Data Science and Advanced Programming

Zildete Landim Crausaz

November 8, 2025

Abstract

The oil market plays a crucial role in supporting global energy needs in sectors such as transportation and the generation of electricity. As such, it is important for businesses and governments to keep records to track the consumption of petroleum products for better planning in production to maintain a balance between demand and supply in order to avoid overproduction or shortage. However, it is challenging to find consolidated data or analyses that businesses and governments could directly use to make decisions, such as pricing. Thus, they need to conduct further analyses to identify trends that can be used to predict future consumption. This project, therefore, analyses one of the petroleum products, gasoline, using data from the Joint Organizations Data Initiative (JODI) website to investigate factors such as volatility, correction, trends, and future forecast of demand and supply of 25 countries in Europe, as well as three key regions: Amsterdam-Rotterdam-Antwerp (ARA), Mediterranean (MED), and Northwest Europe (NWE). To achieve this, the project implements a Python-based data analysis pipeline developed in Visual Studio Code and version-controlled using GitHub.

Keywords: Gasoline, supply, demand, European markets.

Table of Contents

Demand-Supply Analysis for Gasoline

Introduction

Gasoline is a refined product of petroleum, which is mostly used in cars, motorcycles, and small aircraft as a fuel for internal combustion engines. Its production occurs through the fractional distillation and catalytic reforming of crude oil, where the oil is heated and divided into fractions depending on the boiling points. Galosine emerges as a lighter hydrocarbon with 4 to 12 carbon atoms. Today, gasoline usually contains additives such as ethanol to improve the performance of engines and reduce emissions in terms of unburnt carbons into the environment.

The European gasoline market is one of the most active segments of the global energy system, accounting for 14% of the worldwide demand as of 2024 (Fuels Europe 1). Even though Europe, just like the rest of the regions in the world, is gradually transitioning toward cleaner sources of energy for the transport system, gasoline still has a huge share when it comes to consumption in vehicles and aircraft of refined oil products. Gasoline is one of the world's most traded and consumed energy commodities. Its usage is driven by the high demand in the transportation sector, which is experiencing mobility growth in emerging markets like Poland, Czechia, Hungary, and Albania. On the supply side, the factor is affected by refinery configurations concentrated in the Amsterdam–Rotterdam–Antwerp (ARA), Mediterranean (MED), and Northwest Europe (NWE) regions. For a more comprehensive analysis of the demand and supply of galosine in these regions, this project will proceed by exploring the background and identifying gaps in the literature review, followed by a section on data sources and methodology. This will be followed by sections on design and architecture, implementation, evaluation, discussion, and conclusion.

Literature Review

Background

Some of the factors that gave rise to ARA, MED, and NWE hubs were geography, infrastructure, and the development of industries after World War II (late 1940s–1950s). Following the war, European nations rebuilt their energy logistics around deep-water ports that could handle the high volumes of imported crude oil and its refined products, as well as the trade within the regions (Van Den Berghe et al. 315). For instance, the ARA region was the main trading and refining hub due to expansions such as the Botlek industrial complex (1952–1957) and the Europoort terminal (1958–1964), which connected North Sea crude to continental markets via the Rhine River. During the 1970s and 1980s, ARA became the discovery point for natural prices for gasoline and other petroleum products as the oil market grew in Europe and worldwide (Notteboom et al.). Today, the ARA hub is the main storage and distribution centre in Europe. It connects seaborne imports from global suppliers with inland demand using a barge and pipeline system (Port of Rotterdam 3)

The MED hub also evolved during these periods as refineries and shipping lanes that linked Southern Europe, North Africa, and the Middle East, while Northwest Europe (NWE) became a wider commercial region. By the 1980s, financial markets had begun to formalise these hubs through official pricing benchmarks. For example, the International Petroleum Exchange introduced its gasoil futures contract in 1981, and the launch of the Eurobob Oxy Gasoline benchmark in 2010. Currently, this region links economies such as Italy and Spain that have heavy refineries with nations that depend on imports in the eastern basin, such as Greece, Turkey, and Croatia

The demand and supply for gasoline also shift depending on factors such as seasons. For example, during the winter months, NWE usually imports additional crude oil, including from the United States and the Middle East, to offset the demand for its products, such as gasoil, which is used for heating (Strautmann). Higher demand for gasoil alters the supply of gasoline, as refineries in these regions may prioritise the production of gasoil instead of gasoline, whose demand may be lower due to people being less mobile during winter. Concurrently, if there is an unplanned refinery outage, it could lead to a low supply of gasoline in any of the hubs, such as MED. This can lead to higher demand as traders and importers rush to find alternative sources of supply (Park 2).

Gap

Although there is data available for analysis of gasoline consumption in Europe, it is always a challenge to get consolidated datasets or analyses for demand and supply due to varying refinery capacities across different countries, differing fuel standards, and trade dependencies. Also, the differences in imports between the ARA, MED, and NWE regions make it difficult to understand how supply and demand interact to influence price movements (Gkatzoglou et al. 102). In addition, much of the data available on gasoline flows is not comprehensive, and it is mostly dispersed across different websites. The data is also not consistently integrated into analytical models for easier analysis. Therefore, the purpose of the project is to analyse European gasoline supply and demand patterns using JODI data to identify market imbalances and trends that inform decisions, such as an increase in refining of the commodity or changes in prices.

Data Sources and Methodology

Data Sources and Explanation

The data used in this project were obtained from the Joint Organisations Data Initiative (JODI). JODI is an international energy data collaboration, which was established by major global organisations such as the International Energy Agency (IEA), OPEC, Eurostat, and the United Nations Statistics Division (JODI). Its purpose is to improve the level of transparency as well as the consistency, as researchers can collect and harmonize official data from participating countries.

