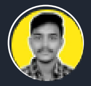
Oil Well Analysis: Findings, Insights, and Recommendations

In this presentation, we'll go through the process of oil well analysis, highlighting key insights and recommendations.



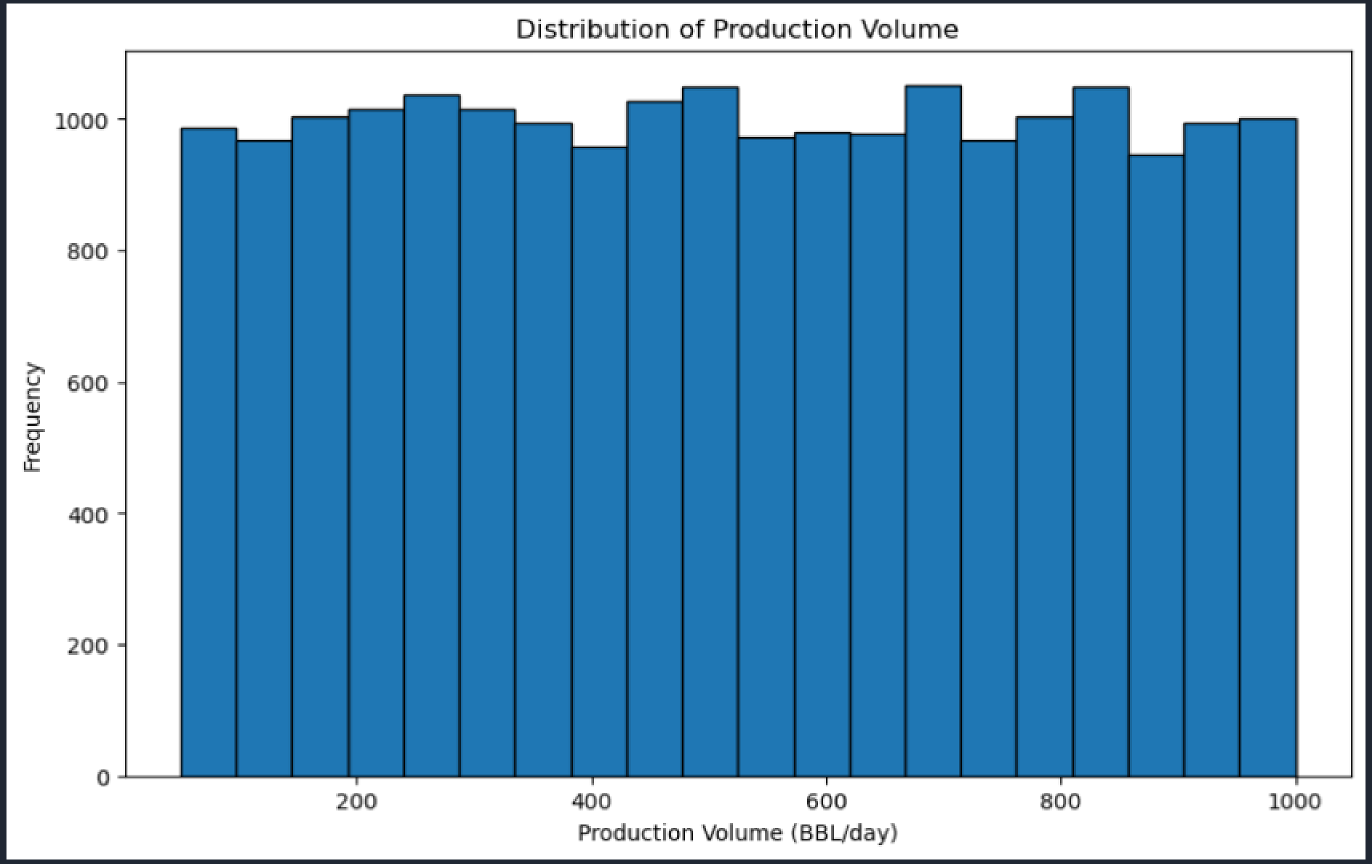
by Parasmani Sahu

Data Generation Process

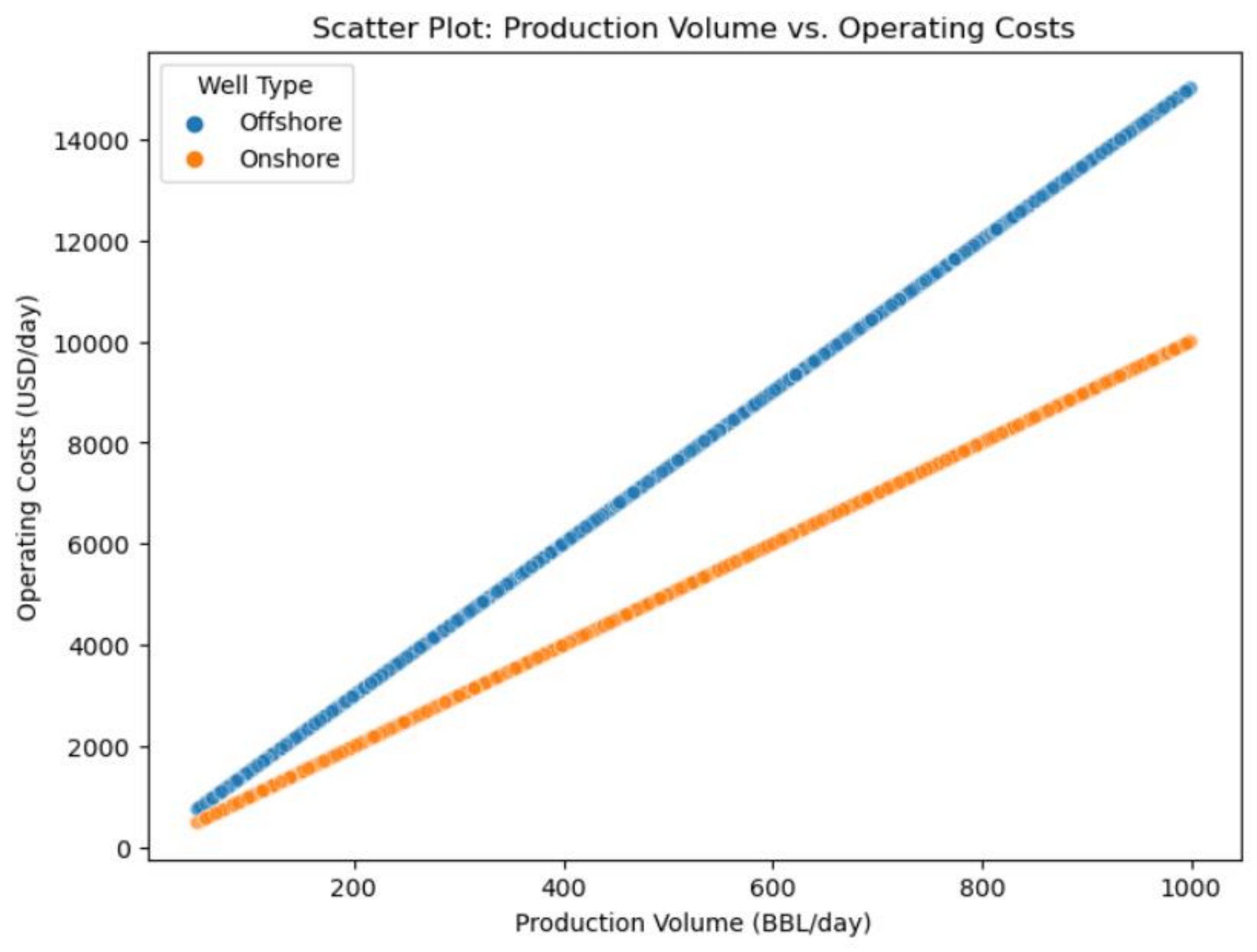
- **Objective:** The primary objective of the data generation process was to create a realistic dataset simulating oil well data for comprehensive analysis and forecasting.
- **Dataset Size:** The dataset was designed to represent 1,000 unique oil wells, each containing 20 randomly generated line items of data, resulting in a total of at least 20,000 data points.
- **Time Range:** To capture diverse production scenarios, the production dates were generated within a substantial time range. The dataset encompasses data from January 1, 2015, to December 31, 2022.
- **Oil Production Volume (BBL/day):** The production volume of each well was randomly generated, with values ranging from 50 to 1000 barrels of oil per day.
- **Operating Costs (USD/day):** To ensure practicality, the operating costs were calculated based on several factors, including the production volume, well type (onshore or offshore), and well pressure.
- **Well Locations:** The well locations were thoughtfully assigned to ensure representation from diverse regions. The dataset includes well locations such as Abqaiq, Ghawar, Haradh, Khurais, Khursaniyah, Manifa, Nuayyim, and Qatif, offering a comprehensive portrayal of different geographical areas.
- **Oil Composition (API Gravity):** The API gravity values of the extracted oil were randomly generated to represent various oil compositions. These values were categorized into light, medium, and heavy crude oil, with ranges of 36 to 45, 23 to 30, and 10 to 18, respectively.
- **Well Depth (Feet):** To reflect real-world conditions, the depth of each oil well was randomly generated, spanning from 500 to 5000 feet. This range accounts for the variations in well depths encountered in actual oil drilling operations.
- **Well Type:** Each well was designated as either onshore or offshore to distinguish between different types of oil wells, each with distinct operational characteristics and costs.
- **Well Pressure (PSI):** The well pressure, an essential parameter in oil extraction, was randomly generated between 2000 and 5000 PSI.

