**Textile Data Analysis**

**1. 📥 Importing Required Library**

import pandas as pd

The script starts by importing pandas, which is essential for handling data.

**2. 📂 Reading the Dataset**

df = pd.read\_csv(r"C:\\Users\\Samiksha\\Downloads\\WICVOCDataAll\_030223.csv\\WICVOCDataAll\_030223.csv")

* A CSV file containing textile data is loaded.
* Make sure the path exists on your machine if rerunning this code.

**3. 🔍 Initial Exploration**

df.head()

df.info()

df.describe()

df.columns

* head(): Shows the first few rows of the dataset.
* info(): Displays data types, number of non-null entries.
* describe(): Statistical summary for numeric columns.
* columns: Lists all column names.

**4. 📉 Missing Value Analysis**

df.isnull().sum()

This checks for the number of missing values in each column. Done twice for cross-verification.

**5. 🧾 Exploring Categorical Variables**

for col in df.select\_dtypes(include='object').columns:

print(f"{col}: {df[col].nunique()} unique values")

Prints the count of unique values in each categorical column.

**6. 🧹 Data Cleaning**

**Remove Duplicates**

df.drop\_duplicates(inplace=True)

Removes repeated rows.

**Handling Missing Data (Suggested)**

# df.fillna(method='ffill', inplace=True)

This line is commented out, suggesting forward-filling could be used to handle null values.

**7. 📊 Visualizing Numerical Data**

df.hist(figsize=(15, 10), bins=20)

plt.tight\_layout()

plt.show()

Creates histograms for all numerical columns to understand their distributions.

**8. 📈 Categorical Visualization (Suggested)**

# sns.countplot(data=df, x='CategoryCol')

This is a suggested count plot using Seaborn, currently commented.

**9. 📉 Group-Based Aggregation**

df.groupby('means\_of\_exchange')['ps\_value\_stuivers'].mean()

Calculates the average value (ps\_value\_stuivers) grouped by the method of exchange.

**10. 🏭 Company Count**

df['company'].nunique()

Finds how many unique companies exist in the dataset.

**✅ Summary & Insights**

* Dataset contains trade or textile data with both categorical and numerical fields.
* Duplicates removed; missing values assessed.
* Numerical distributions plotted.
* Aggregation gives average values based on exchange method.
* Unique company count extracted.

**🧠 Recommendations**

* Add Seaborn-based plots for categorical analysis.
* Explore time trends if date fields are present.
* Study company-specific or regional performance.
* Look into relationships using correlation heatmaps.