The dataset includes monthly statistics on petroleum products for 118 countries that reported their products from January 2002 to August 2025. The balance items include refinery output, receipts, imports, exports, products transferred, interproduct transfers, stock change, statistical difference, demand, and closing stocks (JODI). These balances represent products such as liquefied petroleum gas (LPG), naphtha, motor and aviation gasoline (the main concern in this project), kerosene, gas/diesel oil, fuel oil, other oil products, and total oil products. The units of measure are expressed in thousand barrels per day (kb/d), thousand barrels (kb), thousand kilolitres (kL), and thousand metric tons (kt).

JODI Oil Supply–Demand Data

Since the dataset only contains demand figures, the formula below, as proposed by Suchodolska (6), was used to calculate the supply values for each respective country.

$$\begin{aligned} Supply = {} & Refinery\ Output + Receipts + Imports - Exports \\ & - Product\ Transferred + Interproduct\ Transfers \\ & - Stock\ change \end{aligned}$$

Regional Aggregation Methodology

This approach converts the figures from JODI into comparable regional series for ARA, MED, and NWE. For this analysis, the ARA comprises two countries: Belgium and the Netherlands. MED consists of five major economies, including France, Greece, Italy, Spain, and Turkey. NEW also comprises Belgium, Denmark, France, Germany, Ireland, the Netherlands, Norway, Sweden, and the United Kingdom. Even though some countries may be in more than one region, the main purpose of this analysis is to compare gasoline demand and supply patterns across these distinct regional markets, without overlapping influences that might obscure regional differences.

Analytical Methodology

The workflow and analysis of the project will proceed through three main stages.

Data Cleaning and Validation

The raw Excel sheets were downloaded from the JODI website after filtering out the 43 European countries (refer to the consolidated data Excel dataset) for the period from February 2016 to August 2025. The data was then cleaned using Python to eliminate missing data, where only of 25 countries, Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, and United Kingdom, with complete data remained.

Design and Architecture

I started by establishing a clean, modular project architecture within VSCode, where I created different folders that were well labelled. I then organised the project into distinct directories for source code, data, results, and documentation. This method is suitable as I was able to manage each component separately. It also simplified the debugging process.

The system architecture was designed around three main layers: data management, analysis processing, and output generation. Firstly, I made sure I created a centralised data loader module to handle the clean data, as well as make sure data is loaded into the system for analysis. The analysis layer consists of modules for various market perspectives, including combined analysis for demand and supply balances and regional comparisons, as well as correlation, volatility, yearly, top player, and forecast analyses. Following this structure allowed me to develop and test each component independently.

I also decided to include a folder for the output system, by organising them into figures and tables for easier visualisation once the project is completed. I also embedded Python into VSCode through the code.

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from statsmodels.tsa.holtwinters import ExponentialSmoothing
import seaborn as sns
import os
```

And installed the following packages with the code:

