Yagiz SEZERSAN

Thus, I will use two datasets:

\* 1 - Dataset that provides information about the total world production of mineral raw materials

https://www.world-mining-data.info/?World\_Mining\_Data\_\_\_Data\_Section

\* 2 - Dataset from their database("VanPetroDatabase"). Since they know I am going to submit it to college,

they manipulated the data and if they like what I am trying to do, they will let me to do analysis with their real data in future

Dataset 1

This dataset will serve as valuable resources for understanding the world mining landscape, identifying trends, making informed decisions, and drawing meaningful conclusions about the mining industry on a global scale.

1. PetroDataset.xlsx:

This dataset provides information about the total world production of mineral raw materials for a specific time period. It likely includes data on the overall production quantities of various minerals on a global scale. You can analyze trends, identify major contributors to global production, and understand the distribution of mineral resources worldwide.

2. GroupsCommodities.xlsx:

This dataset presents data on the production of mineral raw materials categorized by the specific type of mineral.This data can be useful for identifying opportunities and challenges in different mineral markets.

3. IncomeLvl.xlsx:

This dataset includes data on the annual per capita income of different countries or regions involved in the mining industry. Analyzing income levels alongside mineral production can provide insights into the economic impact of the mining sector on individuals and communities.

4. Political stability.xlsx:

This dataset offers information about the political stability of various countries or regions where mining activities occur. Analyzing political stability data can help assess the risks associated with mining investments and operations in different regions.

Info about final cleaned dataset and features :

- year (continuous): The year in which the order was collected or recorded.

- value (continuous): A numerical value associated with volume of material production.

- income\_group\_encoded (categorical): An encoded representation of the income group to which the sample belongs.

- groupscommodity\_encoded (categorical): An encoded representation of the commodity group.

- region\_encoded (categorical): An encoded representation of the region.

- politicsest\_encoded (categorical): An encoded representation of the political situation.

Dataset 2

Dataset from their database("VanPetroDatabase.xlsx"). Since they know I am going to submit it to college, they manipulated the data and if they like what I am trying to do, they will let me to do analysis with their real data in future

Info about dataset and features :

- OperationID: Unique identifier for each order.

- Quantity: Quantity of material.

- TotalWeight: Total weight of the material.

- OrderCompleteHour: Total hour duration at which the order was completed.

- Invoice\_number: Invoice number associated with the order

- Profit: Profit earned from operation.

- Price: Price for service given to customers.

- Rating: Rating given by customer about their satisfactory level.

- Customer.companyName: Name of the customer's company.

- Customer.Country: Country where the customer is located.

- Service.Service: Name of the mining service.

- DelayIssues.DelayType: Type of delay issue.

- Order.Material/RockType: Type of material or rock extracted in the order.

- Order.SubmissionDate: Date of order submission.

- Order.CompletionDate: Date of order completion.