Trade and Customs Analytics Report by da-22

# Executive Summary

This report presents a comprehensive analysis of trade and customs data aimed at understanding import trends, taxation efficiency, and logistics performance. Key insights were drawn from rows of 77783/columns of 18, with a focus on HS Codes, CIF/FOB values, tax contributions, and country origins. The goal was to inform better policy decisions and optimize customs operations.

# Introduction

Purpose of the Report:  
To analyze trade and customs data, extract actionable insights, and provide data-driven recommendations for improved logistics, taxation policies, and import compliance. A secondary goal is to evaluate risk signals across HS codes and countries for more robust policy enforcement.  
  
**Dataset Overview:**

1. Source: DataVerse Africa
2. Time Frame Covered: July – August
3. Key Columns: HS Code, Country of Origin, FOB, CIF, Mass (KG), Tax Amount, etc.
4. Total Observations: **77,783.**
5. Total Unique Importers: **11027.**
6. Total Unique HS Categories (2-digit): **63**

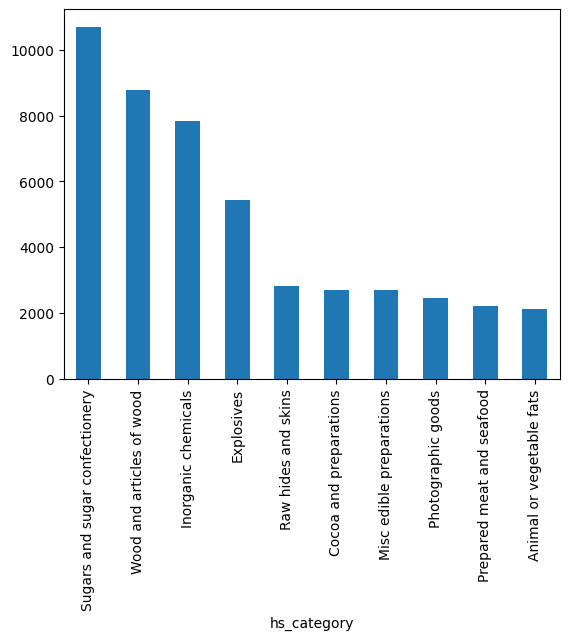
# Data Cleaning Summary

A thorough data cleaning process was conducted to ensure consistency and accuracy. Key steps included:

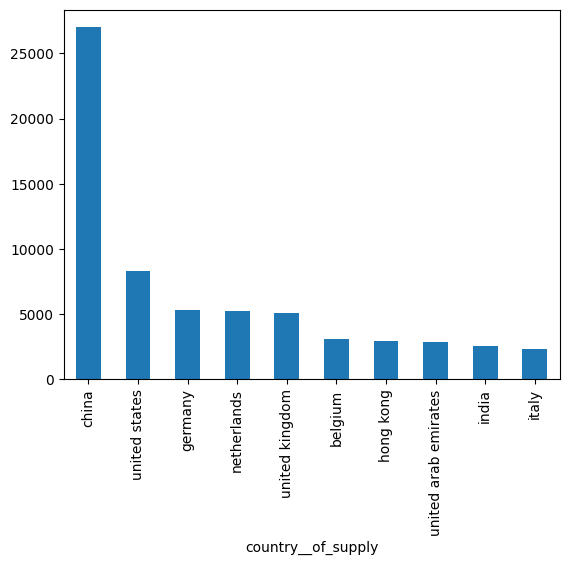
1. Rectified formatting inconsistencies in HS Codes.
2. Truncated to 6 digits, aggregated using first 2 digits to ensures consistency with WCO classification.
3. We used the first two digit of the HS code to assign labels to them.
4. Converted monetary fields (FOB, CIF, Tax) to numeric types.
5. Handled missing and duplicate values.  
   Normalized date and country formats Checked for the correlation between country of supply and country of Origin, which turns out very strong. With this we were able to forward fill the empty values in the country of supply.
6. Used high Pearson correlation (r ≈ 0.998) with Country of Origin to justify substitution.
7. For the receipt date column, we noticed the date were far past and the years were 1866, 1867, 1868, 1869. To standardized the date we formatted them forwarding by 2022, 2023,2024 and 2025 respectively each.
8. Dropped the receipt date missing values due to uniqueness and lack of redundancy impacted less than **1.67%** of total records.
9. Tools Used: Python (Pandas), Excel