```
pip install pandas numpy matplotlib openpyxl statsmodels seaborn
```

The packages installed include pandas for data manipulation and analysis, numpy for numerical computations, matplotlib for creating visualisations and charts, openpyxl for reading Excel files with gasoline data, statsmodels for time series forecasting and statistical analysis, and seaborn for enhanced data visualisation.

Before starting the actual project, I configured Git version control in VSCode to track all code changes and architectural design choices. The Source Control panel from VSCode allowed me to maintain a clean commit history, as well as document any refinements made during the project. In the final phase, I validated the VSCode setup against real trading scenarios and made any adjustments.

Implementation

I began the project by first creating a file named data_loader.py in the VSCode environment. I then wrote code to load the data, but the action was unsuccessful because the path to obtain the data from the original folder was unclear. Nevertheless, after identifying the correct path, through test_path.py, the data was successfully loaded into the system and was ready for analysis. I also created another t file with relevant code in test_setup.py, to test the setup.

Under the evaluation of the balance between supply and demand across countries and regions, I developed a script under the file name combined_analysis.py. To ensure that I have covered the analysis of the market comprehensively, I designed the code to display the top 15 countries based on their average supply-demand balance, over the entire 2016–2025 period, using the formula balance = supply-demand. This helped to identify markets with the largest surpluses and deficits. For regional analysis, that is, ARA, NEW, and MED, I wrote a script to compute the sum of their constituent countries' balances and display the results on graphs and in tables. I also made sure the visuals were within a figure size of (14, 10) to ensure the readability of long country names, but at the same time maintaining an appropriate level of data density. While scripting for the visuals output, I used colour-coding with red for deficits (balance < 0) and green for surpluses (balance > 0).

