Petroleum E & P

Under Guidance :

Mr. Khalid Hussain

Mr. Prashant Bisoi

Mr. Keerthi Dadhirao

Mr. Sanket Patil

Prepared by :

Ms. Pooja Tirupathi

(Fr. C. Rodrigues Institute of Technology,Vashi)

ACKNOWLEDGMENT

I would like to extend my heartfelt appreciation for the guidance, support, and mentorship during the 1 month training carried out from **29/05/2023 to 30/06/2023 at Reliance Industries LTD.** Your expertise and willingness to share your knowledge have been instrumental in enhancing my professional skills and understanding of the industry. I am grateful for the trust you placed in me and for the challenging assignments and projects that allowed me to apply my theoretical knowledge in a practical setting.

I am also grateful to the entire team at Reliance Industries Limited E&P for their warm welcome and for creating an inclusive and inspiring work environment. The collaborative atmosphere and teamwork exhibited by the employees have further contributed to my growth and learning during my internship.

The internship experience at Reliance Industries Limited E&P has provided me with invaluable insights into the oil and gas industry, along with hands-on experience in various aspects of the field. The exposure to real-world projects, technical skills development, and interaction with industry professionals have been transformative for my career aspirations.

RELIANCE INDUSTRIES LTD. – Company Profile

Reliance Industries Limited (RIL) is an Indian multinational conglomerate company headquartered in Mumbai, Maharashtra, India. Founded in 1966 by Dhirubhai Ambani, RIL is one of the largest publicly traded companies in India by market capitalization. The company operates in multiple sectors, including energy, petrochemicals, textiles, natural resources, retail, telecommunications, and media.

Here are some key highlights of Reliance Industries Limited:

1. Energy and Petrochemicals: RIL is the largest private sector company in India involved in refining crude oil and manufacturing petrochemicals. It operates the world's largest refining complex in Jamnagar, Gujarat, with a refining capacity of over 1.4 million barrels per day.
2. Retail: RIL owns and operates one of the largest retail chains in India known as Reliance Retail. It offers a wide range of products and services across various categories, including groceries, fashion, electronics, and more. Reliance Retail operates through different formats such as Reliance Fresh, Reliance Trends, Reliance Digital, and Reliance Jio stores.
3. Telecommunications: RIL launched its telecommunications subsidiary, Reliance Jio Infocomm Limited, in 2016. Reliance Jio revolutionized the Indian telecom industry by offering affordable 4G data services and disrupting the market with its competitive pricing strategies. It is now one of the largest mobile network operators in India.
4. Digital Services: RIL has expanded its presence in the digital space through various initiatives. It offers digital services such as JioSaavn (music streaming platform), JioCinema (video-on-demand platform), JioTV (live TV streaming), and JioMart (online grocery platform).
5. Exploration and Production: Reliance Industries Limited is actively involved in the exploration and production of oil and gas resources. It holds significant interests in multiple oil and gas blocks in India and abroad.
6. Sustainability Initiatives: RIL is committed to sustainable and responsible business practices. The company focuses on environmental conservation, renewable energy, and social initiatives to create a positive impact on society.

Reliance Industries Limited has achieved significant milestones and has played a crucial role in the growth of India's economy. With a strong focus on innovation, technology, and customer-centric approach, RIL continues to be a leader in various industries.

INTRODUCTION

Reliance is one of the largest exploration and production players in India having a balanced domestic conventional and unconventional hydrocarbon portfolio.

RIL's upstream business comprises the complete chain of activity starting from exploration, appraisal, development and production of hydrocarbons.

Reliance entered the Exploration and Production (E&P) business by becoming a 30% partner in an unincorporated joint venture with Shell (erstwhile BG) and ONGC in the Panna Mukta and Mid and South Tapti blocks. As of Jan 1st, 2020, our domestic portfolio comprises of conventional oil and gas blocks in Krishna Godavari and Mahanadi basins and two Coal Bed Methane (CBM) blocks, Sohagpur (East) and Sohagpur (West) in Madhya Pradesh.

