📊 India's Principal Commodity-wise Export Dashboard (2022–23)

Welcome to the **Interactive Data Analysis Dashboard** for India's Principal Commodity-wise Export data (2022–2023)! This project is a comprehensive analysis and visualization tool developed using **Streamlit** to offer insightful, responsive, and user-friendly interaction with real-world export data.

For More Info Access this Github Link: https://github.com/AtharvaKale1/India-s-Principal-Commodity-Wise-Export-Dashboard-2022-23

⚠️ **Note:** This is **Phase 1** of our ongoing project. More enhancements, advanced analytics, and features are coming soon!

🚀 Project Objective

To analyze, cluster, and visually explore India's commodity-wise exports using machine learning and data visualization techniques—helping users gain business and trade insights interactively.

📂 Dataset Overview

- **Dataset Title:** Principal Commodity-wise Exports (2022–23)
- **Source:** Government of India (DGCI&S)
- **Format:** Excel (.xlsx)
- **Key Columns:**
    - COMMODITY_NAME
    - COUNTRY
    - UNIT
    - QUANTITY_KGS
    - VALUE_USD_MILLION
    - PRICE_PER_KG
    - CLUSTER (generated using KMeans clustering)

🧐 Key Features & Insights

✅ **EDA & Data Cleaning**

- Handled missing data and standardized units.
- Added computed columns like PRICE_PER_KG.

✅ **Machine Learning**

- Applied **KMeans Clustering** to identify export patterns.
- Used **PCA** for dimensionality reduction and visual representation.

✅ **Business Insights**

- Top 10 most expensive/cheapest commodities by price per kg.
- Cluster-wise average prices and high-value exports.
- Country-wise export distribution and total export values.

✅ **Streamlit Dashboard**

- Clean UI and interactive filters.
- Responsive plots (Pie, Bar, Box, Line).
- User can explore:
    - Commodity performance
    - Country-wise insights
    - Cluster distributions

🛠️ Tech Stack

Category	Tools Used
📊 Data Analysis	pandas, numpy, matplotlib, seaborn
🤖 Machine Learning	scikit-learn
📉 Visualization	plotly, matplotlib, seaborn
🌐 Dashboard	Streamlit

📸 Sneak Peek

📍 Dashboard Page	✨ Description
📌 Overview	Summary stats, total export value
📈 Charts	Bar, pie, and line charts per commodity/country
🔍 Cluster View	ML-based export segmentation
🌎 Country View	Filter by export partner countries

📁 Project Structure

```
🗓️ DS_ML_Export_Analysis/
├── app.py              # Streamlit app
├── Cleaned_Dataset.xlsx  # Final dataset with clustering
├── cluster_model.pkl     # Saved KMeans model
├── requirements.txt      # Dependencies
└── README.md             # Project documentation
```

⚙️ Installation & Run Locally

```
# 1. Clone the repo
git clone https://github.com/yourusername/DS_ML_Export_Analysis.git
cd DS_ML_Export_Analysis

# 2. Install dependencies
pip install -r requirements.txt

# 3. Run Streamlit app
streamlit run app.py
```

💡 Future Scope (Next Phases)

- Add **forecasting** using time series models.
- Integrate **RAG + LLM**-based analytics assistant.
- Use **interactive maps** for geospatial trade flows.
- Enable **user uploads** for dynamic commodity files.

🤝 ♂️ Team & Contributions

Name	Role
Atharva Kale	Data Science Lead & Developer

We welcome contributions and feature requests! Feel free to fork, contribute, or open issues. 💬

📄 License