For the subsequent evaluations, I applied a similar approach in correlation analysis, that is, the Pearson Correlation Coefficient, where I wrote the code to examine both the strength and direction of demand-supply relationships across European markets. The evaluation uses full-time series data from 2016 to 2025 with the formula corr = demand_country.corr(supply_country) to measure how closely supply adjustments follow demand fluctuations over the period. The visual outputs are expected to give the top 10 positive correlations (near +1.0) and top 10 negative ones (near -1.0) on bar charts with a size of (14, 6). The scripts also rank the countries by the strength of their correlation.

For the volatility, I formulated a script to generate the instability of both demand and supply using the coefficient of variation (standard deviation divided by the mean). I used the formula volatility = data.std(axis=1) / data.mean(axis=1) for demand and supply datasets, respectively. The expectation of the analysis is to display the top 10 most volatile countries. One of the key decisions I made here was to write code to filter out countries with near-zero mean values, to ensure that comparisons are meaningful with complete and relevant data. The code was also designed to calculate the overall market averages (avg_demand_vol and avg_supply_vol) to identify which side of the market, demand or supply, shows greater unpredictability across Europe.

I also carried out other analyses, including top players, to reveal the countries with the highest demand and supply among the 25 countries. This was followed by analysing the balances in each year to identify which period had the highest level of gasoline usage from 2016 to 2017. The final code was written to calculate the supply and demand forecast to determine what the market would look like by the end of 2025 and in 2026. However, I was able to overcome this issue by systematically debugging each file, isolating errors, and using logging to track data flow.

Evaluation

Before discussing and accepting the results of the project, I cross-checked the outputs of the analyses from VSCode against manual calculations in Excel to confirm if they were accurate. The calculations within VSCode proved more precise, meaning the scripts produced the expected outcome.

Results

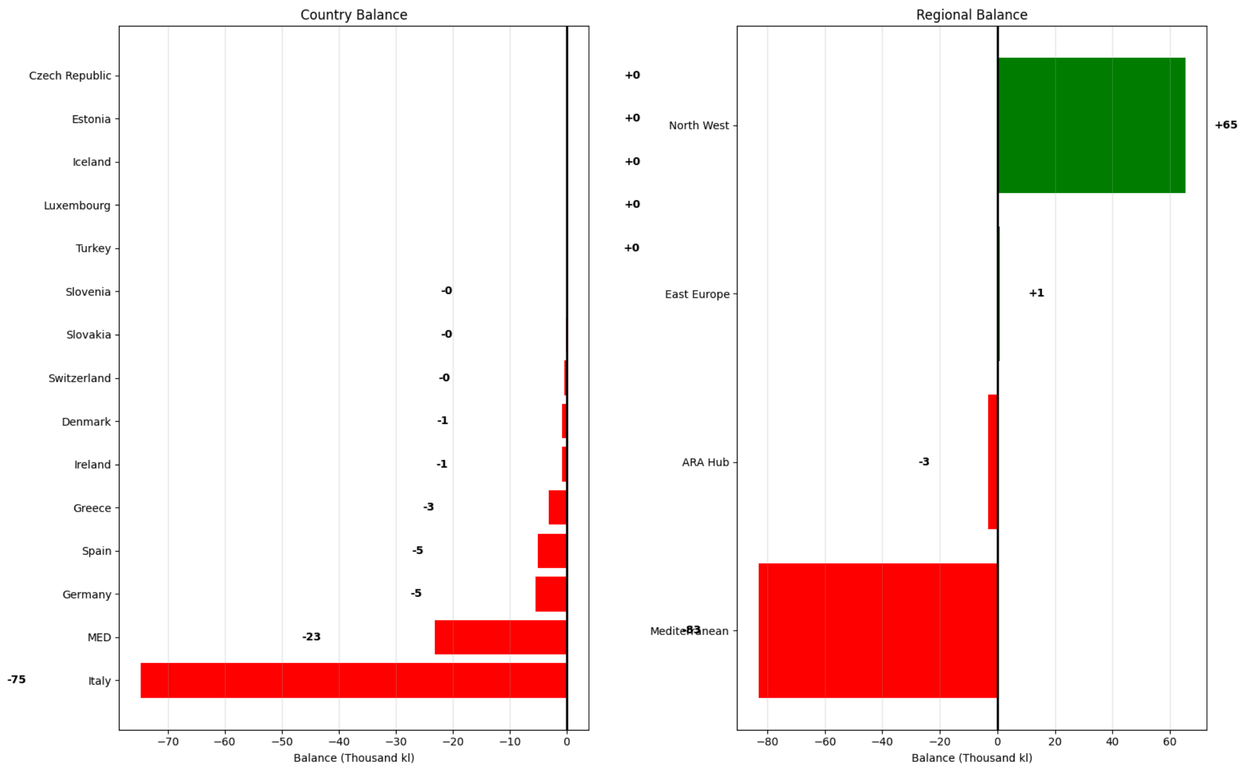

Fig 1: Graph of Combined Analysis: Demand and Supply Balances

Fig. 1 shows the gasoline supply and demand balances across 25 countries and ARA, NWE, and MED regions. Italy has the largest deficit of around 75000 kl, followed by Germany and Spain, each with -5000 kl. MED has the largest deficit, -83,000kl, among the regions. However, the evaluation shows that from 2016 to 2025, NWE had a positive balance, with 65,000kl.

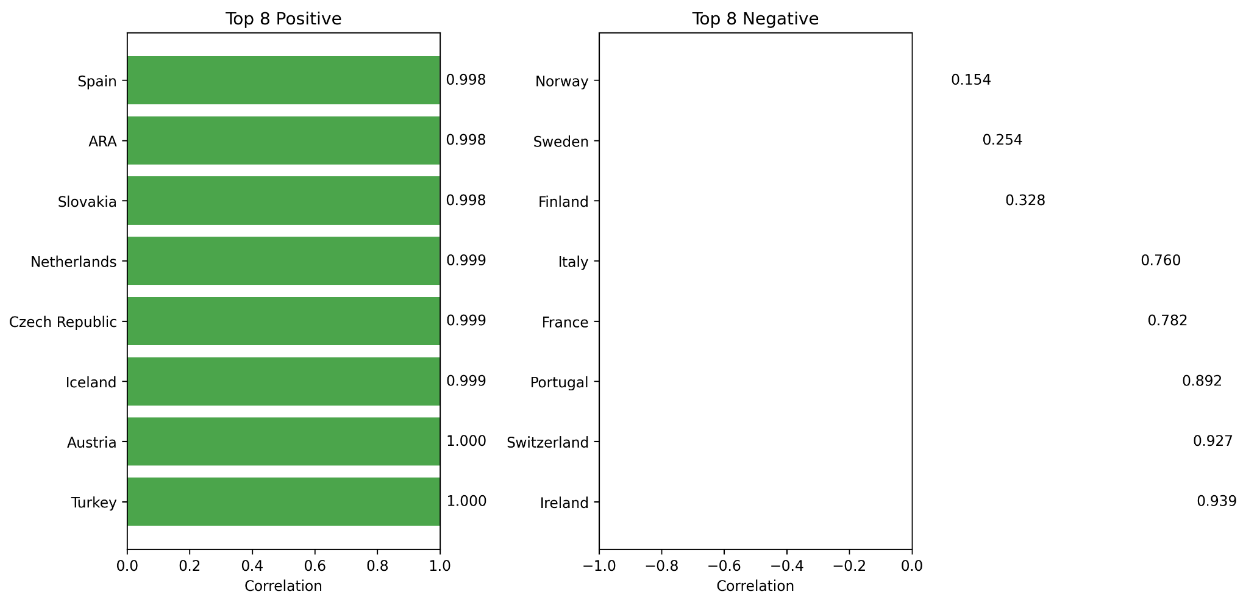

Fig. 2: A Graph of Correlation Analysis Results