Exploratory Data Analysis (EDA)

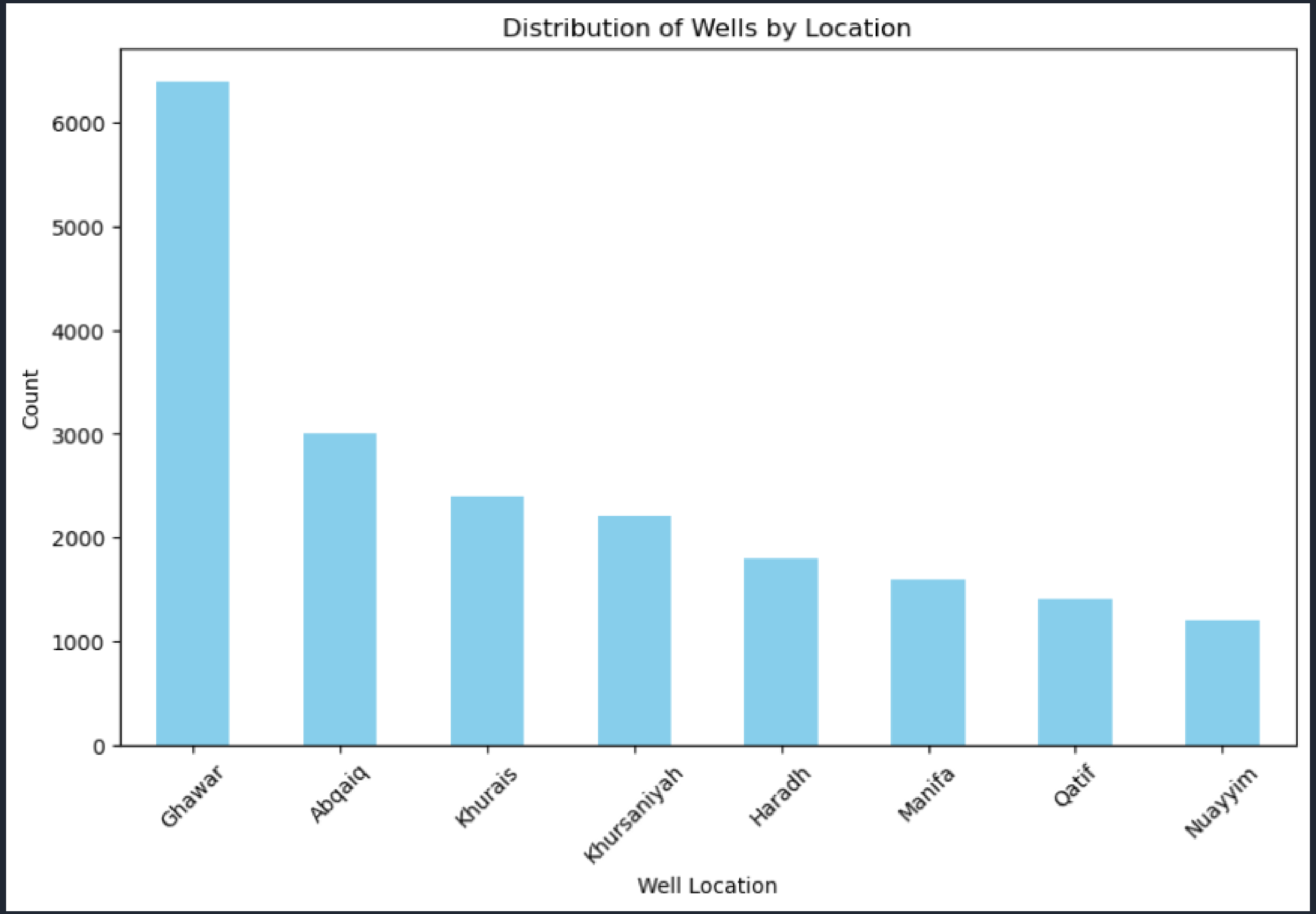
- **Distribution of Production Volume:** The histogram of 'Production Volume' shows that most oil wells have a production volume between 50 to 500 barrels per day.



