A comparative analysis of Gas Consumption and Electricity Generation in Ireland (Before and After COVID-19)

Student Name: Divesh Patil ***Student Id: 23401478***

Student Name: Ruchita Raut ***Student Id: 24240702***

Student Name: Anuja Tawde ***Student Id: 24257788***

Abstract

Reliable electricity and gas supply is essential for everyday life and economic activity, especially in Ireland where natural gas is widely used for electricity generation and heating. This project analyses daily gas consumption and electricity generation data in Ireland to understand how different fuel sources interact, how energy use changes with the seasons, and how the system behaves over time, including during the COVID-19 period. Data from the Central Statistics Office (CSO) were stored and processed using PostgreSQL and MongoDB within Docker to ensure a structured and reproducible workflow. Exploratory data analysis, correlation analysis, and time-series methods were used to identify trends, seasonal patterns, and differences in gas demand across sectors. An ARIMA forecasting model was applied to estimate short-term gas consumption for power plants. The results show clear seasonal behaviour, a strong contribution from renewable sources such as wind, and the continued role of natural gas as a flexible balancing fuel. Comparison of pre-COVID and COVID periods suggests that Ireland's electricity system remained stable, highlighting its resilience to external disruption. Overall, this study demonstrates how data analytics can be used to gain practical insights into real-world energy systems.

Keywords – *electricity generation, natural gas consumption, time-series analysis, EDA, COVID-19*

Introduction

This project analyses Ireland's daily gas consumption and electricity generation data to better understand energy usage patterns, fuel interactions, and system behaviour over time. Given Ireland's strong reliance on natural gas for both electricity generation and heating, understanding how gas demand and electricity generation respond to seasonal changes and external disruptions such as COVID-19 is particularly important.

The study examines daily data from multiple years, including pre-COVID and COVID periods. A complete data analytics pipeline was implemented using Python, with data stored in both PostgreSQL (via Docker) and MongoDB, depending on data structure. Semi-structured gas consumption data were cleaned, transformed, and stored before exploratory data analysis and visualization were applied to identify trends, seasonal patterns, and differences across customer sectors. Time-series analysis techniques were used to study electricity generation behaviour, and an ARIMA-based forecasting model was applied to predict short-term gas demand.

Results show clear seasonal patterns in both gas consumption and electricity generation. Renewable sources, particularly wind, play a significant role in electricity generation, while natural gas acts as a flexible balancing source when renewable output is lower. Comparison of pre-COVID and COVID periods indicates that Ireland's electricity system remained largely stable, with no major disruption to overall generation trends.

Overall, this project demonstrates how data analytics and exploratory analysis can be effectively applied to real-world energy datasets to gain insight into fuel mix dynamics, seasonal behaviour, and system resilience in Ireland.

Related works

Previous research on energy systems has mainly focused on natural gas demand forecasting, renewable energy integration, and the impact of external disruptions such as the COVID-19 pandemic. These studies informed both the analytical focus and methodological choices of this project.

Several studies review natural gas demand forecasting techniques, ranging from traditional time-series models such as ARIMA to advanced machine-learning and deep-learning approaches [1], [3]. Although advanced models often achieve higher predictive accuracy, they typically require large datasets, complex parameter tuning, and offer limited interpretability. Given the limited pre-COVID data and the academic focus of this study, an ARIMA-based approach was selected to provide transparent and reliable short-term forecasts.

Research analyzing the impact of COVID-19 on gas consumption across Europe shows that temperature and seasonality remain the dominant drivers of demand, even during lockdown periods [2]. However, these studies generally focus on aggregated national demand and do not examine consumption differences across customer categories. This project extends existing work by focusing on Ireland and analyzing daily gas consumption by customer type, offering clearer insight into sector-level behaviour prior to and during COVID-19. Other studies investigate energy consumption changes across broad sectors such as transport, industry, and households using EU-wide datasets [4]. While these studies highlight significant reductions in transport and service sectors during COVID-19, their reliance on aggregated data limits detailed interpretation. In contrast, this project adopts a more granular approach using daily data, enabling a clearer understanding of customer-level demand patterns.

Research conducted in industrial settings demonstrates that COVID-19 caused noticeable shifts in electricity and gas demand and that time-series models can capture these changes [5]. Although these findings support the use of pre- and post-COVID comparisons, their results are context-specific and may not directly translate to Ireland's energy system. Similarly, financial-focused studies show that COVID-19 increased uncertainty within the energy sector but do not explain changes in daily energy consumption [6]. Studies of Ireland's electricity system highlight that increasing wind penetration introduces variability and increases reliance on flexible balancing sources, particularly natural gas [7]. While many studies focus on policy frameworks or simulated scenarios [8], [9], this project differs by analyzing observed daily data to assess real-world fuel interactions. Unlike broader cross-country studies [10], this research focuses specifically on Ireland to examine whether COVID-19 coincided with observable deviations in electricity generation behaviour.

Overall, existing literature highlights the importance of renewable variability, seasonality, and system resilience. This project builds on these ideas using an exploratory, data-driven approach based on time-series analysis, correlation analysis, and contribution strength evaluation. While this approach does not establish causality, it provides clear and interpretable insights aligned with the objectives of this module.

Methodology

This section describes the datasets used, database storage approach, data pre-processing steps, and analytical methods applied in this project.

Data sources

The datasets used in this project were obtained from the Central Statistics Office (CSO) of Ireland, accessed via the CSO data portal, data.cso.ie., making it reliable and authoritative source for analysis. Both datasets cover multiple years, including the COVID-19 period, allowing comparison of electricity generation before, during and after the pandemic.

Data descriptions

Dataset 1:

The Dataset Used in this project is NGSD02.20251113T141152.json (JSON Line Format) which is a semi-structured JSON dataset Converted into (JSON Array Format). The dataset gives the daily gas demand values categorised by the customer type, which satisfies the answer to the research question.

Dataset 2:

The Dataset Used for project analysing of Ireland's daily gas consumption is gas_daily_raw in (CSV) format. The dataset includes the information on gas demand over different customer types, which covers the post-COVID period. The dataset was chosen due its time scale, sector-level details, and appropriate to the research question.

Dataset 3:

Two primary datasets were used in this study: an electricity generation dataset and a gas supply dataset for Ireland. The electricity dataset contains daily electricity generation values broken down by fuel type, while the gas supply dataset contains daily gas supply values associated with electricity production.

Both datasets include a common date field, enabling accurate alignment and integration. The data was provided in structured CSV format and includes key variables such as date, total electricity generation, generation by individual fuel sources, and gas supply values.