Oil and gas is currently being produced from our KG D6 and CBM blocks in India.

Operations

Conventional:

In 2002, Reliance struck gas in the D1-D3 field of KG D6 block. These fields were put on production in 2009.

The KG D6 fields rank amongst one of the largest green-field deepwater oil and gas production facilities in the world. These fields were the first deepwater producing fields in India and remains among the most complex reservoirs in the world. Reliance, along with its partners, has committed ~$6 Billion for second wave of projects in KG D6 over the next few years. To supplement the existing asset base, we continue to look at new opportunities that are a strategic fit with capabilities and integrated petroleum value chain.

In December, 2020, gas production was started from the R Cluster field in Block KG D6. This is India's first ultra-deepwater field and Asia's deepest offshore gas field. In April 2021, gas production was started from Satellite Cluster field in Block KG D6.

Coal Bed Methane:

Production from the CBM blocks (Sohagpur East and West) commenced in 2017. As part of CBM development program, Reliance has drilled ~ 300 wells and set up three Gas Gathering Stations.

Reliance Gas Pipeline Limited (RGPL), one of the subsidiaries of RIL operates 300 KM of natural gas pipeline from Shahdol in Madhya Pradesh to Phulpur in Uttar Pradesh to transport gas from RIL's CBM blocks.

The CBM project is the largest surface footprint hydrocarbon project of the country.

SOFTWARE REQUIREMENTS AND SPECIFICATIONS

Software requirements :

Operating system : Windows 10 Enterprise

Coding language : python

Visualization tools : Tableau , Seeq

Hardware requirements :

Processor : Intel(R) Core(TM) i5-6200U CPU @ 2.30GHz 2.40 GHz

Installed RAM : 8.00 GB (7.90 GB usable)

System type : 64-bit operating system, x64-based processor

DATA VISUALIZATION TOOLS USED DURING THE INTERNSHIP PROGRAM -

TABLEAU:

Tableau is a powerful data visualization and business intelligence tool that enables users to analyze and present data in a visually compelling and interactive manner. It provides a user-friendly interface that allows individuals, regardless of their technical expertise, to work with data and create insightful visualizations.

One of the key strengths of Tableau is its ability to connect to a wide range of data sources, including databases, spreadsheets, and cloud services. This flexibility allows users to bring together data from multiple sources and blend them seamlessly for analysis. Tableau's data connection capabilities also include live connections, where the visualizations dynamically update as the underlying data changes.

Once connected to the data, Tableau offers a variety of tools and features to manipulate and transform the data for analysis. Users can perform data cleaning, filtering, aggregation, and calculations within Tableau, without the need for complex programming or SQL queries. This empowers users to quickly prepare and shape their data to uncover meaningful insights.

The true power of Tableau lies in its data visualization capabilities. With a drag-and-drop interface, users can easily create a wide range of charts, graphs, maps, and dashboards to represent their data visually. Tableau's visualizations are highly customizable, allowing users to control colors, labels, and other design elements to create compelling and engaging visuals. Interactive features such as filters, parameters, and tooltips enable users to explore the data and gain deeper insights by interacting with the visualizations.

Tableau also provides advanced analytics capabilities, including statistical analysis, forecasting, and clustering. These features allow users to perform more sophisticated analysis within the tool, without the need to switch to external statistical software.

Furthermore, Tableau has a strong focus on collaboration and sharing. Users can publish their visualizations and dashboards to Tableau Server or Tableau Public, making them accessible to others in their organization or the wider public. This

promotes data-driven decision-making by enabling stakeholders to explore and interact with the data visualizations on their own.

To support users in their Tableau journey, the platform offers comprehensive training resources. These include online tutorials, documentation, training courses, and a vibrant community where users can exchange knowledge, seek assistance, and share best practices.

