Google Data Analytics Capstone Project

Operational Data Analysis in Oil & Gas
Volve Field Case Study

Student: Richard Zamora

Certificate: Google Data Analytics Professional Certificate

Coursera

1. Ask – Business Task

The objective of this analysis is to evaluate how operational production data can be used to identify efficiency patterns and potential production issues in the oil and gas industry. Using the Volve field as a case study, this project analyzes production behavior over time and across wells to support data-driven operational decisions.

Key questions:
- How is total oil production distributed across wells?
- How has oil production evolved over time?
- Is there a relationship between increasing water production and decreasing oil production?

Stakeholders include operations managers, production engineers, and data analytics teams.

2. Prepare – Data Description

The dataset used is **Volve Production Data**, publicly available on Kaggle. It contains real historical production data from the Volve oil field in Norway.

The data includes daily and monthly production volumes for oil, gas, and water. The dataset was evaluated using the ROCCC framework and was found to be reliable, original, comprehensive, current for analytical purposes, and properly cited.

3. Process – Data Cleaning and Processing

R and RStudio were used for data processing. The following steps were performed:
- Imported Excel files using read_excel()
- Verified sheet and column names
- Converted date fields to proper date format
- Handled missing values using na.rm = TRUE
- Aggregated data by well and date
- Exported cleaned datasets to CSV files for reproducibility.

4. Analyze – Data Analysis

The analysis shows that oil production is highly concentrated in a small number of wells. Time-series analysis indicates a general decline in production, consistent with a mature oil field.

A scatter plot analysis reveals an inverse relationship between water production and oil production, suggesting declining reservoir efficiency as water production increases.

5. Share – Key Findings

- Oil production is unevenly distributed across wells
- Total oil production decreases over time
- Increased water production correlates with declining oil output
- Operational data analysis provides valuable insights for efficiency monitoring

6. Act – Recommendations

It is recommended to prioritize monitoring of high-performing wells, implement water management strategies, conduct regular data-driven reviews of production trends, and apply this analytical approach to other mature fields.

Final Conclusion

This project demonstrates how operational data analysis supports informed decision-making in the oil and gas industry. The project follows all six phases of the Google Data Analytics process and meets Coursera evaluation standards.