	Statistic Label	Day	Supply Type	UNIT	VALUE
lapse Output	ked Gas Daily Supply	2018 January 01	Indigenous Gas Production	Gigawatt hours	111.682190
1	Networked Gas Daily Supply	2018 January 01	Gas Imports	Gigawatt hours	22.021667
2	Networked Gas Daily Supply	2018 January 01	All Networked Gas Supply	Gigawatt hours	133.703857
3	Networked Gas Daily Supply	2018 January 02	Indigenous Gas Production	Gigawatt hours	115.101624
4	Networked Gas Daily Supply	2018 January 02	Gas Imports	Gigawatt hours	21.883611

	Statistic Label	Day	Time Bands	Primary Fuel Output	UNIT	VALUE
0	Metered Electricity Generation	2020 January 01	All time periods	Battery Storage	Megawatt hours	NaN
1	Metered Electricity Generation	2020 January 01	All time periods	Biomass/Peat	Megawatt hours	6218.2800
2	Metered Electricity Generation	2020 January 01	All time periods	Coal	Megawatt hours	926.1700
3	Metered Electricity Generation	2020 January 01	All time periods	Distillate	Megawatt hours	0.0000
4	Metered Electricity Generation	2020 January 01	All time periods	Gas	Megawatt hours	32304.0271

Figure 1: Raw Datasets

Database Management

Dataset 1:

MongoDB Usage and DOCKER:

After conversion of clean JSON file into Array it was stored in a MongoDB database which is running inside a Docker container. Due to its flexibility in handling semi-structured data MongoDB database was chosen. The data was accessed in python with the help of PyMongo library and converted into pandas Data frame for analysis.

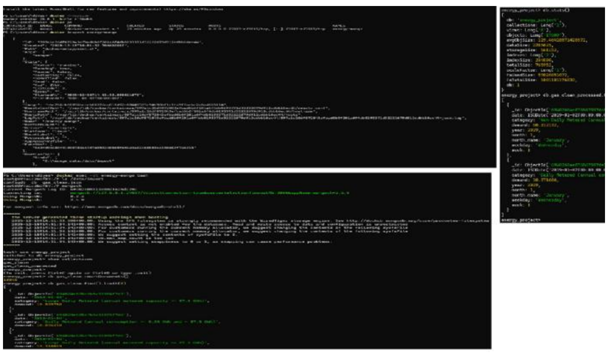


Figure 2: Docker and MongoDB Deployment

Dataset 2:

PostgreSQL Usage and DOCKER:

PostgreSQL was deployed using docker to show structure, data storage, and real database interaction. The docker volume was used to make sure continuous storage of database, which allow the data to remain available, even when the container was stopped. The PostgreSQL database was examined programmatically from Python using SQL Alchemy. Whereas SQL queries, where implemented to extract gas consumption data from the database into pandas' data frame for further analysis.

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\HP> docker ps -a
CONTAINER ID   IMAGE         COMMAND                  CREATED       STATUS       PORTS                                         NAMES
10f8e1a62213   postgres:15   "docker-entrypoint.s…"   2 weeks ago   Up 2 hours   0.0.0.0:5432->5432/tcp, [::]:5432->5432/tcp   energy-postgres
PS C:\Users\HP> docker start energy-postgres
energy-postgres
PS C:\Users\HP> docker ps
CONTAINER ID   IMAGE         COMMAND                  CREATED       STATUS       PORTS                                         NAMES
10f8e1a62213   postgres:15   "docker-entrypoint.s…"   2 weeks ago   Up 2 hours   0.0.0.0:5432->5432/tcp, [::]:5432->5432/tcp   energy-postgres
PS C:\Users\HP>
```

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka
.ms/PSWindows

PS C:\Users\HP> docker exec -it energy-postgres psql -U student -d customer_
type
psql (15.15 (Debian 15.15-1.pgdg13+1))
Type "help" for help.

customer_type=# \c customer_type
You are now connected to database "customer_type" as user "student".
customer_type=# \dt
                  List of relations
 Schema |            Name             | Type  |  Owner
--------+-----------------------------+-------+---------
 public | gas_customer_type           | table | student
 public | gas_daily_consumption_clean | table | student
 public | gas_daily_raw               | table | student
(3 rows)
```

Figure 3: Docker and PostgreSQL deployment

Dataset 3:

To support a structured and reproducible workflow, the datasets were stored programmatically in a PostgreSQL database via Docker, ensuring consistency across systems. Raw datasets were inserted into database tables using Python scripts, and SQL queries were used to retrieve data for preprocessing and analysis. This approach ensured that we followed good data management practices.

```
(base) anuja@Anujas-MacBook-Air Desktop % docker exec -it apdv_postgres psql -U
postgres -d apdv_db
psql (15.15 (Debian 15.15-1.pgdg13+1))
Type "help" for help.

apdv_db=# \dt
            List of relations
 Schema |       Name        | Type  |  Owner
--------+-------------------+-------+----------
 public | electricity_raw   | table | postgres
 public | gas_supply_raw    | table | postgres
(2 rows)
```

Figure 4: Raw data tables created in SQL

Data per-processing

Dataset 1:

Data Processing:

The following steps were applied for the preprocessing:

After successful connection the id was automatically created, as it was unnecessary column, so it was removed. Handling missing and duplicate values where required. Conversion of date columns into datetime format. From 2019 to 2020 the data was used for visualization as per the research question. Created columns for more clear understanding of visualization. Attributes include date, category, demand, year, month, month name, weekday, week.

Saved the cleaned dataset as(gas_clean) back to the MongoDB.

Algorithms and Processing Logic:

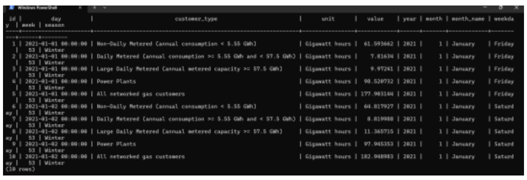
Group-by aggregations used to compute daily, monthly and yearly demand. Whereas Statistical summaries were generated to understand the distribution and variability.