Fig. 2 shows the correlation analysis of gasoline demand and supply across European markets. The left side of the figure displays the Top 8 Positive Correlations, with countries such as Austria, Turkey, and Switzerland exhibiting a near-perfect positive correlation (values close to 1.0). This indicates that their gasoline supply closely tracks demand. On the right side, the Top 8 negative correlations show countries with weaker or inverse relationships between supply and demand, including Italy and France, with lower correlation values such as 0.760 and 0.782, respectively.

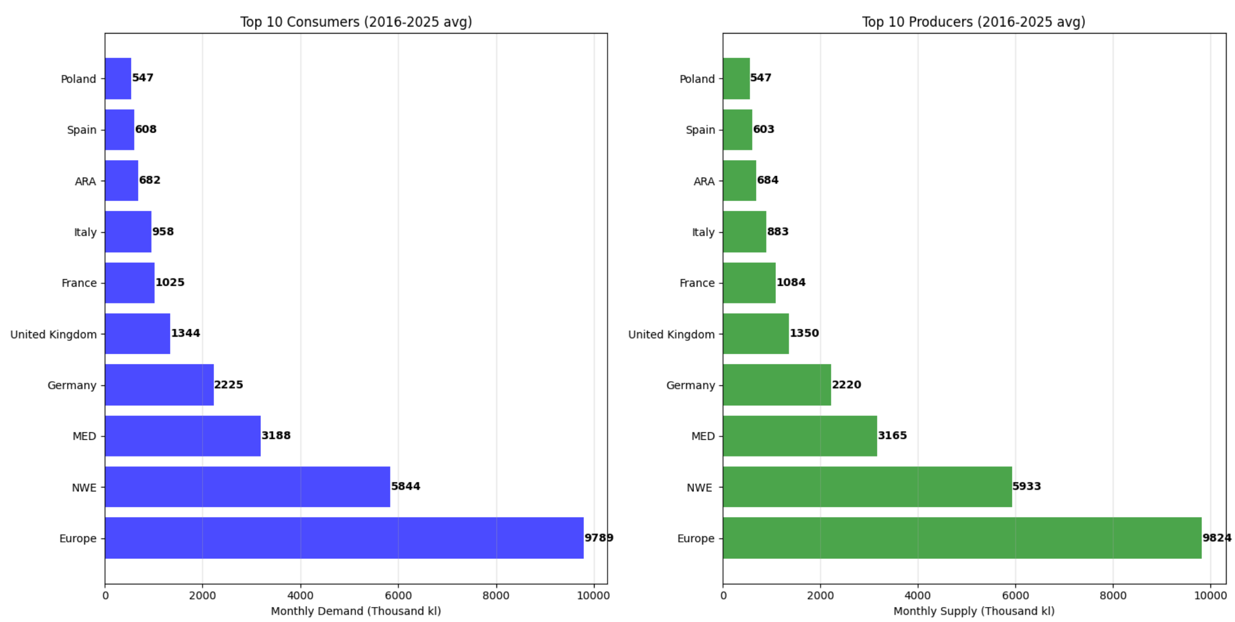

Fig. 3: A graph of the Top Players in the Gasoline Market in Europe

Fig. 3 represents the Top 7 gasoline consumers and top 7 gasoline producers across European markets, averaged over the period February 2016 to July 2025. Germany has the highest demand and supply at 2,225 and 2220 thousand kl per month, respectively. NEW also tops in both cases, demand and supply, with 5,844 and 5,933 thousand kl respectively. The other top countries include the United Kingdom and France.

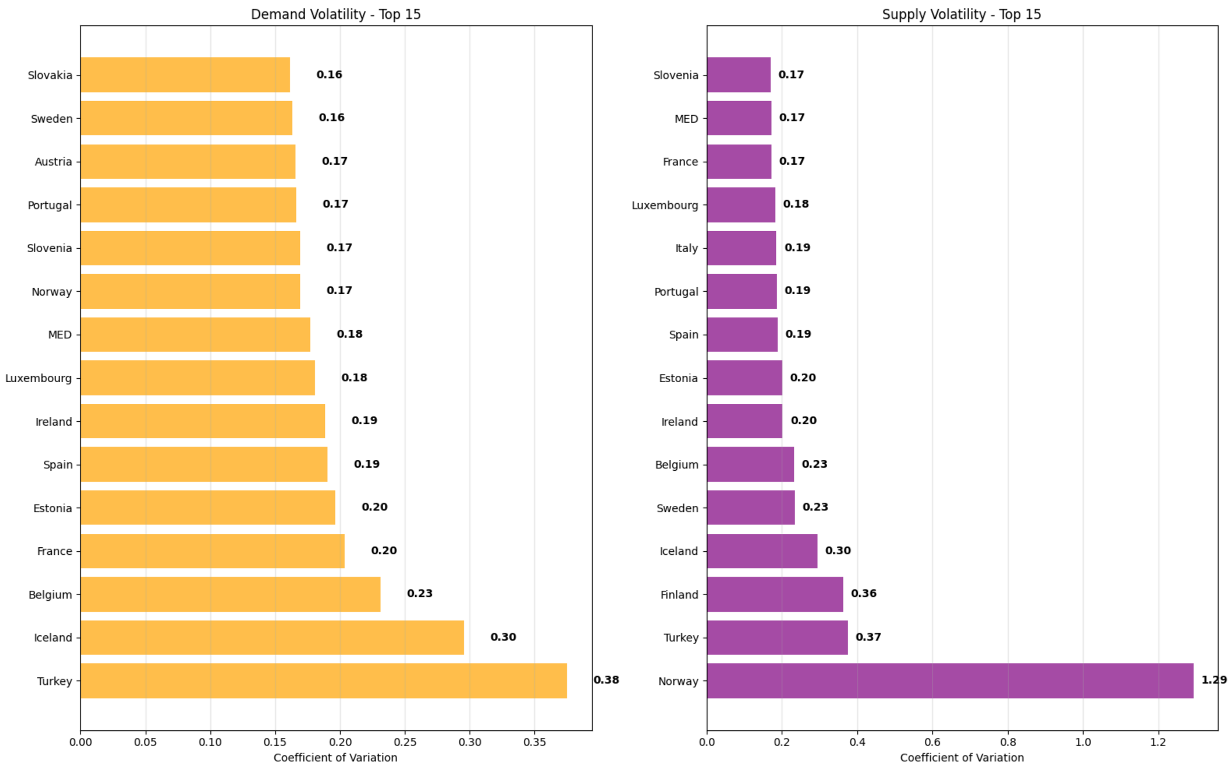

Fig. 4: A Graph of Volatility Analysis

Fig. 4 shows the top 15 countries with the highest demand and supply volatility in gasoline within Europe. The analysis indicates that Turkey and Iceland have the highest volatility, with scores of 0.38 and 0.30, respectively. This is an indication that these markets experience huge changes in demand patterns, a factor that risks the price stability of gasoline in these countries. On the other side, Norway's supply market is more volatile, with a value of 1.29. This is a suggestion that there are supply disruptions that exceed those in other countries, including Turkey and Finland, with 0.37 and 0.36.

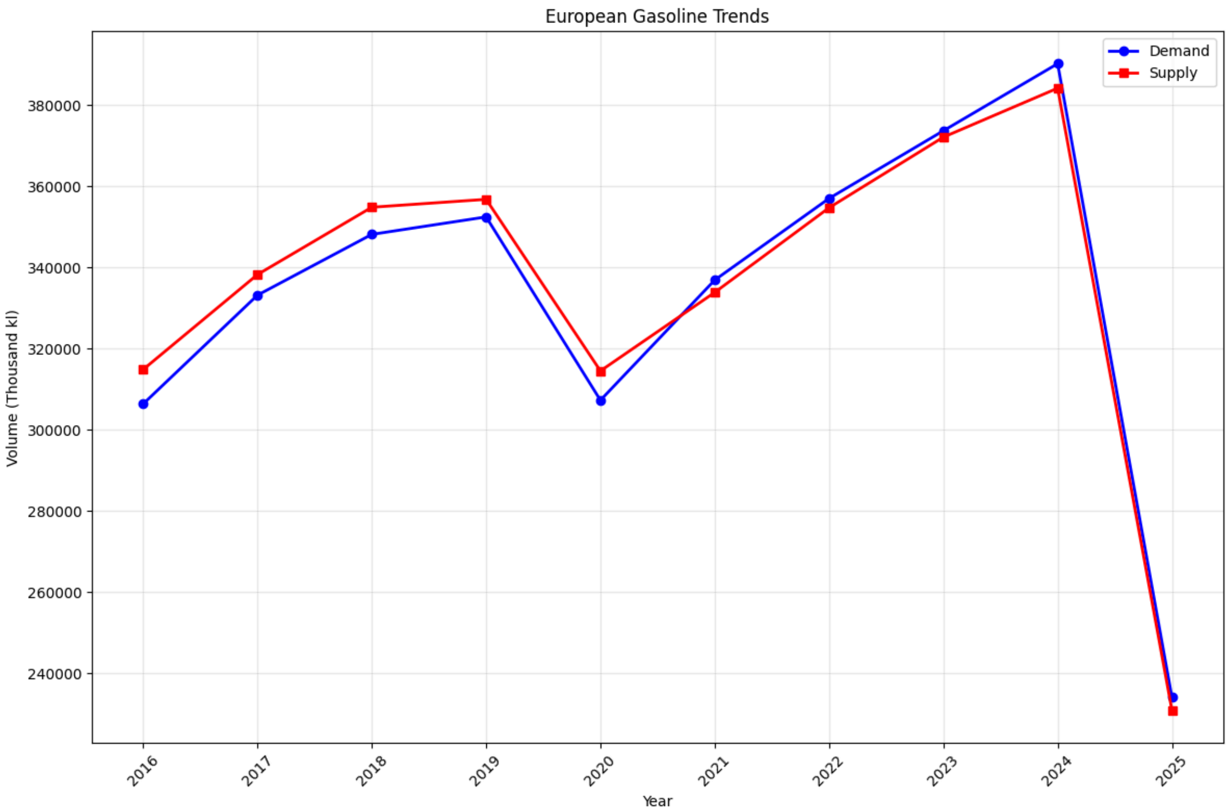

Fig. 5: A Graph of Yearly Trends from 2016 to 2025

Fig. 5 shows demand and supply trends from February 2016 to July 2025. The analysis indicates that there was a sharp decline in both gasoline demand and supply around 2020, a period that corresponds to the onset of the COVID-19 pandemic. This drop can be attributed to the low mobility of individuals in Europe due to lockdowns and travel restrictions, as well as shutdowns of refineries, which led to low consumption and production of gasoline. However, both demand and supply show a gradual recovery moving forward from 2021, which is a reflection of the return of normalcy in economic and transportation activities**.**

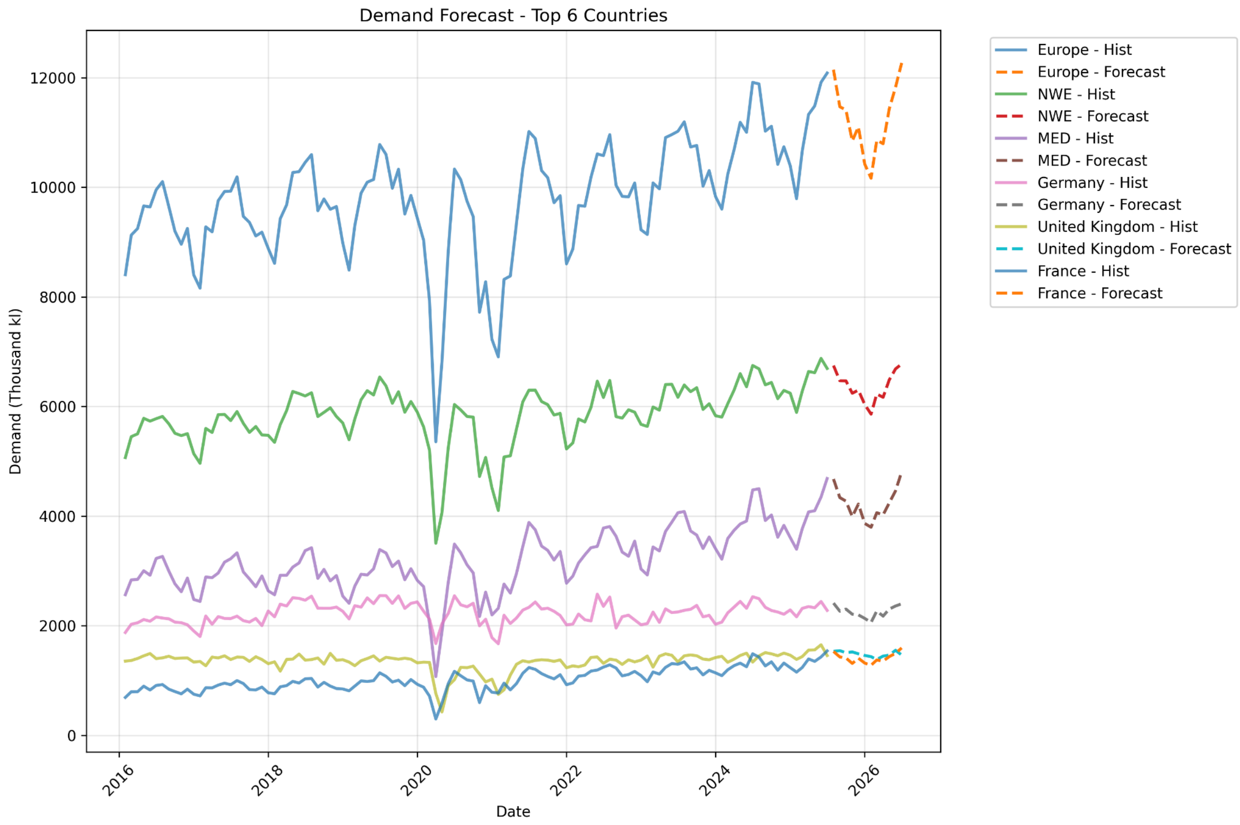

Fig. 6: A graph of Demand Forecast

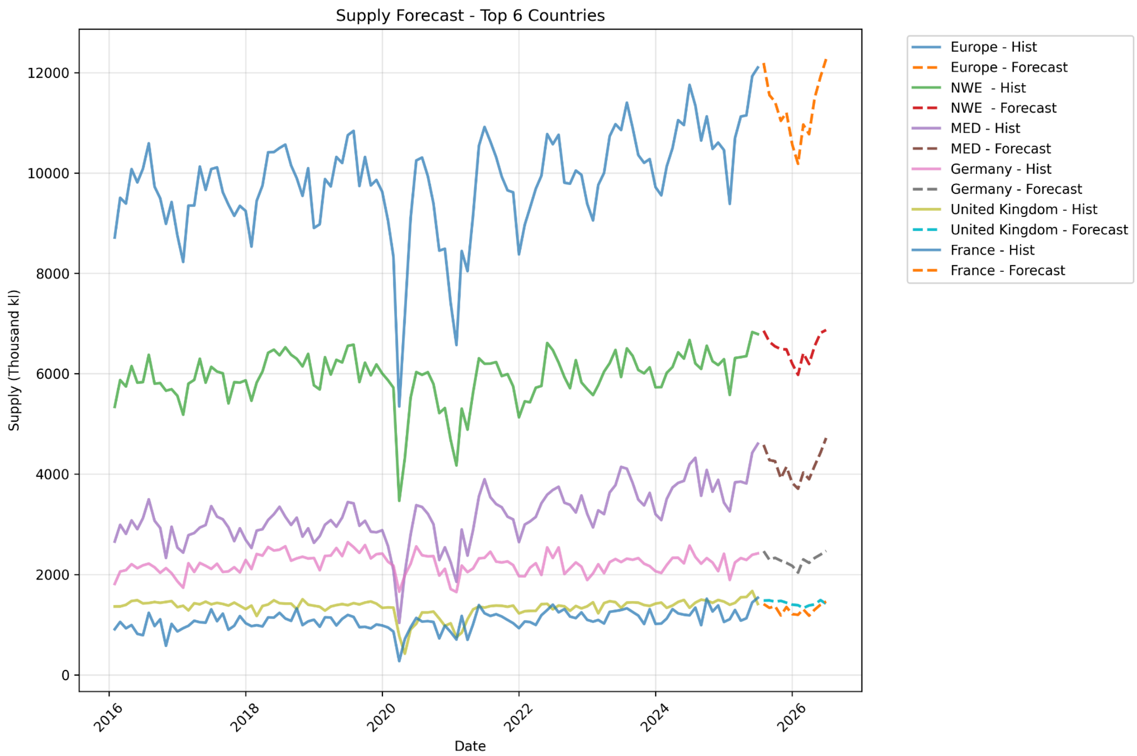