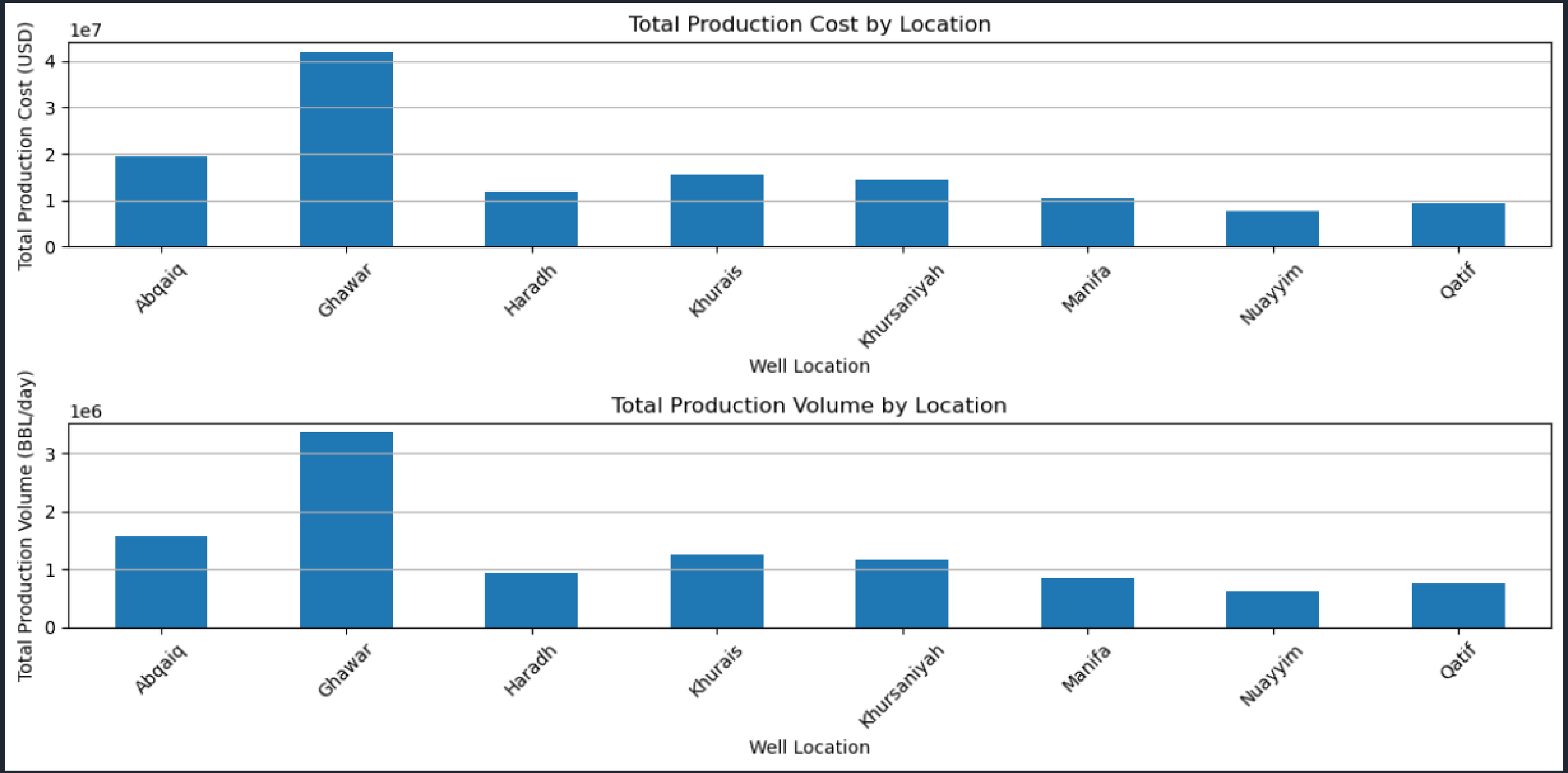
- **Offshore vs. Onshore**: The high seas demand a higher price! Operating costs are greater for offshore wells, making onshore wells the more cost-effective option.
- **Profitable Partnership**: Witness a fruitful partnership! Production Volume and Operating Costs exhibit a positive correlation, especially for onshore wells.

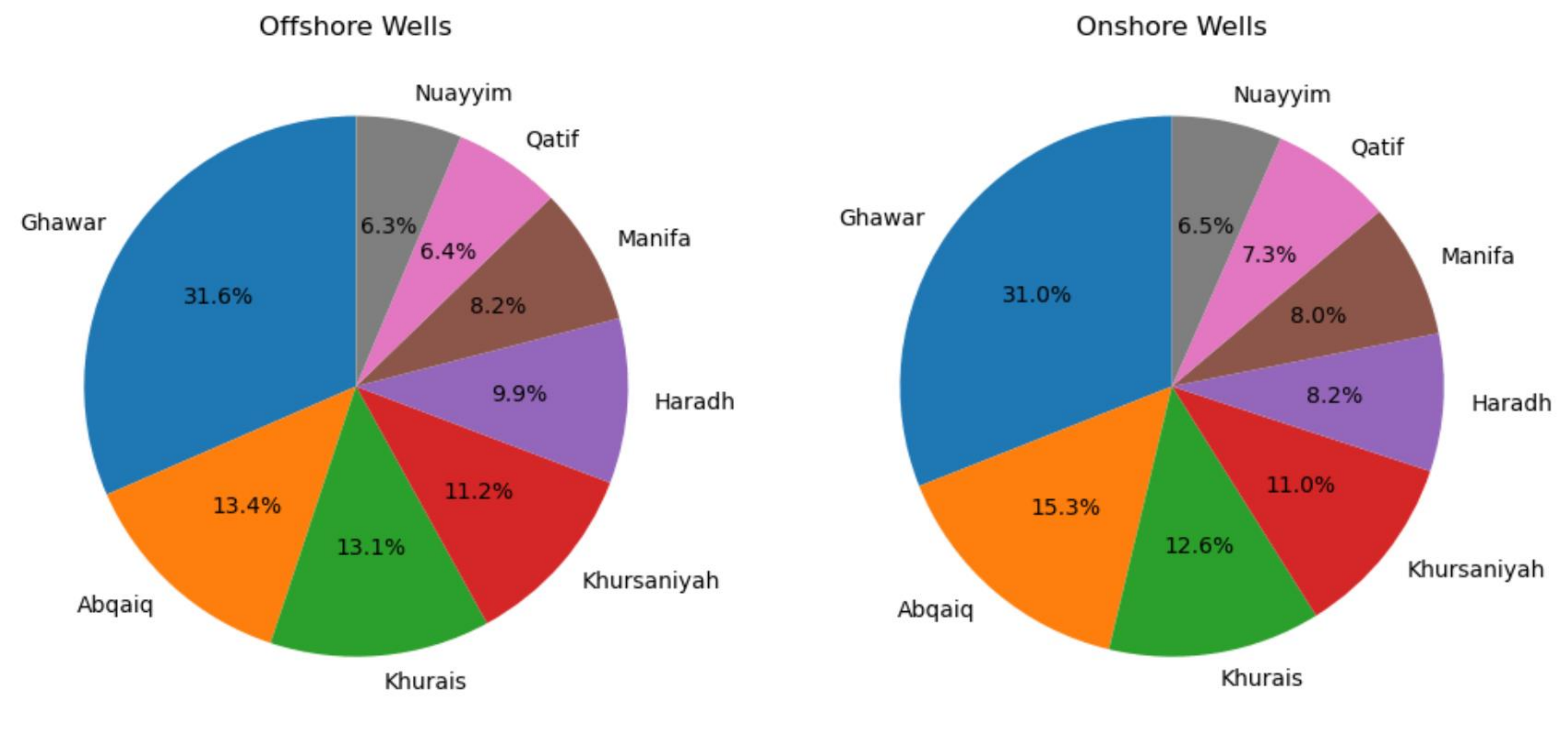


- **Majestic Oil Havens**: Behold Ghawar, Abqaiq, and Manifa! These oil havens reign supreme in total production cost and volume.