Technology used: Python (pandas, matplotlib, seaborn, statsmodels). MongoDB (NoSQL database).Docker(containerised MongoDB).Jupyter Notebook (development environment)

Dataset 2:

Data Extraction and Processing:

The Data was extracted programmatically from PostgreSQL sing the SQL Alchemy, whereas the tools and Technologies used for the analysis is Python (Pandas, Matplotlib, Seaborn, Auto-Arima), PostgreSQL (structured data storage), DOCKER (database deployment), SQL Alchemy was used for database connection and stream lit is used for dashboard visualization. The processing of dataset is the Conversion of date columns into datetime format then Aggregation of gas demand by customer type ,Handling missing and duplicates where required and created columns for more clear understanding of visualization such as id, day, customer type, unit, value, year, month, month name, weekday, week, season and saved the cleaned dataset back to the PostgreSQL which name as Gas_daily_consumption_clean. The Time Period of the study is Post-COVID Gas Consumption from the year (2021 -2023).



id	day	customer_type	unit	value	year	month	month_name	weekday	week	season
1	2021-01-01 00:00:00	Non-Daily Metered (annual consumption < 5.55 GWh)	Gigawatt hours	61.593662	2021	1	January	Friday	53	Winter
2	2021-01-01 00:00:00	Daily Metered (annual consumption >= 5.55 GWh and < 57.5 GWh)	Gigawatt hours	7.81634	2021	1	January	Friday	53	Winter
3	2021-01-01 00:00:00	Large Daily Metered (annual metered capacity >= 57.5 GWh)	Gigawatt hours	9.97241	2021	1	January	Friday	53	Winter
4	2021-01-01 00:00:00	Power Plants	Gigawatt hours	98.520732	2021	1	January	Friday	53	Winter
5	2021-01-01 00:00:00	All networked gas customers	Gigawatt hours	177.903144	2021	1	January	Friday	53	Winter
6	2021-01-02 00:00:00	Non-Daily Metered (annual consumption < 5.55 GWh)	Gigawatt hours	64.817927	2021	1	January	Saturday	53	Winter
7	2021-01-02 00:00:00	Daily Metered (annual consumption >= 5.55 GWh and < 57.5 GWh)	Gigawatt hours	8.819988	2021	1	January	Saturday	53	Winter
8	2021-01-02 00:00:00	Large Daily Metered (annual metered capacity >= 57.5 GWh)	Gigawatt hours	11.365715	2021	1	January	Saturday	53	Winter
9	2021-01-02 00:00:00	Power Plants	Gigawatt hours	97.945353	2021	1	January	Saturday	53	Winter
10	2021-01-02 00:00:00	All networked gas customers	Gigawatt hours	182.948983	2021	1	January	Saturday	53	Winter

(10 rows)

Figure 5: After Feature Engineering the dataset saved to PostgreSQL

Dataset 3:

Data preprocessing was carried out using Jupyter notebook. This included converting date columns into a consistent datetime format, handling missing values, dealing with duplicates, selecting relevant variables, merging two datasets for efficiency. The electricity and gas datasets were merged using the date field to create a unified processed dataset. New columns were created, such as monthly indicators and a COVID-period flag, to support time-series analysis. The final processed dataset was used for all further analysis and visualization.

```
(base) anuja@Anujas-MacBook-Air APDV_Project % cd ~/Desktop
tree -L 4 APDV_Project -I "venv|__pycache__"

APDV_Project
├── APDV_References
│   ├── A_parametric_analysis_of_domestic_electricity_consumption_patterns_in_Ireland.pdf
│   ├── Effects_of_the_COVID-19_Pandemic_on_Energy_Systems.pdf
│   ├── Increased_penetration_of_wind_generated_electricity_using_real_time_pricing_amp_demand_side_management.pdf
│   └── On-site_energy_generation_in_the_manufacturing_industry_effects_on_the_variability_of_the_net_electricity_demand
.pdf
├── data
│   ├── processed
│   │   └── final_data.csv
│   └── raw
│       ├── Electricity.csv
│       ├── README.md
│       └── Supply.csv
├── Database_Evidence.docx
├── docker-compose.yml
├── notebook
│   ├── APDV_Code_File.ipynb
│   ├── APDV_Code_File.nbconvert.ipynb
│   ├── net_generation.gif
│   └── wind_vs_gas.gif
├── README.md
├── requirements.txt
└── scripts
    ├── create_tables.py
    ├── db_config.py
    ├── export_final_data.py
    ├── load_raw_data.py
    ├── pipeline
    │   ├── dagster_pipeline.py
    │   ├── pipeline_diagram.png
    │   ├── README_pipeline.md
    │   └── Run_details.png
    ├── process_data.py
    └── verify_queries.py

8 directories, 26 files
(base) anuja@Anujas-MacBook-Air Desktop %
```

Figure 6: Final folder structure

Data analysis and visualization

Dataset 1:

Visualizations:

After preprocessing the Exploratory data analysis was performed to identify the trends, distributions and the sector differences in gas demand. Whereas the Key Analysis and the Insights were:

A) Trend of daily gas demand:

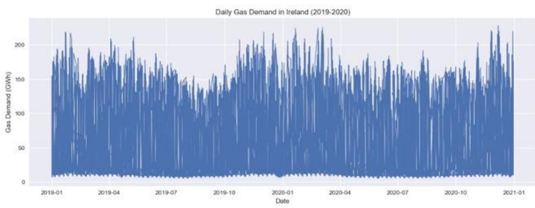


Figure 7: Daily Gas Demand in Ireland (2019-2020)

The output of the line chart displays Ireland daily gas demand (GWh) from 2019 to 2020. The daily gas demand shows the strong fluctuation. Each and every point represents gas consumption for every single day throughout the two years. We can see that higher demand is clearly seen during winters (2019-10) and (2020-10) and the lower demands during summer (2019-04) and (2019 -04) this pattern shows a seasonal dependency of gas consumption in Ireland

B) Compare customer type means average demand per customer category:

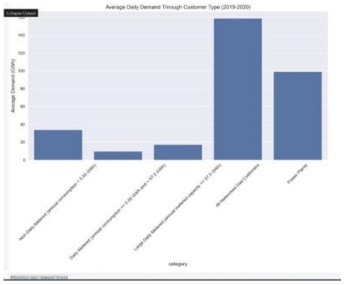


Figure 8: Average Daily Demand Through Customer Type

The output of the average daily demand by customer type shows that all network gas customers show the highest average demand. Power plants are the second largest daily gas consumers, whereas daily metered and non-daily metered customers consume significantly less gas, while large daily customers show moderate but consistent demand. This highlights that large scale consumers dominate overall gas usage.