Fig. 7: A graph of Supply Forecast

Fig. 6 shows the future prediction of gasoline in Europe, illustrating that by the end of 2025 and early 2026, demand is expected to decrease slightly. This could be due to the new trend of electric cars in the transportation industry. However, by the mid- and end of 2026, it is projected to increase, likely due to increased travel to attend activities such as the World Cup in the USA. The same trend is anticipated in supply, as shown in Fig. 7. This is because, as demand rises, refineries will likely adjust their output to meet the higher consumption, ensuring that supply keeps pace with the growing needs of the market.

Discussion

The results show that the project was successful in producing the expected outcomes. Fig. 1 shows the regions and countries with the largest gasoline surpluses and deficits, which help to identify key areas for market opportunities. Similarly, Fig. 2 demonstrates strong positive correlations between demand and supply in countries such as Austria and Turkey, which indicates that their markets are more predictable. Fig. 4 has revealed that some volatile markets, such as Norway for supply and Turkey for demand, can help decision-makers in the oil industry pinpoint regions with higher market risk in terms of price changes.

Although this analysis has been comprehensive, there are some limitations since it is heavily dependent on the JODI data. The dataset may not always be fully accurate, as it is from a single source, with no other source to compare it with to determine if the data is consistent. Therefore, if there are errors in collecting and reporting the figures across countries, they may be propagated when analyses are carried out. Also, analysing 25 European countries does not reflect the state of the entire market, as the majority of countries are left out due to missing data.

For future work in projects like this, it is important to integrate more diverse data sources, including macroeconomic indicators (GDP, inflation, population growth) and real-time refinery capacity data. This approach would improve the dependability of the results. It would also be beneficial to explore the impact of emerging alternative fuels, such as biofuels and hydrogen, on gasoline supply and demand, as these technologies are gaining popularity in Europe. A researcher could also expand the analysis to explore the impact of environmental regulations, such as the European Green Deal, on the gasoline market by investigating whether their implementation has an immediate effect on demand and supply.

Conclusion