SEEQ:

Seeq is a powerful data analytics software platform designed specifically for industrial organizations. Its primary focus is on analyzing time-series data to extract valuable insights and drive data-driven decision-making. Seeq integrates with various data sources commonly found in industrial environments, allowing users to gather and consolidate data from multiple systems. With features for data preparation and cleansing, users can ensure the quality and reliability of their data for analysis. Seeq offers advanced analytics capabilities, including pattern recognition, statistical analysis, regression, clustering, and forecasting, empowering users to uncover hidden patterns, anomalies, and relationships within their data. The platform emphasizes visual data exploration, providing an intuitive interface for users to interactively explore and manipulate data through charts, trends, and summaries. Seeq also facilitates collaboration by enabling users to share analysis workbooks, insights, and results, fostering knowledge exchange and data-driven decision-making across teams. Additionally, Seeq offers industry-specific applications tailored to sectors like manufacturing, energy, pharmaceuticals, and utilities, providing pre-built analytics templates and workflows to address common challenges within those industries. Overall, Seeq enables industrial organizations to leverage their operational data effectively, leading to improved performance, reduced costs, and increased efficiency.

TASKS PERFORMED :

1. Tableau training :

- Started by understanding the basics of Tableau, including its interface, terminology, and key concepts like data connection, dimensions, measures, and visualizations.
- Took advantage of Tableau's online resources, including video tutorials, documentation, and knowledge base. These resources provide a solid foundation for learning Tableau and its functionalities.
- Worked with sample datasets provided by Tableau or available online. Practiced importing data, creating basic visualizations like bar charts and scatter plots, and applying filters and calculations.
- Once I had a good grasp of the basics, delved into Tableau's advanced features. Learned about advanced calculations, blending data from multiple sources, using parameters and table calculations, working with maps, and implementing advanced visual analytics techniques.

2. Created a tableau dashboard using a sample dataset downloaded from Kaggle :

- About Dataset

  This dataset contains a list of video games with sales greater than 100,000 copies. It was generated by a scrape of vgchartz.com.

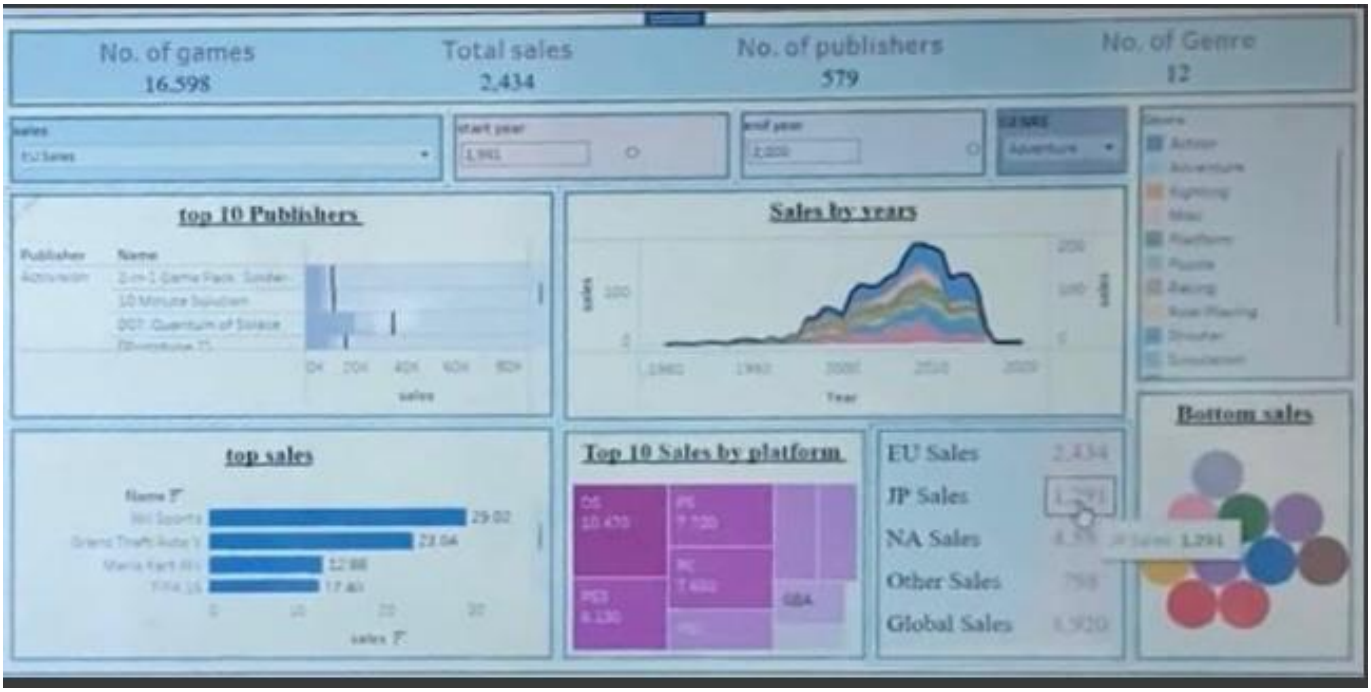
  Fields include :