C)Monthly gas demand trend:

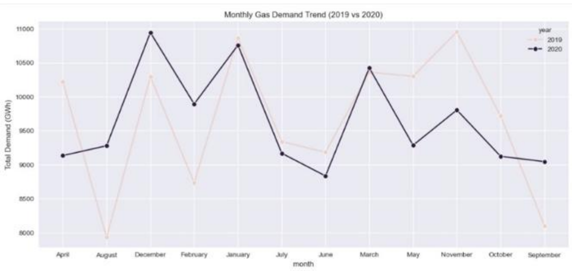


Figure 9: Monthly Gas Demand Trend (2019 vs 2020)

Monthly gas demand trend of (2019 vs 2020) of both years, which shows the similar seasonal trends that pick during the colder months. Winter month, such as January, November and December shows the highest demand in both years where as summer months shows lower gas usage. 2020 follows the similar pattern to 2019 with some variations likely impact by early COVID effects.

D)which category contributes the most:

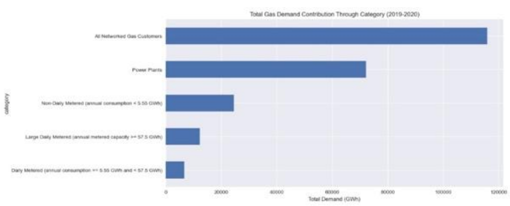


Figure 10: Total Gas Demand Contribution Through Category

The total gas consumption calculates across 2019 to 2020 for each customer category. From the output, we can see that all network gas customers contribute the highest of total demand, whereas PowerPoint contribute to the second last to total consumption and the smaller customer categories share less amount of total demand of gas consumption. So, this tell that a few sectors are responsible for most gas demand.

E) Heatmap category vs month:

Figure 11: Average Demand through Category and Month.

Heat map is used for category versus month where we can see. Dark colors show the high demand during winter months over main category types. where we can see the power plans and network customers shows continuous high usage throughout the year. light colors show low demand. The heat map clearly visualizes the seasonal category wise variation.

F) Gas consumption by weekday:

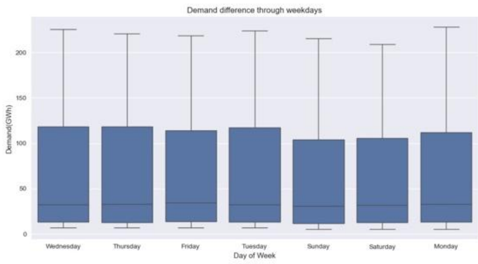


Figure 12: Demand difference through weekdays

This shows the distribution of gas demand over different days of the week. The gas remains consistent over weekdays. It's rightly lower demand on weekdays, especially for Sunday and Saturday. This tells that industrial and commercial impact the daily demand.

G) Total demand per year:

The output of the total gas demand shows the total gas demand in 2019 and 2020 is very similar. As there is no major decline seen in 2020, which tells that pre-Covid demand is stable. This supports the use of 2019-2020 as a dependable baseline period.

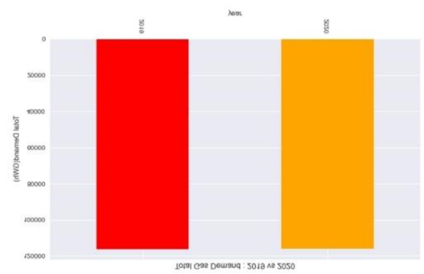


Figure 13: Total gas demand from 2019-2020

H) Category trend line plot:

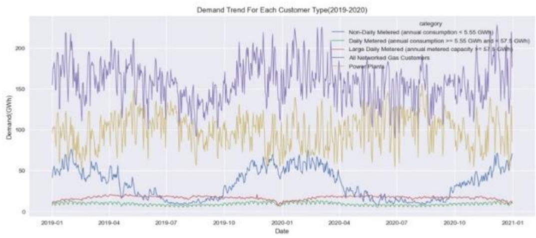


Figure 14: Demand Trend for Each Customer Type

The output of the daily demand train for each customer type shows that the all-customer types show clear seasonal fluctuation. The power plants and networked customers shows higher inconsistency. Daily metered and non-daily metered customer category shows the stable and low demand.

I) Forecasting for next 30 days was used with the help of ARIMA model:

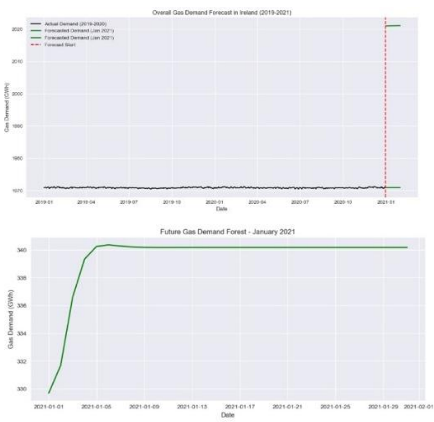


Figure 15: Forecasting for next 30 days

The output of the forecast shows the overall gas demand with the help of Arima model, which provide a stable gas demand for January 2021. The forecast follows the past average values without extreme fluctuations. Where the output shows forecast stable after initial days. The straight-line behaviour in the output shows the Arima smoothing nature. This tells the expected continuous demand in the short term. Where is the Arima focus on Trend and seasonality.

Dataset 2:

Visualizations:

After preprocessing the Exploratory data analysis was performed to identify the trends, distributions and the sector differences in gas demand. Whereas the Key Analysis and the Insights were:

A) Total Gas consumption by customer types (Bar Chart) :

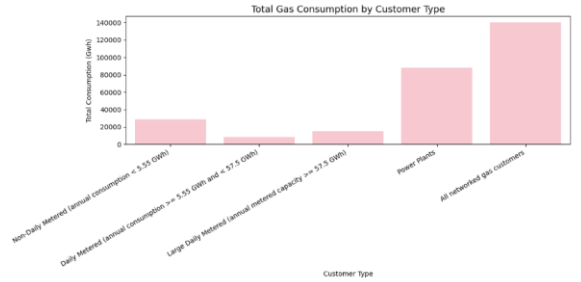


Figure 16: Total Gas Consumption by Customer Type

The output shows that the customer type non-daily metre, such as small consumers like homes, small businesses. where gas use is slow, which changes slowly over time. The daily meter, means, medium sized industries or business which uses more gas and commercial customers which measure every day because consumption is higher and more variable. The Large delimiter are big industries customers like factories with very high gas usage. Power plants mean electricity-generation plants that is highest gas consumption. Huge peaks in winter where when more energy is needed. And all networked gas customers is the total consumption across all customer types combined. Here we can see that power plants is showing very high consumption, large daily metered shows moderate and the small consumers the lowest.