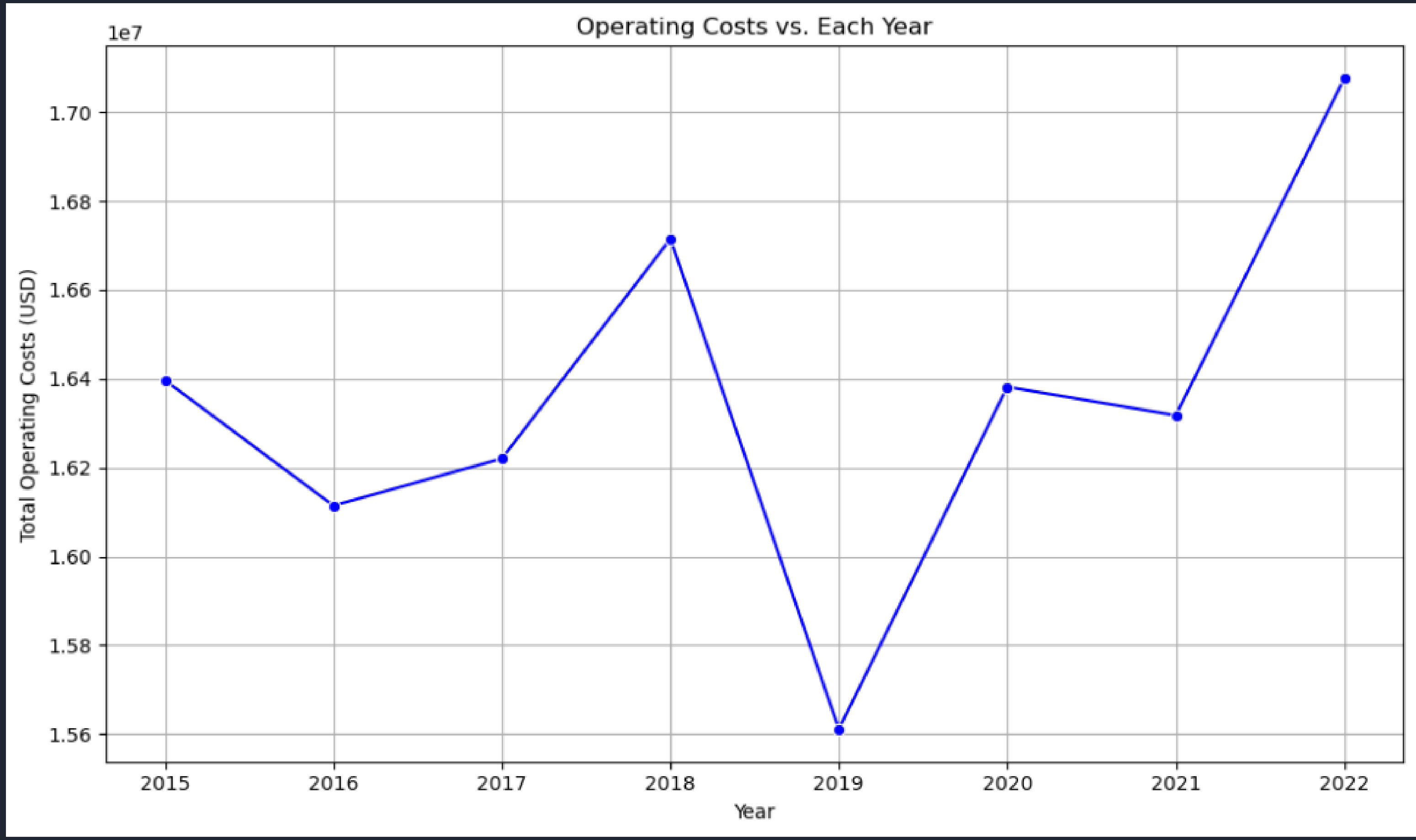


- **Total Production Cost by Location:** The bar chart highlights the total production costs for each location. Ghawar emerges as the region with the highest cumulative production cost, indicating its strategic importance in oil production.
- **Total Production Volume by Location:** The bar chart showcases the total production volume for each location. Once again, Ghawar proves its dominance with the highest cumulative production volume.



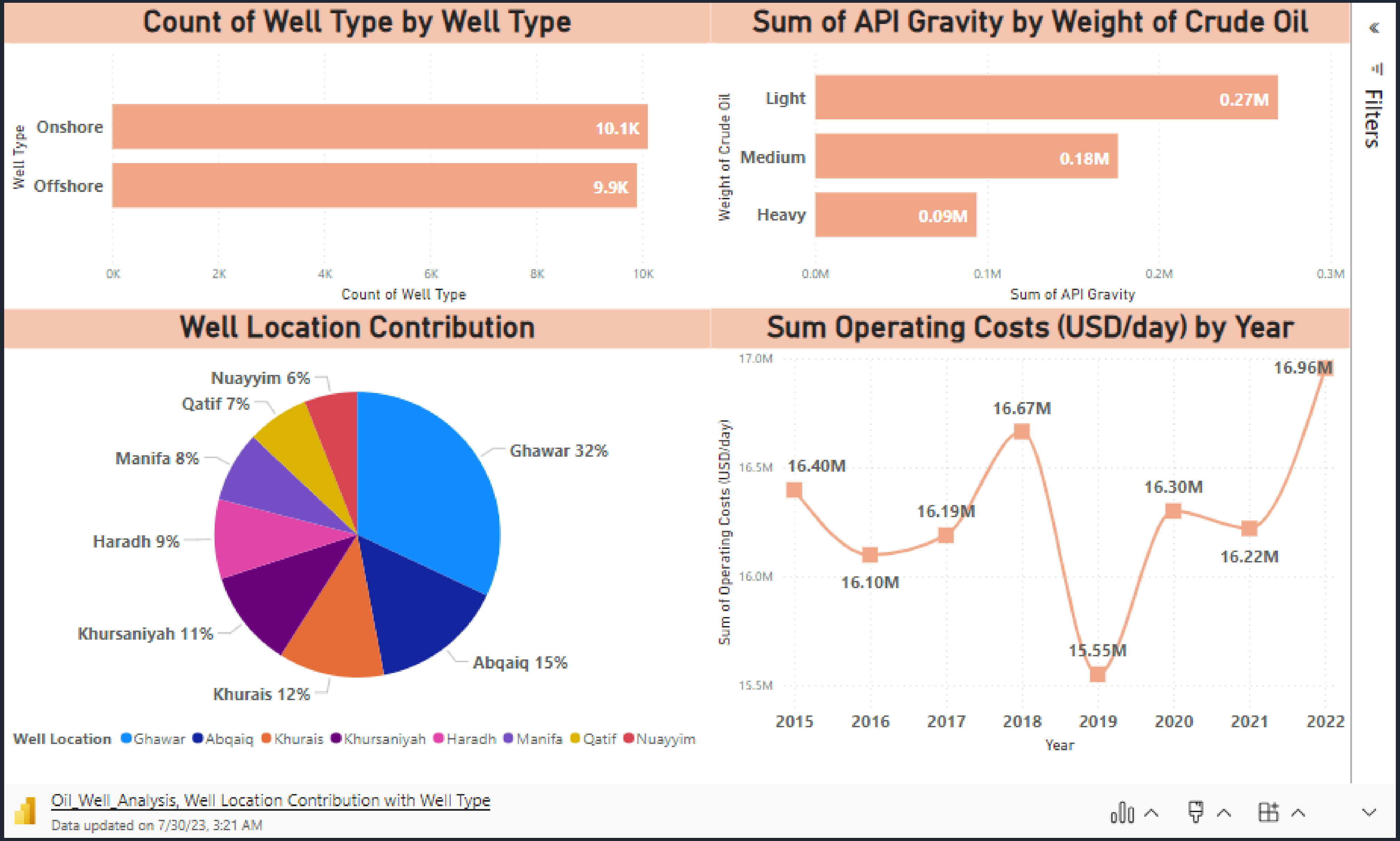

Offshore Wells
Nuayyim
Qatif
Ghawar
Manifa
Haradh
Khursaniyah
Abqaiq
Khurais
6.3%
6.4%
31.6%
8.2%
9.9%
11.2%
13.4%
13.1%
Onshore Wells
Nuayyim
Qatif
Ghawar
Manifa
Haradh
Khursaniyah
Abqaiq
Khurais
6.5%
7.3%
31.0%
8.0%
8.2%
11.0%
15.3%
12.6%