  - Rank - Ranking of overall sales
  - Name - The games name
  - Platform - Platform of the games release (i.e. PC,PS4, etc.)
  - Year - Year of the game's release
  - Genre - Genre of the game
  - Publisher - Publisher of the game
  - NA_Sales - Sales in North America (in millions)
  - EU_Sales - Sales in Europe (in millions)
  - JP_Sales - Sales in Japan (in millions)
  - Other_Sales - Sales in the rest of the world (in millions)
  - Global_Sales - Total worldwide sales.

It is based on BeautifulSoup using Python. There are 16,598 records. 2 records were dropped due to incomplete information.

- Created a dashboard using the above dataset by using different functionalities of tableau :

  - Obtained the necessary data related to video game sales. This includes information about game name, year, sales figures, genres, publishers, etc.
  - Then we perform data preparation process , that is to clean and organize the data, to ensure it's in a suitable format for analysis. This involves removing duplicates, handling missing values, and performing data transformations or calculations as required.
  - Firstly we opened Tableau and connected to the dataset . Tableau supports various data sources, including spreadsheets, databases, and online platforms.
  - Determined the layout and structure of my dashboard. Considered the key metrics and insights which were important to be displayed and planned the visual arrangement of different components in the form of charts, filters, and text boxes.
  - Started building the visualizations for my dashboard. Tableau provides a wide range of chart types, including bar charts, line charts, scatter plots, maps, and more.
  - Enhanced the dashboard by adding interactivity elements. This included filters, parameters, and actions that allow users to dynamically explore and analyze the data. For example, we added filters to select specific time periods, genres, or platforms and observe how the sales metrics change accordingly.
  - Identified the important metrics which was required to be highlighted and included them in the dashboard. This includes total sales, sales by platform or genre, top-selling games, etc.
  - Formatted the dashboard to make it visually appealing and easy to understand. Adjusted colors, fonts, labels, and titles to ensure consistency and clarity.



3. Gas sales prediction model using slug catchers parameter in seeq :

- Analyzing and studying the previous usecase of gas sales prediction model by using R cluster and SAT cluster:

  - Model was created by using individual wells based on choke Pressures, THP, PCVPG and WGFMQG.
  - They created a model, by considering all the wells of a cluster (RDWPLEM & S-Manifold separately) and understand the impact of these wells over the header pressure for a cluster.
  - Based on the new input signals they compared the Actual and predicted sales ratios for more accuracy.
  - The five parameters which were used to create this model are as follows :
  - 'PCVPG','PTT1PT','PTT4PT','WGFMQG' and 'SMAN'.
  - From these parameters WGFMQG was predicted.
  - Plotted the Daily Sales & Actual Sales and the difference was minor (<0.05) except for few days.
  - Dashboard :
    Daily Sales and Predicted Sales - 1-week data was populated.
    Daily Sales and Predicted Sales (i Day) - Cumulative sales are displayed on hourly basis (which includes the technical difficulties in the process and project the revised Sales values on hourly basis.) rather than plotting the sale rate on hourly basis for entire day.