# Exploratory Data Analysis (EDA)

1. Import Overview by HS Code

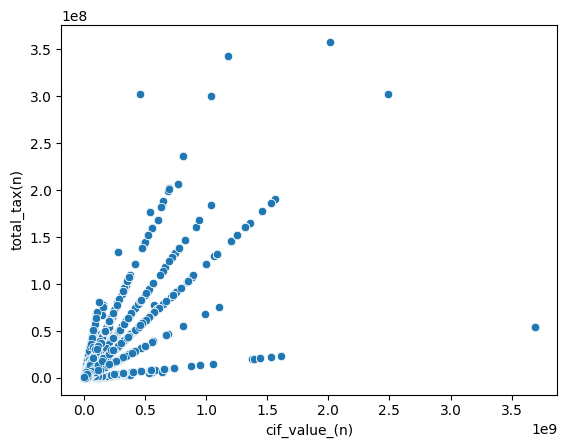


Sugars and sugar confectionery happens to be the most traded goods  
  
B. Top Countries of Origin

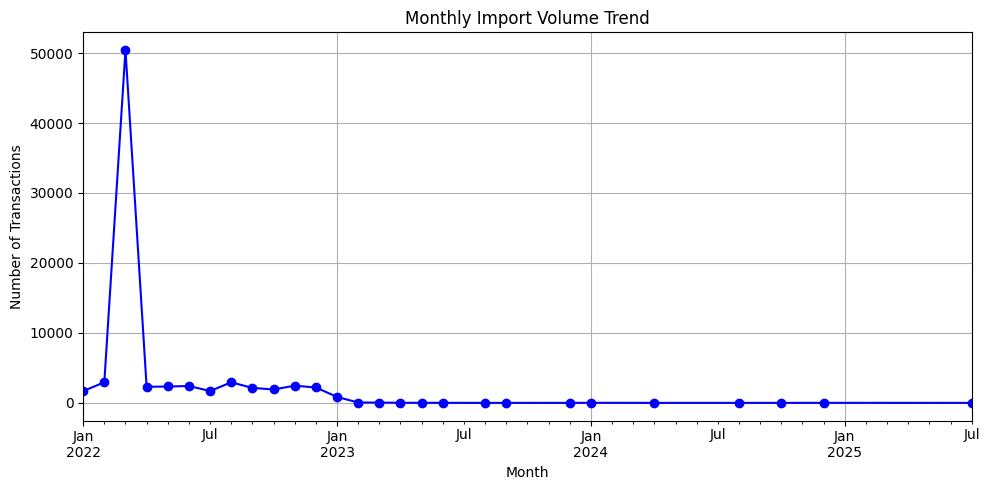
  
- Insight: China leads in volume; these are the top ten traded country by volume

# Advanced Analysis & Policy Evaluation

1. Correlation Insights  
    CIF vs Tax:



B. Trend Analysis  
Line chart showing monthly volume trends



# KPI Evaluation

**Import Volume & Value KPIs**

1. Total Import Value (FOB): $1,702,257,979,596.00
2. **Total Import Value (CIF): $**1,820,915,950,052.00
3. Average Import Value: Avg FOB per transaction:
   1. **FOB:** $22,253,774.59
   2. **CIF:** $23,805,001.11
4. **Average Mass per Transaction: 17,391.14 kg**

Top Importing Countries by FOB:

|  |  |  |
| --- | --- | --- |
| Countries | Total FOB ($) | % of Total |
| China | 667,423,362,775 | 39.2% |
| Lebanon | 114,343,994,910 | 6.7% |
| Italy | 110,300,513,789 | 6.5% |
| United Kingdom | 92,498,330,110 | 5.4% |
| Korean Republic | 78,740,311,239 | 4.6% |
| United States | 74,733,363,981 | 4.4% |
| Sweden | 68,901,520,503 | 4.0% |
| South Africa | 61,894,973,861 | 3.6% |
| Hong Kong | 39,777,956,468 | 2.3% |
| India | 10,452,019,535 | 0.6% |

China dominates import value, contributing over **39%** of total FOB. The top 3 countries account for over **52%**, suggesting import dependence on a narrow group of trading partners.

The **import value distribution** is highly **right-skewed**, with a small number of countries responsible for a disproportionate share of trade.

**Taxation & Revenue KPIs**

1. Total Tax Collected: $ 251,393,817,308.00
2. Average Tax per Transaction: $ 3,286,494.42
3. Tax Efficiency Ratio:
4. Tax-to Ratio: 14.77%
5. Tax-to-CIF Ratio: 13.81%

Tax collection appears proportionate to FOB values, with a fairly stable average tax-to-value ratio around 14%. However, certain **HS categories and countries** significantly deviate from this ratio, indicating potential under-taxation or exemptions. Example: “Raw hides and skins” has a **Tax-to-CIF ratio of just 2.0%.**

**Logistics & Shipment KPIs**

1. Total Shipments: 76493
2. **Average Containers per Transaction:** 44.52
3. **Most Common Container Size:** 40ft
4. Total Weight of Imports by top Country of Supply:

|  |  |  |
| --- | --- | --- |
| Country | Total Mass (KG) | % of Total |
| China | 682472717 | 51.3% |
| India | 139141749 | 10.5% |
| Italy | 134743493 | 10.1% |
| United State | 77375724 | 5.8% |
| Lebanon | 63275649 | 5.8% |
| United Kingdom | 33270141 | 2.5% |
| Hong Kong | 32952076 | 2.5% |
| Korean Republic | 23888790 | 1.8% |
| Germany | 23672158 | 1.8% |
| South Africa | 15593207 | 1.2% |

Mass distribution also reflects China’s dominance, accounting for over **51%** of total import weight. This aligns with FOB value trends and underscores China's centrality in trade volume and logistics dependency.