B) Daily gas consumption Trend by Customer Type (Time-series Line Chart):

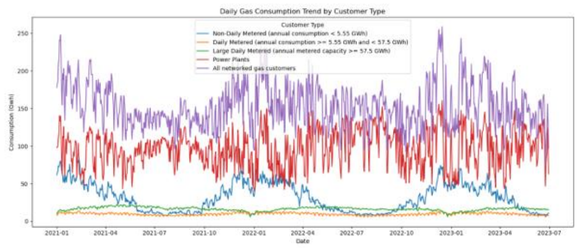


Figure 17: Daily Gas Consumption Trend by Customer Type

The output shows that the time changes over for daily gas consumption to customer type. Where Power Plants which is shown by red line i.e much higher than all other customer types, also highly fluctuating in winter months, showing the sharp peaks during colder seasons (Dec-Feb). Here the Non-Daily Metered consumers which shows the blue line tell that it is higher in winter and lower in summer. where now we can say that residential demand is predictable according to weather. The Daily Metered and Large Daily Metered which is shown in orange and green colour it represents medium and large industries, which shows the stable and low fluctuation and also it doesn't show extreme peak. They are steady because the operations run continuously and they are not strongly influenced by seasonal temperature.2021-2023 winter (DEC-FEB) it shows the highest demand and in summer (Jun-Aug) shows the lowest demand.

C) Monthly Gas consumption Analysis (Line plot):

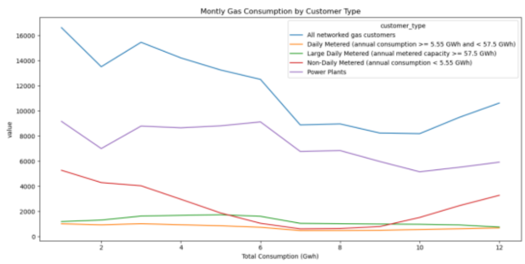


Figure 18: Monthly Gas Consumption by Customer Type

The output shows that the how gas consumption changes month-to-month for each customer type, The x-axis shows the numbers 1 to 12 that are month and y-axis shows the how much gas (gigawatt hr) each customer type used during that month. A networked gas customers with blue line shows the highest because it shows the total gas usage across all customers. Where we can see a peak in January and a drop in Summar and rise towards winter. Power plant which is in purple colour use a very large amount of gas. Where we see the very high in January and drop in summer with rise in winter ,Non-Daily Metered(small consumers) which is in red colour use less gas than industries and high in winter and low in summer , Non-Daily Metered(small consumers) which is in red colour use less gas than industries and high in winter and low in summer.

D) Seasonal Gas Consumption Distribution(Box plot):

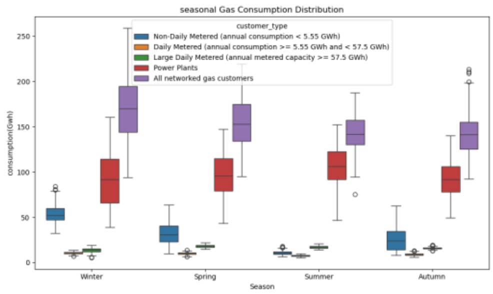


Figure 19: Seasonal Gas Consumption Distribution

The output shows that power plant in red boxplot have tall boxes in winter where we can say that power plants are the dominant contributions to Ireland's gas demand. All Network boxes in purple shows the aggregated customers like industrial + commercial + household. which shows very high values and strong seasonality and here the winter peaks nearly matching power plants. Daily metered and large Daily metered customers shows the average usage which is in orange and green colour which shows winter with highest values, summer with lowest and much smaller range than power plants. Non-daily metered in blue like which are mostly household shows the lowest consumption with very low medians in winter increases slightly and in summer almost it is flat where households use less gas than industry/power plants. now we can say that Seasonal Variation is Strongest for Power plants & All networked Customers where residentials users have very small consumption of gas and overall low values.

E) Weekday pattern of Gas consumption(Grouped Bar Chart):

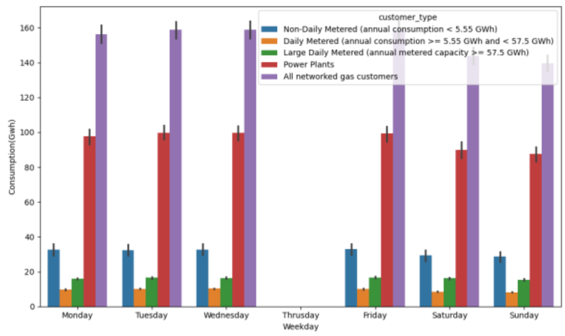


Figure 20: Weekday pattern of Gas consumption

The bar chart shows that the gas consumption is higher in weekdays than weekends because industries reduce operations on weekends, whereas also the offices, commercial buildings operates less and power plants bring down electricity demand. This is in shot strong behavioural pattern medium and large industries reduce the production or we can say shift load during weekends. And for Thursday which shows the empty bar, it means that in the dataset there is no any data for the day Thursday so is shows the no bar. Households which hold the blue bar is most stable across the week around 30 to 35 Gwh every day. as it is consistently use for cooking, heating, hot water demand stays constant.

F) Finding the top 10 highest consumption records:

	id	day	customer_type	unit	value	year	month	month_name	weekday	week	season
3569	3570	2022-12-15	All networked gas customers	Gigawatt hours	258.821017	2022	12	December	Thursday	50	Winter
3539	3540	2022-12-09	All networked gas customers	Gigawatt hours	254.821945	2022	12	December	Friday	49	Winter
3554	3555	2022-12-12	All networked gas customers	Gigawatt hours	248.479445	2022	12	December	Monday	50	Winter
39	40	2021-01-08	All networked gas customers	Gigawatt hours	248.034477	2021	1	January	Friday	1	Winter
3744	3745	2023-01-19	All networked gas customers	Gigawatt hours	243.272513	2023	1	January	Thursday	3	Winter
29	30	2021-01-06	All networked gas customers	Gigawatt hours	240.853722	2021	1	January	Wednesday	1	Winter
3534	3535	2022-12-08	All networked gas customers	Gigawatt hours	240.553187	2022	12	December	Thursday	49	Winter
3564	3565	2022-12-14	All networked gas customers	Gigawatt hours	240.097329	2022	12	December	Wednesday	50	Winter
34	35	2021-01-07	All networked gas customers	Gigawatt hours	239.993926	2021	1	January	Thursday	1	Winter
3574	3575	2022-12-16	All networked gas customers	Gigawatt hours	237.183444	2022	12	December	Friday	50	Winter

Figure 21: Top 10 highest Consumption Records

The output of the Table shows the top 10 days with the highest gas consumption records in the dataset, where this all records belongs to "All networked gas customers, which represents the total gas demand across all customer types combined. Winter month has the highest gas usage mainly in the month of December and January. The values of Gas consumption peak the range from about 237 to 259 GWh, which are the maximum values observed in the data. Most of the high-consumption falls on weekdays, showing the higher energy demand at the time of regular working days. This tells that winter season and colder weather strongly increase overall gas demand.