- To achieve better accuracy for prediction of gas sales new usecase was formulated :

  Instead of using five parameters used in the previous model , we will consider slug catcher, inlet pressure and totalized gas flow (of individual wells) to predict the gas sales. So the step by step procedure is as follows :

  - **First we created a worksheet in the seeq workbench**
    SAT-Cluster (5 wells) :
    It consists of **Totalized gas slug catcher1**, **PTT1PT** (inlet pressure of individual wells) and we created a formula by adding the gas flow of all the wells using seeq tool which is **GASTOT2**.

    R-Cluster (6 wells) :
    It consists of **Totalized gas slug catcher2**, **PTT1PT** (inlet pressure of individual wells) and we created a formula by adding the gas flow of all the wells using seeq tool which is **GASTOT1**.

    MJ-Cluster (8 wells) :
    It consists of **Totalized gas slug catcher3**, **PTT1PT** (inlet pressure of individual wells) and we created a formula by adding the gas flow of all the wells using seeq tool which is **GASTOT3**.

  - **Model Building in Seeq Data Lab (SDL)-**
    Split the data into training and testing sets. Build separate prediction models for each cluster using the selected features. Choose an appropriate machine learning algorithm, such as linear regression, decision trees, random forests, or gradient boosting, to train the models on the cluster-specific data.

    Experimenting with the different Models in SDL:
    - Linear Regression
    - Decision Tree Regressor
    - Random Forest Regressor
    - KNN Regressor
    - Gradient boost Regressor
    - Naive Bayes

    Metrics used to evaluate the model
    - R2 Score

- MAE
- RMSE

We achieved r2 score (One of the major Validation metrics) around 0.98 for train and test both for almost all the above algorithms.

As **Random forest regressor** has the highest accuracy (R2 Score) therefore it was chosen to predict the gas sales.

- Using Random forest regressor predicted the values of the gas sales of R-cluster and SAT-Cluster. Then we added the predicted values of both the clusters which is predtot (SATPRED + RPRED)

  Then we compared the values of **Actual slug catcher** with **predicted SC** and achieved a difference of less than 0.1.

- **Compared the actual sales and predicted sales :**
  The formula used for predicted sales is [Predicted SC – (Extrapolated Internal Consumption for current day + Flare)].

OUTCOME AND CONCLUSION

The primary outcome of my internship is the successful development of a gas sales prediction model. I have demonstrated my proficiency in data analysis, machine learning, and predictive modeling techniques.

Throughout the internship, I have honed the data analysis skills by gathering and preprocessing historical data related to gas sales. This includes handling missing values, performing data transformations, and feature engineering.

I gained expertise in using machine learning algorithms to build predictive models. This includes selecting appropriate algorithms, training the models, and evaluating their performance using relevant metrics.

My internship involved the application of cluster analysis techniques to identify distinct customer segments or sales patterns within the data. This allowed for personalized gas sales predictions for different well groups.

The outcome highlights the proficiency in evaluating the performance of the gas sales prediction model using evaluation metrics like mean absolute error (MAE), root mean square error (RMSE), or coefficient of determination (R-squared).

The gas sales prediction model holds practical business applications, such as enabling better resource planning, inventory management, and decision-making related to gas sales.

My work during the internship has the potential to create a positive impact on the organization or industry by providing accurate gas sales predictions, leading to improved efficiency and optimized operations.

As an intern, I have not only developed technical skills but also enhanced my teamwork, communication, and problem-solving abilities, contributing to my overall personal and professional growth.
The successful completion of the gas sales prediction model project opens up possibilities for further research and future opportunities in data science, machine learning, and related fields.

The outcome of my internship program reflects a successful journey of learning, exploring, and applying data science techniques to address real-world challenges in the gas sales domain. My efforts and contributions during the internship have laid the foundation for continued growth in the field of data analytics and predictive modeling.