

End-to-End Project: Minerals and Energy Resources

Dataset Link: <https://shorturl.at/fbepE>

LinkedIn: <https://www.linkedin.com/in/vikaspandey001/>

Github: <https://github.com/vikaspandey-001>

Project Overview

This project aims to analyze mineral and energy resources using a dataset based on the System of Environmental-Economic Accounting (SEEA) Central Framework (2012). The dataset contains details about stocks and flows of resources such as oil, gas, and minerals, measured in various units. The analysis focuses on converting units for consistency, cleaning and preparing the data, and visualizing the insights using Microsoft Power BI.

Dataset Description

Dataset Name: Mineral and Energy Resources

Source:

- [Classification Mapping Document](#)
- [OECD Working Paper](#)

Dataset Key Details

- **Unfiltered Data Points:** 11,731
- **Columns in Processed Data:**
 - Reference Area: Country origin of resources.
 - Unit measures: Measures of unit in which resource founded.
 - Resource2: Name of minerals and resources that founded.
 - Stocks and flows: Contains 10 different types:
 - Closing stock
 - Discoveries

- Discrepancies
 - Downward reappraisals
 - Extractions
 - Opening stock
 - Reclassifications
 - Total additions to stock
 - Total reductions to stock
 - Upward reappraisals
 - Class:
 - There are 6 classes present:
 - Class A: Commercially recoverable resources
 - Class A_B: Commercially recoverable resources + potentially commercially recoverable resources
 - Class ATC: Commercially recoverable resources + potentially commercially recoverable resources + non-commercial and other known deposits
 - Class B: Potentially commercially recoverable resources
 - Class B_C: Potentially commercially recoverable resources + non-commercial and other known deposits
 - Class C: Non-commercial and other known deposits
 - Time Period: Year in which the resources were founded
 - Observation value: Observation value of the respective resource
 - Observation Status: Observation status of the respective resource.
 - Unit mult and Unit multiplier: The Unit and measure in which the resource founded.
-

Project Steps

Step 1: Data Collection

- Collected the dataset containing 11,731 rows of raw data from reliable sources, ensuring it was comprehensive and aligned with SEEA and UNFC standards.
-

Step 2: Data Cleaning and Preprocessing (Using Excel)

1. Duplicate Column Check:

- Identified and removed duplicate columns to ensure a clean structure.

2. Data Refinement:

- Checked for duplicates row-wise and column-wise.
- Ensured no loss of essential data during refinement.

3. Unit Conversion:

- Problem:** The dataset used multiple units of measure (Barrels, Tonnes, Cubic Meters), which needed standardization.
- Solution:** Converted all measurements to a single unit: **Cubic Meters**.
- Conversions:**
 - 1 Barrel = 0.159 Cubic Meters.
 - 1 Volume (Cubic Meters) = Mass (Tonnes) ÷ Density (Tonnes/m³).

4. Excel Formula for Conversion:

```
=IF(C2="BBL_US", R2*0.159, IF(C2="T", R2/VLOOKUP(F2, $X$3:$Y$19, 2, FALSE), R2))
```

- Converted the **UNIT_MEASURE** column into a new column called **Temp Unit**.
- Pasted the values into a final column named **UNIT_M_CUBE** for use in analysis.

5. Value-Based Dataset:

- Created a new sheet with value-only data to avoid errors when imported into Power BI.
-

Step 3: Data Import (Power BI)

- Imported the Excel Sheet (Data without Formula Sheet) in Power BI.
-

Step 4: Visualization (Power BI Dashboard)

Created an interactive dashboard with the following key insights:

- Total Unit wise Stocks and flows:
 - Chart used: Clustered Column chart.
 - Sum of Total Unit by Resource:
 - Chart used: Pie chart
 - Time Period by Resource found:
 - Chart used: Line chart
 - Country wise Total Unit:
 - Chart used: Funnel chart
 - Count of Stocks and flows by Reference area and Class:
 - Chart used: 100% stacked column chart
-

Key Challenges

1. **Unit Standardization:**
 - Addressed inconsistencies in unit measures (BBL_US, Tonnes) with precise conversion formulas.
 2. **Handling Large Data:**
 - Optimized dataset for Power BI processing by using value-only data.
-

Key Learnings

- Gained expertise in unit conversion techniques for environmental datasets.
 - Developed advanced report page in Power BI
-

Final Deliverables

1. **Power BI Report File (.pbix):**

An interactive dashboard showcasing resource stocks and flows.

2. **Documentation:**

A step-by-step explanation of data processing, modeling, and insights.

3. **Insights Summary:**

- Highlighted trends in resource stocks over time.
 - Identified regional variations in resource distribution.
 - Provided actionable insights for policymakers and environmentalists.
-