# Key Findings & Recommendations

Key Findings:  
**1. Total Tax Revenue**

* The dataset reflects a **total customs tax revenue** of approximately **$251.4 billion**.

**2. Top Product Categories by Tax Contribution**

* The product category **"Inorganic chemicals"** generated the **highest share of total tax revenue**, contributing approximately **17.5%** of the entire customs tax.
* The **Top 5 highest tax-contributing categories** are:

|  |  |
| --- | --- |
| HS Category | % Total Tax Revenue |
| Inorganic chemicals | 17.5% |
| Wood and articles of wood | 12.0% |
| Sugars and sugar confectionery | 10.8% |
| Miscellaneous edible preparations | 10.1% |
| Iron and steel | 6.7% |

**3. Under-Taxed Categories**

* Several categories show significantly **low tax-to-CIF value ratios**, suggesting potential under-taxation:

|  |  |  |  |
| --- | --- | --- | --- |
| HS Category | Avg CIF Value ($) | Avg Tax ($) | Tax/CIF Ratio |
| Raw hides and skins | $22.4 million | $442,322 | 4.0% |
| Tobacco | $6.2 million | $419,918 | 6.8% |
| Printed books | $15.4 million | $993,925 | 7.3% |
| Animal or vegetable fats | $8.9 million | $690,661 | 7.9% |
| Cork and articles of cork | $20.7 million | $1.6 million | 9.3% |

* These sectors may be **subject to policy leniency**, **misclassification**, or **tax evasion risks**.

**4. High-Risk Importing Countries (Low Avg. Tax Rate)**

* Some countries are importing large CIF values but generating **exceptionally low average tax rates**, potentially indicating:
  + Favourable trade treaties
  + Possible misreporting or fraud
  + Gaps in current tax policy

|  |  |  |  |
| --- | --- | --- | --- |
| Country | CIF Value ($) | Tax Paid ($) | Avg. Tax Rate |
| Luxembourg | ₦40.97 billion | ₦607 million | 1.48% |
| Argentina | ₦49 million | ₦1 million | 2.14% |
| Thailand | ₦22.5 billion | ₦486 million | 2.17% |
| Portugal | ₦3.39 billion | ₦78 million | 2.30% |
| India | ₦108.8 billion | ₦6.6 billion | 6.08% |
|  |  |  |  |

**Recommendations**

1. **Optimized Tax Policy for Under-Taxed Categories**
   * Conduct a thorough review of **HS classifications** and declared values for categories such as *Raw hides and skins*, *Tobacco*, and *Cork articles*.
   * Implement **benchmark pricing** and **minimum tax thresholds** for historically under-taxed goods.
   * Consider using AI/ML models to flag unusually low tax/CIF ratios in real time.
2. **Scrutiny on Imports from High-Risk Countries**
   * Establish a **customs intelligence unit** to monitor trade from countries like **Luxembourg**, **Thailand**, and **India**, where import volumes are high but taxation is low.
   * Introduce **audits and data cross-verification** with trading partners to identify potential misreporting.
   * Review existing **bilateral trade agreements** and waivers for loopholes that might reduce tax yield.
3. **Invest in Digital Customs Infrastructure**
   * Strengthen **automated valuation** and **risk assessment** tools.
   * Enable **real-time alerts** for transactions where tax/CIF ratios fall below threshold levels.
4. **Public-Private Collaboration**
   * Work with importers and brokers to **educate and enforce transparency**, reducing unintentional underreporting.

**Conclusion**

The comprehensive analysis of the trade and customs dataset reveals systemic imbalances, untapped revenue potential, and logistical concentration risks that call for urgent policy and operational attention.

Key findings highlight a **highly skewed tax structure**, where a small number of product categories notably Inorganic Chemicals, Wood Articles, and Sugar Confectioneries, account for a disproportionate share of the total tax revenue. Conversely, several **high-CIF categories** such as Raw Hides and Skins, Tobacco, and Printed Books demonstrate **alarmingly low Tax-to-CIF ratios**, suggesting potential **under-taxation**, policy leniency, or deliberate misreporting. These inefficiencies translate into lost revenue and compliance vulnerabilities that compromise fiscal stability and fair trade enforcement.

Moreover, the dataset exhibits strong concentration in both trade value and volume. Over **50% of import weight and nearly 40% of FOB value** originate from **China**, underscoring a **strategic dependency risk** in logistics and supply chains. This concentration, while efficient for scale, exposes national trade resilience to geopolitical and economic disruptions.

From a data infrastructure standpoint, this analysis reveals opportunities to **integrate advanced analytics and automation into customs operations**. Predictive models, anomaly detection algorithms, and tax benchmarking systems can proactively flag suspicious transactions, reduce manual oversight, and improve revenue assurance.