- **Operating Costs vs. Year (Line Plot):** The line plot visualizes the trend of total operating costs over the years. A gradual increase in operating costs can be observed, suggesting the impact of various factors influencing oil well operations over time.



PowerBI Report

PowerBI is used to visualize and explore for a deeper understanding of the data. EDA was performed in PowerBI to include interactive graphs and visualizations for an even deeper analysis.



Click

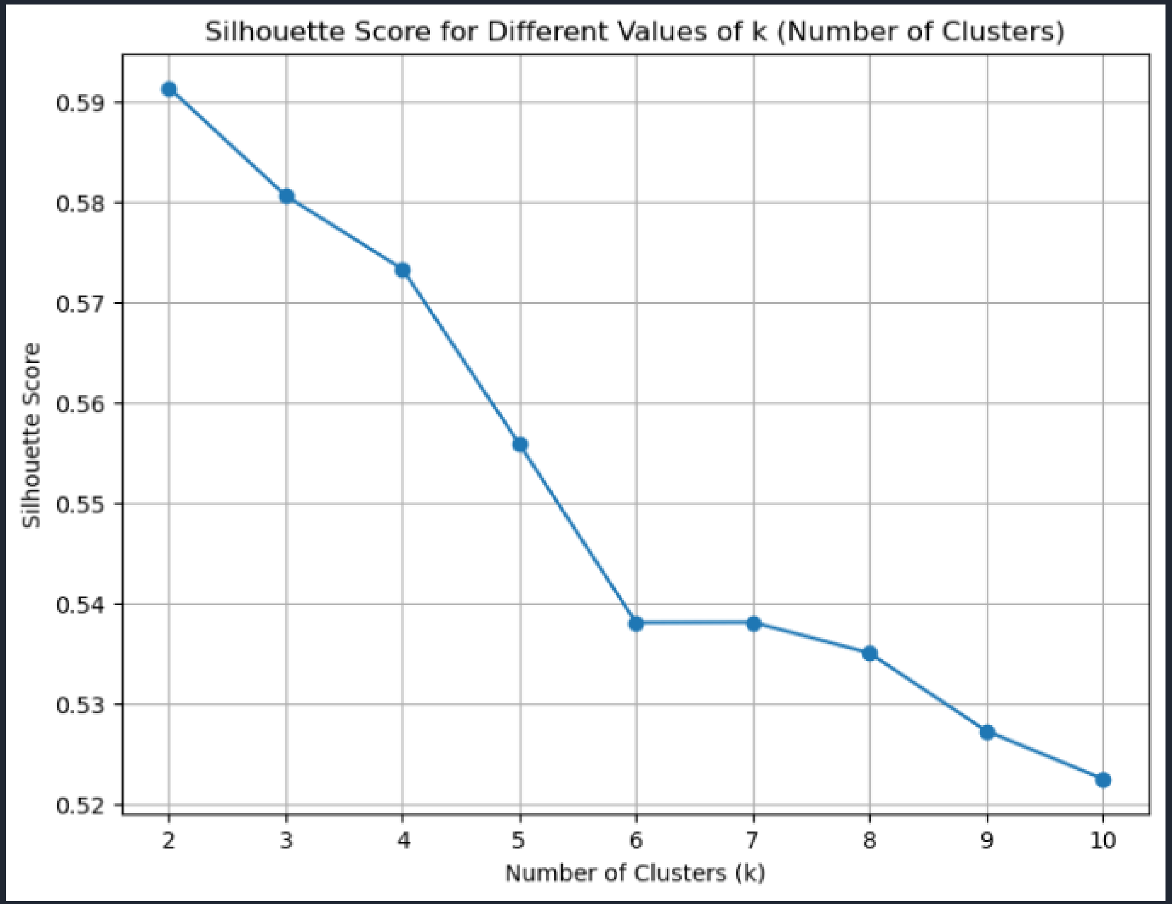
In Report:

1. Well Location Contribution with Well Type
2. Operating Cost and Volume
3. Well Depth and Pressure
4. 10 Most Operating Cost Well
5. 10 Least Operating Cost Well

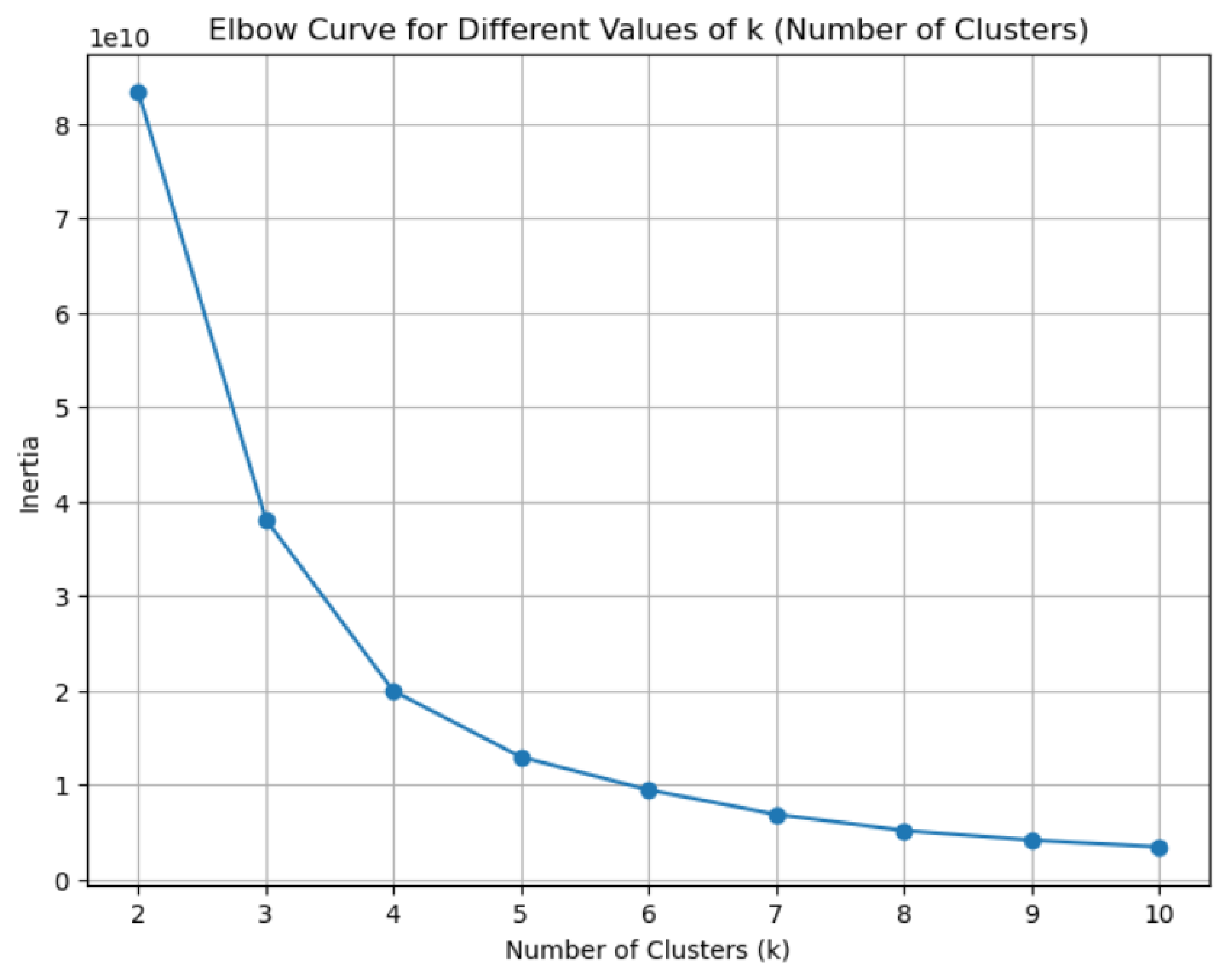
Oil Well Segmentation using K-Means Clustering