An analysis of the demand and supply of petroleum products, like gasoline, is important for businesses in the oil industry as it provides insights that help them make decisions on whether to alter production and pricing. The findings from this study reveal that there are imbalances between these two factors across European regions, such as Italy and the MED region. However, NEW exhibits a consistent surplus. The study further shows strong correlations in countries like Austria and Turkey, indicating that these markets are stable. The volatility analysis identifies regions with high instability, such as Norway, which may present higher risks for traders as prices may not be predictable. The forecasting models suggest a slight decrease in gasoline demand towards the end of 2025, followed by a recovery.

Works Cited

Fuels Europe. *Worldwide Refined Product Demand Averaged 101.7 Million Barrels per Day in 2023*. 2024, www.fuelseurope.eu/uploads/files/modules/documents/file/1720681837_hqsK59auHpdzm5XdKtuDhE0nHIG4mgDxYfxtlLHk.pdf.

Gkatzoglou, Fotios, et al. "Fuel Price Networks in the EU." *Economies*, vol. 12, no. 5, Apr. 2024, p. 102. https://doi.org/10.3390/economies12050102.

JODI. *Beyond 20/20 WDS - Table view*. 2022, www.jodidb.org/TableViewer/tableView.aspx?ReportId=93905.

Notteboom, Theo, et al. "Evolution of the Port of Rotterdam." *Port Economics, Management and Policy | a Comprehensive Analysis of the Port Industry*, 20 Jan. 2024, porteconomicsmanagement.org/pemp/contents/part3/changing-geography-of-seaports/evolution-port-rotterdam

Park, Wesley. "The Diesel Divide: Why Middle Distillates Offer a Contrarian Bet in a Flat Oil Market." *Ainvest*, 27 July 2025, www.ainvest.com/news/diesel-divide-middle-distillates-offer-contrarian-bet-flat-oil-market-2507.

Port of Rotterdam. *Pipeline Network*. www.portofrotterdam.com/en/logistics/connections/intermodal-transportation/pipeline-network.

Strautmann, Mick. *Middle Distillates Set to Find Further Support Into Q4*. 3 Sept. 2025, www.vortexa.com/insights/middle-distillates-find-support-into-q4.

Suchodolska, Liliana. *Hands-on Exercise Using Sample Data*. 2012, www.jodidata.org/_resources/files/training-workshops/7th-regional-jodi-training-workshop/07-handson-exercise-using-sample-data.pdf.

Van Den Berghe, Karel, et al. "Friends With Benefits: The Emergence of the Amsterdam–Rotterdam–Antwerp (ARA) Polycentric Port Region." *Territory Politics Governance*, vol. 11, no. 2, Jan. 2022, pp. 301–20. https://doi.org/10.1080/21622671.2021.2014353.

Appendices

Appendix 1: Combined Analysis Code

