in MySQL environment, which achieved the objective of providing an overview of various energy sectors across Canada including solar, wind, thermal (fuel, coal, gas, uranium) industries. The queries we ran individually and collectively also helped us answer our research questions and provided valuable insights into this topic.

As a team, we recognize that this is a very broad topic and we only scratched the surface. There are more opportunities to keep exploring this topic by obtaining a broader dataset in the future as described in our individual discussions above.

References:

[1]Government of Canada. (2020). *Energy and the Economy*. Retrieved from https://www.nrcan.gc.ca/science-data/data-analysis/energy-data-analysis/energy-facts/energy-and-economy/20062

[2]Solar energy: Markets, economics and policies, Govinda R. Timilsina, Lado Kurdgelashvili, Patrick A. Narbel, Renewable and Sustainable Energy Reviews 16, 2012, DOI number: https://doi.org/10.1016/j.rser.2011.08.009

[3]Canada Energy Regulator. (2019). *Canada's Energy Future 2019: Energy Supply and Demand Projections to 2040* (EF2019). Retrieved from https://www.cer-rec.gc.ca/en/data-analysis/canada-energy-future/2019/index.html

[4]Electricity facts - Natural Resources Canada (2020). https://www.nrcan.gc.ca/science-data/data-analysis/energy-data-analysis/energy-facts/electricity-facts/20068

[5]Government of Canada. (2018). *Economics of Solar Power in Canada.* Retrieved from https://open.canada.ca/data/en/dataset/7b569daa-59cd-4c14-8ffe-6bc8ebe8d8bb

[6]Government of Canada. (2018). *Pan-Canadian Wind Integration Study.* Retrieved from https://open.canada.ca/data/en/dataset/b1e5aeb7-913d-49b1-bbfb-16864bf5c6bf

[7]Government of Canada. (2018). *Monthly Wind Speed at weather station.* Retrieved from : https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data/wind-speed-access.html

[8]Government of Canada. (2018). *Monthly Electric power selling price index.* Retrieved from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810020401&pickMembers%5B0%5D=1.8&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2008&cubeTimeFrame.endMonth=01&cubeTimeFrame.endYear=2010&referencePeriods=20080101%2C20100101

[9]Government of Canada. (2020). *Electricity from fuels, annual generation by electric utility thermal plants.* Retrieved from https://open.canada.ca/data/en/dataset/0c930a3c-fb06-4bda-b81b-8031eed34a0b

[10]Government of Canada. (2020). *Electric power generation, annual fuel consumed by electric utility thermal plants.* Retrieved from https://open.canada.ca/data/en/dataset/5fcea0b0-b0a7-4421-b871-0b7cbf5b51f9

[11]Government of Canada. (2020). *Electric power generation, annual cost of fuel consumed by electric utility thermal plants*. Retrieved from https://open.canada.ca/data/en/dataset/42e522a6-950b-47ff-ae71-31dbbfce5646

[12]Government of Canada. (2020). *Electric power, electric utilities and industry, annual supply and disposition*. Retrieved from https://open.canada.ca/data/en/dataset/835ef477-1ab6-4dbd-a3e2-229a529bcd1e

[13]Government of Canada. (2020). *Renewable energy facts.* Retrieved from https://www.nrcan.gc.ca/science-data/data-analysis/energy-data-analysis/energy-facts/renewable-energy-facts/20069