We are using the K-Means clustering technique for well segmentation based on production volume and operating costs. K-Means is chosen for its simplicity, efficiency, and ability to partition data into distinct clusters, allowing us to identify meaningful groups of oil wells with similar production characteristics.

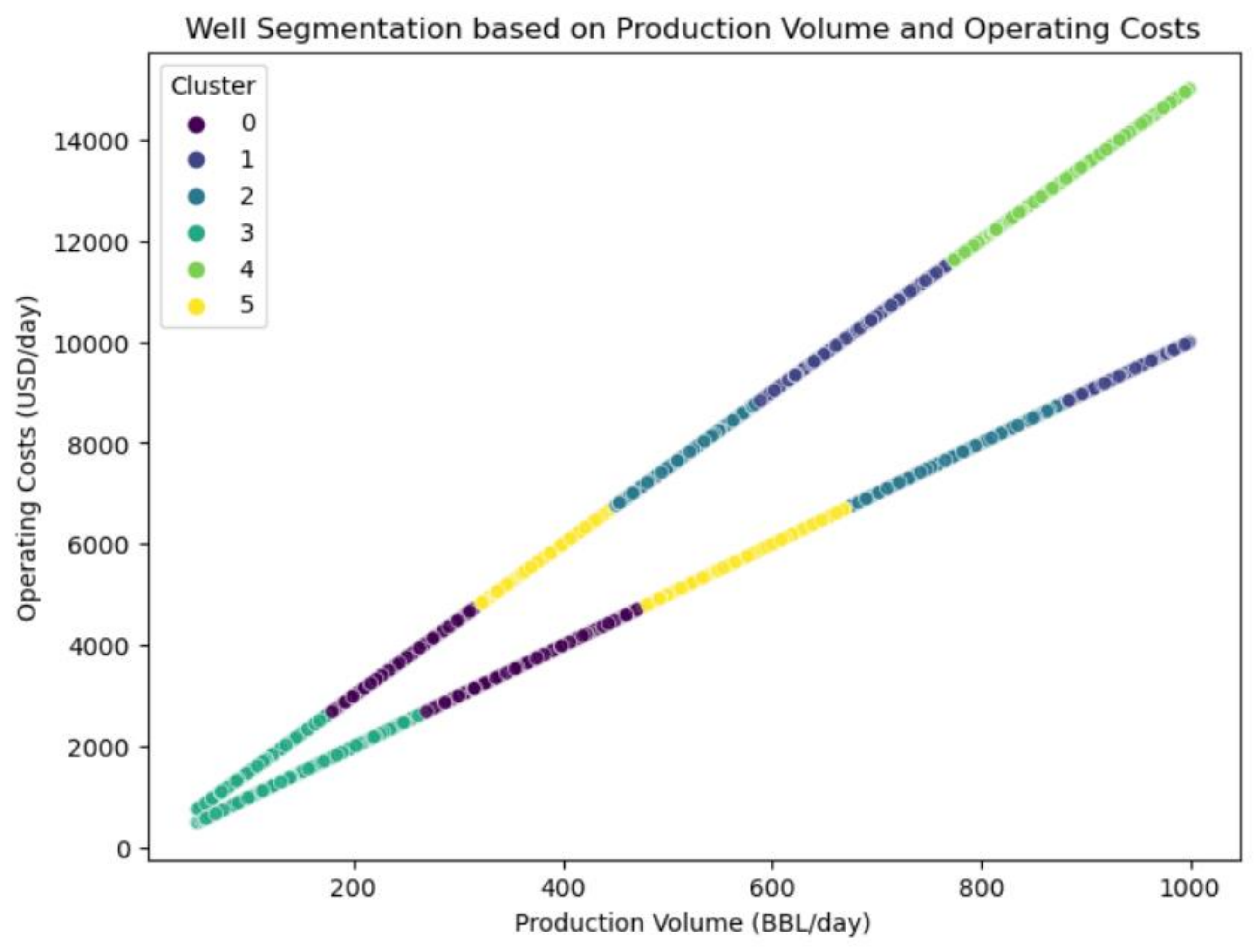
- **Silhouette Score for Optimal Clusters**:
    - Utilized the silhouette score, resulting in the optimal number of clusters (k=6) for well segmentation.
    - The silhouette score technique helped identify the number of clusters that provide well-defined and distinct groups.



- **Elbow Curve for Optimal Clusters**:
  - Employed the elbow method, confirming k=6 as the ideal number of clusters based on inertia values.
  - The elbow curve visually demonstrated the point where additional clusters no longer significantly reduce inertia.