```python
"""
Quick country & regional balance analysis
For trader market positioning
"""

import pandas as pd
import matplotlib.pyplot as plt
import os
from data_loader import load_gasoline_data

# Setup output dir
os.makedirs('../results/figures/combined_analysis', exist_ok=True)

print("=== COUNTRIES & REGIONS ANALYSIS ===")

# Load market data
demand, supply = load_gasoline_data()

if demand is not None and supply is not None:
    # Calculate supply-demand gaps
    balance = supply - demand
    country_avg = balance.mean(axis=1)

    # Get top 15 imbalanced markets
    sorted_countries = country_avg.sort_values()
    top_countries = sorted_countries.head(15)

    # Define trading regions
    regions = {
        'ARA Hub': ['Netherlands', 'Belgium', 'Germany'],
        'North West': ['United Kingdom', 'France'],
        'Mediterranean': ['Spain', 'Italy', 'Greece'],
        'East Europe': ['Poland', 'Czech Republic', 'Hungary']
    }

    # Calculate regional totals
    regional_data = []
    for region_name, countries in regions.items():
        region_total = 0
        for country in countries:
            if country in country_avg.index:
                region_total += country_avg[country]
        regional_data.append([region_name, region_total])

    regional_df = pd.DataFrame(regional_data, columns=['Region', 'Balance'])
```

```
    regional_sorted = regional_df.set_index('Region')['Balance'].sort_values()

    # Create comparison chart
    fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(16, 10))

    # Country balance chart
    colors1 = ['red' if x < 0 else 'green' for x in top_countries.values]
    bars1 = ax1.barh(top_countries.index, top_countries.values, color=colors1)

    # Add value labels
    for bar, value in zip(bars1, top_countries.values):
        ax1.text(bar.get_width() + (10 if value >= 0 else -20),
                 bar.get_y() + bar.get_height()/2,
                 f'{value:+.0f}',
                 ha='left' if value >= 0 else 'right',
                 va='center',
                 fontweight='bold')

    ax1.axvline(x=0, color='black', linewidth=2)
    ax1.set_title('Country Balance')
    ax1.set_xlabel('Balance (Thousand kl)')
    ax1.grid(axis='x', alpha=0.3)

    # Regional balance chart
    colors2 = ['red' if x < 0 else 'green' for x in regional_sorted.values]
    bars2 = ax2.barh(regional_sorted.index, regional_sorted.values,
color=colors2)

    for bar, value in zip(bars2, regional_sorted.values):
        ax2.text(bar.get_width() + (10 if value >= 0 else -20),
                 bar.get_y() + bar.get_height()/2,
                 f'{value:+.0f}',
                 ha='left' if value >= 0 else 'right',
                 va='center',
                 fontweight='bold')

    ax2.axvline(x=0, color='black', linewidth=2)
    ax2.set_title('Regional Balance')
    ax2.set_xlabel('Balance (Thousand kl)')
    ax2.grid(axis='x', alpha=0.3)

    plt.tight_layout()
    plt.savefig('../results/figures/combined_analysis/countries_regions.png')
    plt.show()

    print("Chart saved: countries_regions.png")

else:
```

```
    print("No data loaded")
```

Appendix 1: Correlation Analysis Code

```
import pandas as pd
import matplotlib.pyplot as plt
import os
from data_loader import load_gasoline_data

# Quick setup for output folders
os.makedirs('../results/figures/correlation', exist_ok=True)
os.makedirs('../results/tables', exist_ok=True)

def get_market_correlations(demand_data, supply_data):
    """Check how demand and supply move together for each market"""
    results = []

    for market in demand_data.index:
        if market in supply_data.index:
            # Calculate correlation for this market
            corr = demand_data.loc[market].corr(supply_data.loc[market])
            results.append({
                'market': market,
                'correlation': corr
            })

    return pd.DataFrame(results)

def plot_market_correlations(corr_data, top_markets=8):
    """Show which markets have strongest supply-demand relationships"""
    # Sort by correlation strength
    sorted_data = corr_data.sort_values('correlation', ascending=False)
    top_pos = sorted_data.head(top_markets)
    top_neg = sorted_data.tail(top_markets)

    fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(14, 6))

    # Markets where supply tracks demand well
    bars1 = ax1.barh(top_pos['market'], top_pos['correlation'],
                     color='green', alpha=0.7)
    for bar, val in zip(bars1, top_pos['correlation']):
        ax1.text(bar.get_width() + 0.02, bar.get_y() + bar.get_height()/2,
                f'{val:.3f}', va='center') ax1.set_title(f'Top
    {top_markets} - Supply Tracks Demand')
    ax1.set_xlabel('Correlation')
    ax1.set_xlim(0, 1)

    # Markets with inverse relationships
```

```
    bars2 = ax2.barh(top_neg['market'], top_neg['correlation'],
                    color='red', alpha=0.7)
    for bar, val in zip(bars2, top_neg['correlation']):
        ax2.text(bar.get_width() - 0.03, bar.get_y() + bar.get_height()/2,
                f'{val:.3f}', va='center') ax2.set_title(f'Top
    {top_markets} - Inverse Relationship')
    ax2.set_xlabel('Correlation')
    ax2.set_xlim(-1, 0)

    plt.tight_layout()
    return fig

# Run the analysis
print("Checking market correlations...")

# Load the data
demand, supply = load_gasoline_data()

if demand is None or supply is None:
    print("No data - check files")
else:
    # Calculate all market correlations
    correlations = get_market_correlations(demand, supply)
    correlations = correlations.sort_values('correlation', ascending=False)

    # Create the chart
    fig = plot_market_correlations(correlations)
    plt.savefig('../results/figures/correlation/demand_supply_correlation.png'
,
                dpi=300, bbox_inches='tight')
    plt.show()

    # Save the results
    correlations.to_csv('../results/tables/correlation_results.csv',
index=False)

    # Print key insights
    avg_corr = correlations['correlation'].mean()
    print(f"Average market correlation: {avg_corr:.3f}")

    print("\nMarkets with strongest tracking:")
    for _, row in
    correlations.head(3).iterrows():
        print(f" {row['market']}: {row['correlation']:.3f}")

    print("\nMarkets with weakest tracking:")
    for _, row in correlations.tail(3).iterrows():
        print(f" {row['market']}: {row['correlation']:.3f}")
```

```python
# Quick market efficiency note
if avg_corr > 0.7:
    print("\n✅ Markets generally efficient - supply follows demand")
elif avg_corr > 0.4:
    print("\n⚠️      Mixed efficiency - some markets
disconnected") else:
    print("\n❌ Low efficiency - supply/demand often move independently")
```