G) Heatmap for monthly gas consumption by customer type:

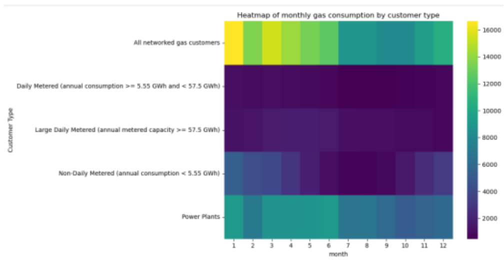


Figure 22: Heatmap for monthly gas consumption by customer type

The output of the heatmap shows the monthly gas consumption patterns for different customers types. Whereas the colours show the real meaning of it. Bright colours like Yellow /light green this colour shows very high consumption of "All networked gas customers" which shows the total gas consumption is highest specially in winters months like January to march. whereas Green/teal colours represent the medium gas consumption shows high gas usage with darker shades in summer which shows the reduced demand. The Dark colours like Dark blue represents low gas consumption for non-daily metered customers which is household showing the average consumption higher in winter and lower in summer. The dark purple for Daily Metered and Large Daily Metered customers shows the lower consumption throughout the year with the seasonal variation.

H) share of Gas Consumption by sector (Pie chart):

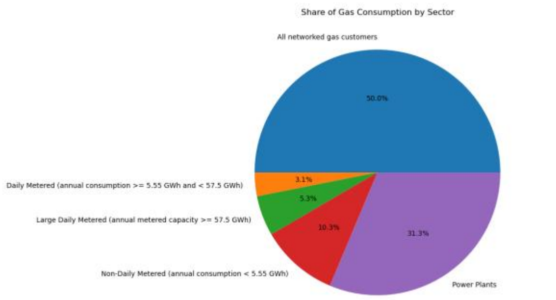


Figure 23: Share of Gas Consumption by sector

The output of the pie chart shows the share gas consumption across different sectors. Where the "All-networked gas customers" take the largest share of about 50% that means half of the total gas demand comes from all customers combined. The second-largest consumer is power plant using about 31% of the total gas, which shows the major role in electricity generation. The 10 % of the shared shown by the non-daily metered customers which are households and small users which is much lower than industrial and power generation use. The Large daily metered customers that is large industries contribute around 5 % of total gas consumption. The smallest share 3 % of Daily metered customers that is medium sized business.

I)Forecast future Gas Demand:

Auto-Arima model was used due to its ability to automatically recognize the optimal model parameters and handles the seasonality.

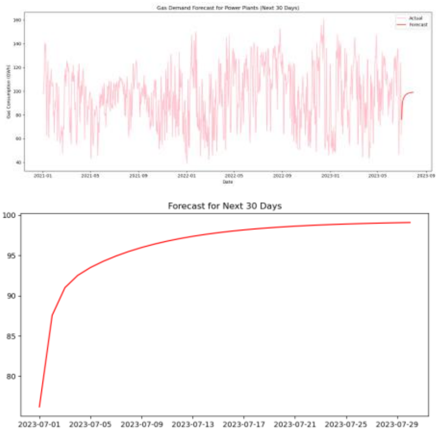


Figure 24: Gas Demand forecast for next 30 days

The output of the forecast shows the actual gas consumption of power plants over time and a forecast for the next 30 days. The line in pink colour shows the actual gas consumption, which changes a lot due to changing in the electricity demand. The forecast of gas demand for the next 30 days is shown by the red colour line. The conclusion of the forecasting is that it advice about the power plant gas consumption will remain stable, with no sudden increase or decrease in future.

Dataset 3:

Exploratory data analysis (EDA) was conducted to examine trends, variability, seasonality and distributions in electricity generation and gas supply. Time-series analysis was used to identify seasonal and long-term patterns. Correlation analysis was applied to explore relationships between gas supply, total electricity generation, and individual fuel sources. A contribution strength analysis was performed to assess which fuel sources were most strongly associated with total electricity generation. Comparisons between pre-COVID and COVID periods were also carried out to examine whether noticeable deviations from established patterns were observed.

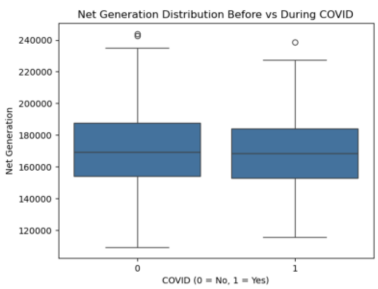


Figure 25: Before vs. After COVID

This boxplot compares the distribution of daily net electricity generation in Ireland before and after COVID-19 period. Overall, this figure suggests that the pandemic did not lead to a major disruption in Ireland, despite the ongoing demands and economic activities during the lockdown periods.

Figure 26: Time-series (with COVID-19 shading)

The time-series plot shows daily net electricity generation in Ireland over the entire time-period available. A clear seasonal pattern is visible, with higher levels of electricity generation occurring during winter months like January, February, November and December while, lower levels during summer months like May to August, indicating strong seasonal effects in electricity generation. In the shaded part of the time-series, we can see the COVID-19 period, with no clear structural break or long-term deviation from the existing trend. Electricity generation during the COVID period continues to follow the same seasonal pattern and variability seen in the pre-COVID and post-COVID periods. The short-term fluctuations are present, these are consistent with normal day-to-day activities, rather than a disruption caused by the pandemic. Overall, this figure suggests that although COVID-19 was a disruptive change in our lives, the electricity generation system remained stable. Seasonal patterns and long-term trends persisted throughout the COVID period, indicating resilience in electricity generation despite changes in demand and broader economic conditions.

	Fuel Source	Correlation with Net Generation
0	Wind	0.664519
1	Renewable Hydro	0.490167
2	Oil	0.057479
3	Coal	0.029626
4	Gas	-0.225064
5	Solar	-0.284505

Figure 27: Contribution Strength of fuel sources

The above table shows the correlation between individual fuel sources and total net electricity generation. We can conclude that renewable sources, particularly wind, is strongly linked to total electricity generation, while gas plays as a back-up source.