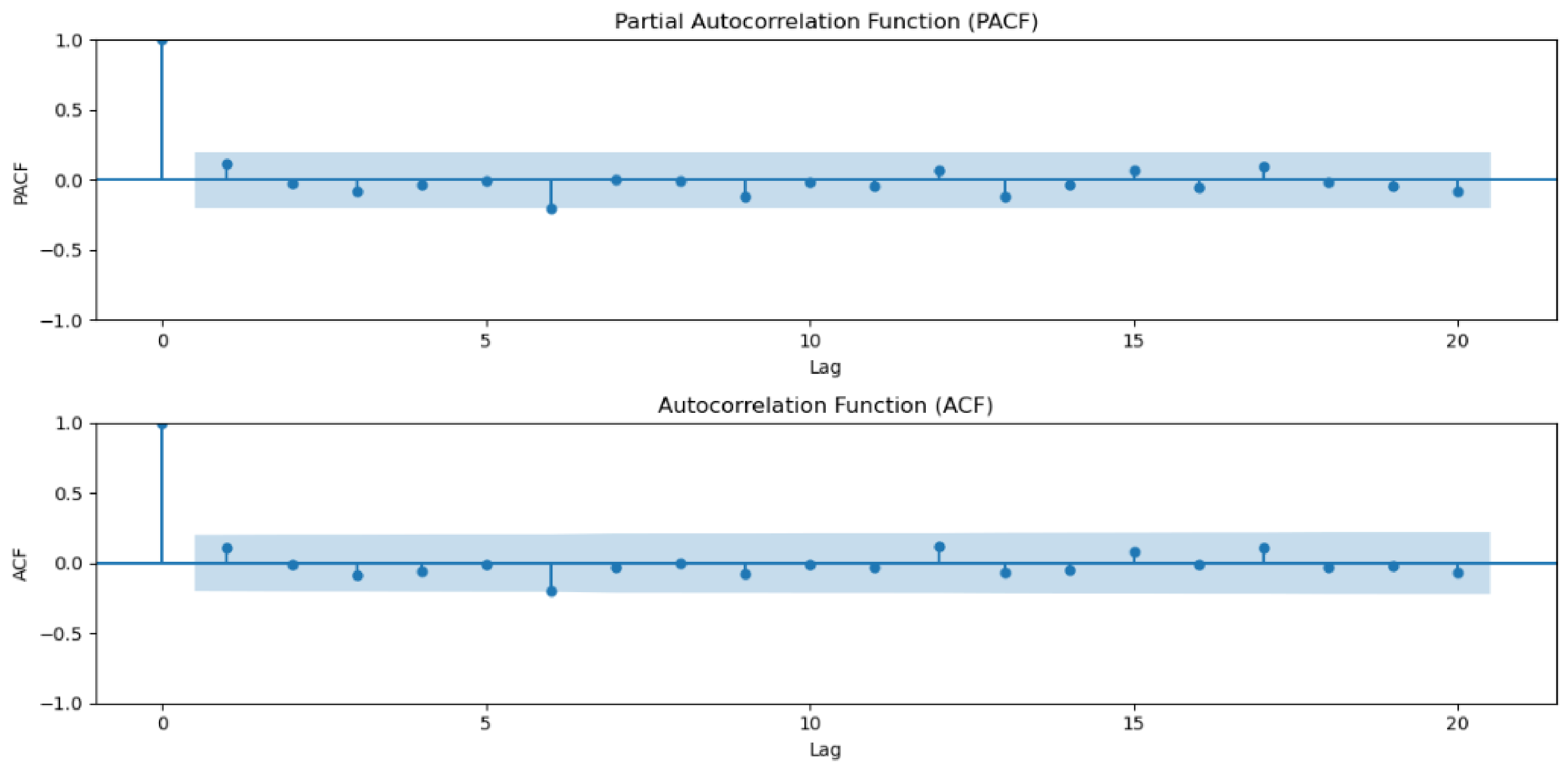
- **Well Segmentation using K-Means Clustering**:
  - Performed K-Means clustering with k=6, based on the silhouette score and elbow curve analysis.
  - Created a new 'Cluster' column in the DataFrame to store the cluster assignments for each well.



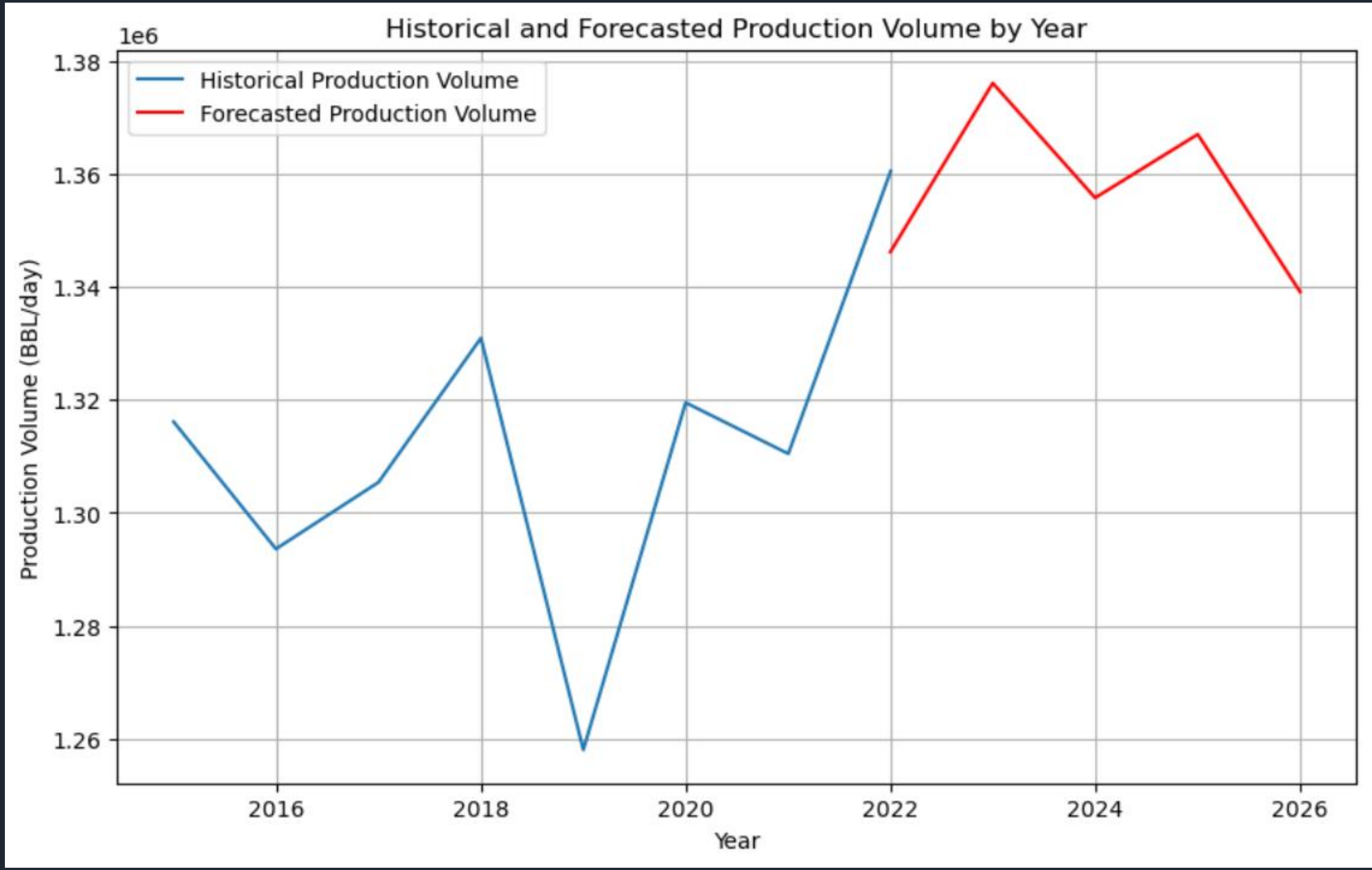
- **High-Producing Wells**:
    - Identified high-producing wells within clusters 1 and 4 for both onshore and offshore well types.
    - Highlighted potential areas of strong oil production performance.
- **Low-Producing Wells**:
    - Discovered low-producing wells within cluster 3 for both onshore and offshore well types.
    - Identified wells that may require further evaluation and optimization.
- **Regional Variation**:
    - Explored the distribution of clusters across different well locations, providing insights into regional variation.
    - Analyzed the composition of clusters in each region to understand production patterns.
- **Operating Costs Analysis**:
    - Analyzed the average operating costs for each cluster, offering insights into cost efficiencies for different well groups.
    - Identified clusters with potential cost-saving opportunities.