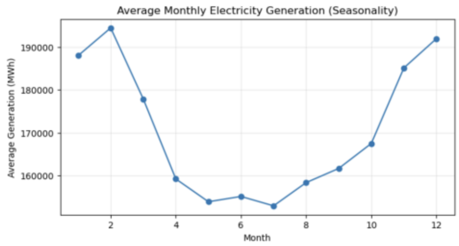


Figure 28: Monthly electricity generation

This line chart shows the average monthly electricity generation in Ireland and highlights clear seasonal patterns. Electricity generation is highest during the winter months, particularly in January, February, November, and December. This likely reflects increased energy demand for heating during colder periods. Generation levels gradually decline during the spring season and reach their lowest values in the summer months, especially between May and July. This indicates reduced demand during warmer periods, along with possible changes in fuel usage. These results align with the predictable climate behavior.

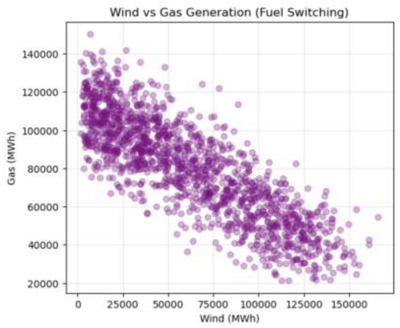


Figure 29: Wind vs. Gas

This scatterplot highlights the important balancing role of gas in Ireland's power system and supports the view that gas adjusts to changes in energy availability rather than driving total generation levels.

Overall, the exploratory data analysis provided a clear overview of electricity generation in Ireland across the study period. The analysis revealed strong seasonal patterns, with higher generation during winter months and lower levels during summer. Renewable sources, mainly wind and hydro, showed strong relation with total electricity generation, while gas generation displayed clear balancing behavior in response to renewable source. Comparisons between pre-COVID and post-COVID periods indicated that total electricity generation remained broadly stable, with no major structural changes observed. Time-series and distribution-based visualizations further suggested that short-term fluctuations during the COVID period were consistent with normal operational functionality rather than sustained disruption. These exploratory findings address the key research questions by establishing baseline relationships, trends, and seasonal effects.

ETL pipeline (Dagster)

An ETL (Extract-Transform-Load) pipeline was implemented to demonstrate workflow automation and reproducibility. Dagster was used to show the pipeline and ensure that each step was executed in the correct order. (See at appendix for screenshots of Dagster)

Results and Evaluation

Dataset 1:

The study shows that Ireland's gas consumption from 2019 to 2020, which follows a clear seasonal pattern with the higher demand during winter months and the lower demand during summer. Whereas all network gas customers and power plants contributes the largest share of total demand over the customer types.

The monthly and weekday analysis tells that demand remains stable across weekdays as on the weekends, such as Sunday and Saturday shows reduction. Comparison between 2019 and 2020 shows the equal total gas demand, which does the stable pre-consumption patterns.

For a stable short-term demand train for January 2021, the Arima forecasting model was predicted, which provides a useful baseline forecast. finally, the project successfully meets with its objectives by integrating database storage, data processing, visualisation and forecasting to generate essential information into Ireland gas demand.

Dataset 2:

The study of the project shows that Daily gas consumption varies over customer types, where the power plants and industrial users occupy the most gas. So, the demand becomes more regular, telling the changes in how gas is used. The result of the forecasting for the next 30 days shows that the gas demand will continue to change, showing the need for flexible energy planning.

Finally, the project achieves its goal by building a complete data analytics pipeline and generating the useful information. There were some challenges such as understanding, docker and pose, Grace equal integration as well as matching the date across datasets, handling time, series data, and designing a compact single page dashboard. Whereas the learning from the project was real database integration, the end to end data pipeline learning, practical stream, dashboard, learning, and to know more about industry relevant data analytic workflow.

Dataset 3:

Analysis Area	**Method Used**	**Key Observations**	**Evaluation / Interpretation**
Electricity generation trends	Time-series analysis	Clear cyclical patterns observed across years, with higher generation in winter and lower generation in summer	Indicates strong seasonal behaviour and stable long-term trends in electricity generation
COVID-19 period comparison	Boxplots and time-series with COVID shading	Median generation levels and variability remain similar before and during COVID	Suggests electricity generation in Ireland was resilient, with no major disruption during COVID
Gas supply and electricity relationship	Correlation analysis	Gas shows weaker or negative correlation with renewable sources	Indicates gas acts as a balancing source rather than a primary driver of total generation
Fuel switching behaviour	Scatter plot (wind vs gas)	Higher wind generation corresponds to lower gas generation	Demonstrates operational fuel switching to balance renewable variability
Fuel contribution strength	Contribution strength analysis	Wind and renewable hydro show strongest association with total generation	Confirms the dominant role of renewables in the electricity generation mix
Seasonal effects	Monthly aggregation analysis	Higher average generation in winter months, lower in summer	Reinforces predictable seasonal demand patterns and the need for flexible generation

Figure 30: Summary table

Overall, the results show that electricity generation in Ireland is characterized by strong seasonality, a significant contribution from renewable sources and power system stability during the pandemic.

Limitations

This analysis is exploratory and focuses on identifying patterns rather than proving cause-and-effect relationships. The dataset does not include weather conditions or policy factors, which limits the ability to fully explain changes in electricity generation. In addition, the data is analyzed at a daily level, so short-term operational changes are not captured.

Conclusion

Dataset 1:

Gas daily demand consumption in Ireland varies essentially over customer category type through all network gas customers and power plants contributing the most demand. This project shows clearly shows seasonal and time-scale patterns were observed in daily and monthly gas used during both years. The study successfully implement's a full data analytics pipeline from MongoDB storage to visualization and Arima forecasting.

Dataset 2:

The project successfully analysed gas consumption patterns after COVID-19 according to theme. The post -COVID demand of gas consumption shows the increase differences across customer types. The study also shows that the gas demand differs by sector with power generation and large industrial customers coming up with the most overall gas consumption. The use of real PostgreSQL database, Docker deployment, and a stream lit dashboard showing an end -to-end analytics and visualisation pipeline.

Dataset 3:

This project explored electricity generation and gas supply in Ireland using an exploratory, data-driven approach. The analysis showed clear seasonal patterns, a significant contribution from renewable energy sources and the role of gas as a back-up fuel. Across the COVID19 period, the electricity remained stable, demonstrating stability in the Irish power generating system.

Future Work

Applying advanced forecasting models like SARIMA, LSTM for improving the prediction accuracy. Integrate weather and temperature data to improve demand prediction accuracy.

To extend the Auto-Arima forecasting model by comparing it with other time-series models like SARIMA or Prophet. To apply the forecasting at the customer-type level. Enhancing the streamlit dashboard with real-time data updates and advanced interactivity. Also expanding the analysis which will additional energy sources.

Future work could include additional data such as electricity demand and weather variables to better explain generation patterns. More advanced statistical or forecasting models can be applied to improve the analysis. Using higher quality data may provide deeper insight into power system behavior.

References:

[1] Liu, Jinyuan, et al. “Natural Gas Consumption Forecasting: A Discussion on Forecasting History and Future Challenges.” Journal of Natural Gas Science and Engineering, vol. 90, June 2021, p. 103930, https://doi.org/10.1016/j.jngse.2021.103930. Accessed 15 Dec. 2025.

[2] Works CitedBianco, Vincenzo, et al. “Analysis and Future Outlook of Natural Gas Consumption in the Italian Residential Sector.” Energy Conversion and Management, vol. 87, Nov. 2014, pp. 754–764, https://doi.org/10.1016/j.enconman.2014.07.081. Accessed 15 Dec. 2025.

[3] Ciais, Philippe, et al. “Impact of Lockdowns and Winter Temperatures on Natural Gas Consumption in Europe.” Earth’s Future, vol. 10, no. 1, Jan. 2022, https://doi.org/10.1029/2021ef002250. Accessed 15 Dec. 2025.

[4]Rokicki, Tomasz, et al. “Changes in Energy Consumption and Energy Intensity in EU Countries as a Result of the COVID-19 Pandemic by Sector and Area Economy.” Energies, vol. 15, no. 17, 26 Aug. 2022, p. 6243, https://doi.org/10.3390/en15176243. Accessed 15 Dec. 2025.

[5]Cihan, Pınar. “Impact of the COVID-19 Lockdowns on Electricity and Natural Gas Consumption in the Different Industrial Zones and Forecasting Consumption Amounts: Turkey Case Study.” International Journal of Electrical Power & Energy Systems, vol. 134, Jan. 2022, p. 107369, https://doi.org/10.1016/j.ijepes.2021.107369. Accessed 15 Dec. 2025.

[6]Works CitedAli, Anis, et al. “Pre and Post-COVID-19 Financial Performance of Oil and Gas Companies: An Absolute and Relational Study of Financial Variables.” International Journal of Sustainable Development and Planning, vol. 18, no. 1, 31 Jan. 2023, pp. 69–77, https://doi.org/10.18280/ijsdp.180107. Accessed 15 Dec. 2025.

[7] F. McLoughlin, A. Duffy and M. Conlon, "A parametric analysis of domestic electricity consumption patterns in Ireland," 2011 10th International Conference on Environment and Electrical Engineering, Rome, Italy, 2011, pp. 1-4, doi: 10.1109/EEEIC.2011.5874814.

[8] P. Finn, C. Fitzpatrick and M. Leahy, "Increased penetration of wind generated electricity using real time pricing & demand side management," 2009 IEEE International Symposium on Sustainable Systems and Technology, Tempe, AZ, USA, 2009, pp. 1-6, doi: 10.1109/ISSST.2009.5156783.

[9] A. Sgobba, C. Meskell and G. Manzolini, "On-site energy generation in the manufacturing industry: effects on the variability of the net electricity demand," CIRED 2020 Berlin Workshop (CIRED 2020), Online Conference, 2020, pp. 717-720, doi: 10.1049/oap-cired.2021.0202.

[10] Navon, Aviad, et al. “Effects of the COVID-19 Pandemic on Energy Systems and Electric Power Grids—a Review of the Challenges Ahead.” Energies, vol. 14, no. 4, 17 Feb. 2021, p. 1056, https://doi.org/10.3390/en14041056. Accessed 15 Dec. 2025.

[11] Data Source: Central Statistics Office.

Data.cso.ie, 2020, data.cso.ie/.

Appendix:

Stream lit Dashboard:

The Stream lit dashboard was created to visualise the analysis of gas and electricity consumption trends before and after the COVID-19 period. For this, the cleaned and pre-process, data set, which is originally extracted from oppose race, sequel database that was stored in CSV and Jason format to support efficient data loading. The dashboard uses the python libraries such as panda’s, matplotlib and seaborn to perform time-based calculation, customer type comparison, and distribution analysis. The gas consumption data further was integrated with electricity generation data to study the relationship between gas demand and power production. The final dashboard presents information through interactive visualisation, showing the clear comparison of trends over time, periods and customer categories which make sure about fast and reliable performance.

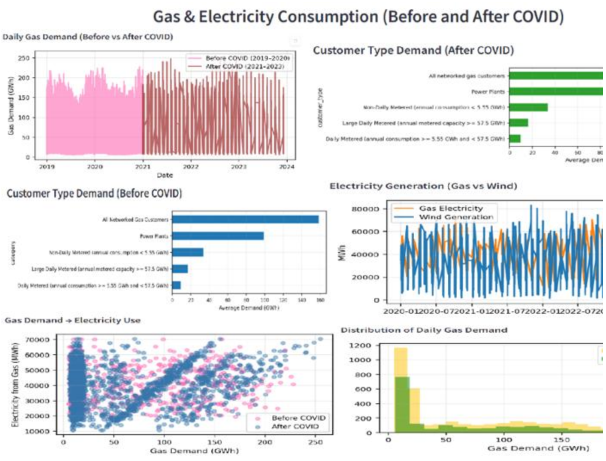


Figure 31: Stream lit Dashboard

1)Daily gas demand (Before vs After Covid):

It is a time comparison of daily gas demand for two periods before Covid vs after Covid .The output shows that gas demand changes over time and shows the clear fluctuations as we can point.

2)Customer Type Demand (Before COVID):

Customer type demand shows the horizontal bar chart of average demand by customer category for before Covid data set. It explains the answer to the question part, which customer type contribute the most so the longest bar “All networked gas demand” contributing customer types before Covid

3) Customer type demand (After COVID):

Customer type demand shows the horizontal bar chart of average demand by customer category for after Covid data set. It explains the answer to the question part, which customer type contribute the most after COVID . so the longest bar “All Networked Gas Demand” contributing customer types after Covid

4)Gas Demand with Electricity Use:

Gas demand and electricity use which shows a scatterplot comparing gas demand with electricity from gas usage. The output of this explains if a point trends upward, it tells higher gas demand is related with higher electricity generation from the gas. The two colours show whether the relationship changed after Covid.

5)Electricity Generation (Gas vs Wind):

It shows two times series lines, electricity from gas vs wind. The output shows how gas was electricity behaves compare to wind and it supports interpretation like when wind varies gas me act as balancing source .

6) Distribution of Daily Gas Demand.

It shows the histogram distribution for daily gas demand before versus after Covid. Hear the output shows the shape of demand. If one period has a wider spread. There is more viability in that period.