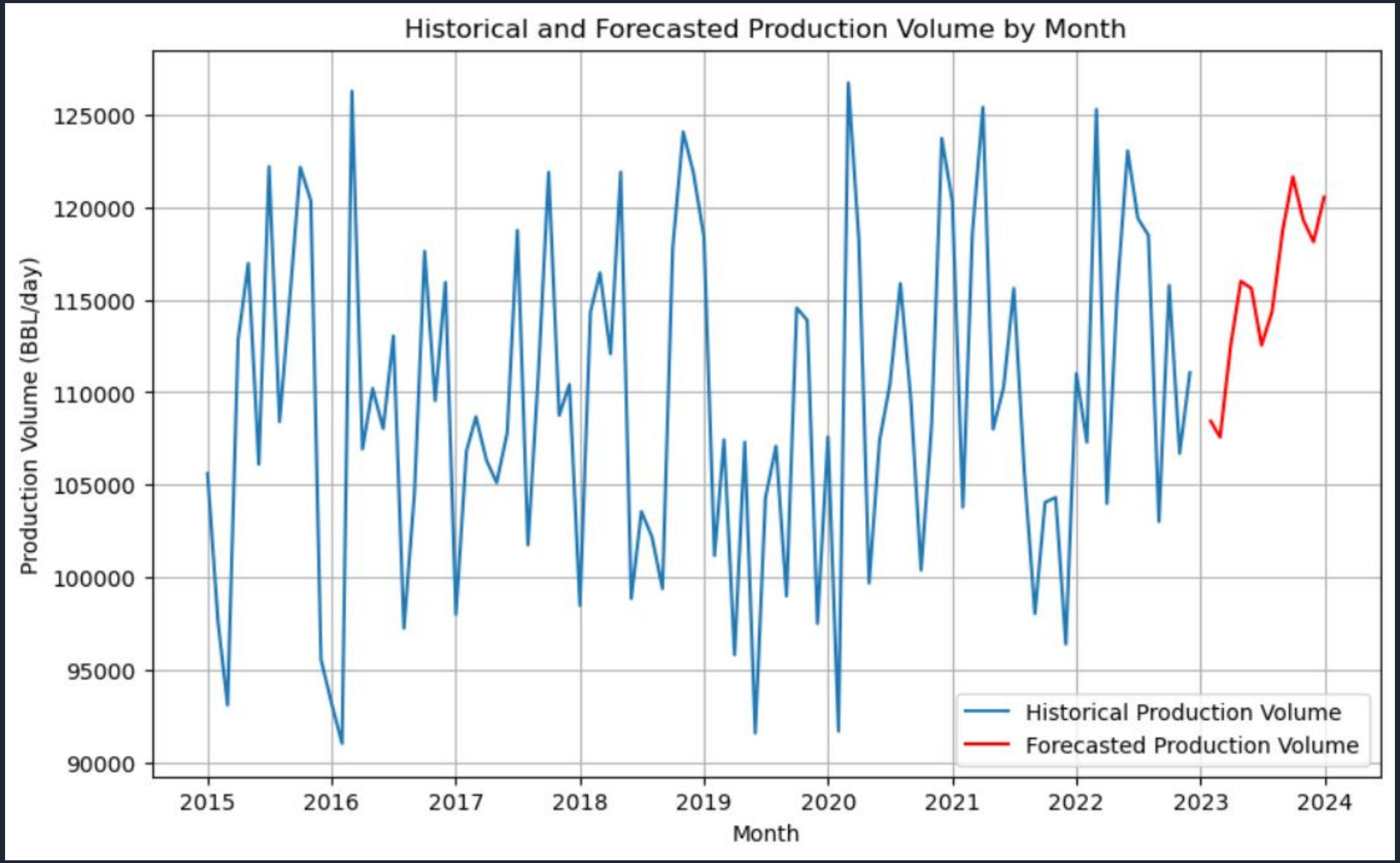
Oil Production Forecasting using ARIMA

- **Time Series Analysis for Forecasting**:
  - Applied ARIMA (AutoRegressive Integrated Moving Average) for time series analysis to forecast oil production volume.
  - Utilized Partial Autocorrelation Function (PACF) and Autocorrelation Function (ACF) plots to identify the optimal order (p, d, q) for the ARIMA model.



- **Yearly Production Forecast**:
    - Forecasted the production volume for the next 5 years based on the yearly aggregated data.
    - Fitted the ARIMA model with order (2, 3, 2) for the time series.
- **Monthly Production Forecast**:
    - Forecasted the production volume for the next 12 months based on the monthly aggregated data.
    - Fitted the ARIMA model with the order (4, 2, 4) for the time series.
- **Visualization of Forecasted Trends**:
    - Plotted historical and forecasted production trends by year and month to visualize the predicted volumes.
    - Provided insights into future production trends, enabling better resource planning and decision-making.


Historical and Forecasted Production Volume by Year
1e6
1.38
1.36
1.34
1.32
1.30
1.28
1.26
Production Volume (BBL/day)
Historical Production Volume
Forecasted Production Volume
2016
2018
2020
2022
2024
2026
Year


Historical and Forecasted Production Volume by Month
Production Volume (BBL/day)
125000
120000
115000
110000
105000
100000
95000
90000
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
Month
Historical Production Volume
Forecasted Production Volume

Insights from Analysis

- The majority of oil wells have a production volume ranging from 50 to 500 barrels per day, with some highly productive wells reaching up to 1000 barrels per day.
- Onshore wells tend to have lower operating costs compared to offshore wells, regardless of their production volume, indicating potential cost-efficiency in onshore drilling.
- Ghawar, Abqaiq, and Manifa are significant regions with a high number of oil wells, substantial production volumes, and significant production costs, making them strategic areas for further exploration and optimization.
- K-Means clustering identified six distinct well clusters with similar production characteristics. High-producing wells are found in clusters 1 and 4, while low-producing wells are located in cluster 3.
- Well clusters exhibit regional variation, implying diverse production patterns in different locations. Tailoring strategies based on regional insights can optimize production efficiency.
- Some clusters show cost-saving potential, emphasizing the importance of data-driven decision-making in well operations. API gravity analysis highlights the impact of oil composition on production characteristics.
- ARIMA modeling enables accurate oil production forecasting. The 5-year yearly forecast indicates steady growth, while the 12-month monthly forecast reveals short-term trends for resource planning.

Recommendations based on Insights

- Focus on optimizing high-producing wells (clusters 1 and 4) to maximize output and revenue generation.
- Implement strategies to improve the productivity of low-producing wells (cluster 3) through careful analysis and optimization efforts.
- High Operating Costs Wells: Implement efficiency measures, predictive maintenance, and staff training to optimize operations and reduce expenses.
- Low Operating Costs Wells: Replicate best practices, optimize resource allocation, and encourage continuous improvement for cost-effectiveness.
- Tailor strategies for each region to maximize production while minimizing operational costs.
- Regularly monitor well performance and operating costs to identify trends and patterns that may impact future production.

THANK YOU