# AIV, HIV/HRDT Treatment Distribution Analysis (Africa)

By Swapnil Herwadkar

## INTRODUCTION

This study focuses on different regions (majorily African) to analyze the distribution of Antiretroviral (ARV) treatments and HIV/HRDT lab testing supplies.

HIV/AIDS remains a significant public health issue globally, with a particularly profound impact on the African region.

There are 25.6 million people living with HIV in the African region. East and Southern Africa account for 20.8 million of these cases, while West and Central Africa have 4.8 million. In 2022, about 380,000 people died from AIDS-related illnesses, and 760,000 were newly infected with HIV. (WHO, 2022)

The healthcare supply chain is essential for delivering medical products efficiently and on time. As a supply chain analyst at our healthcare company, I analyzed ARV and HIV lab shipments to various countries. This report aims to provide key insights, identify areas for improvement, and offer practical recommendations based on the data analysis.

This report primarily focuses on shipment distribution by country, vendor performance, delivery timeliness, and the popularity of brands and different dosage forms. Using Tableau to create an interactive dashboard, I have visualized these insights with various charts and graphs, providing a clear and comprehensive view of our supply chain operations.

The following sections will detail the analysis and findings:

1.Shipment Distribution by Country: Understanding the geographical spread and volume of shipments.

2.Vendor Performance: Evaluating vendor efficiency and identifying top performers.

3.Delivery Timeliness: Assessing the punctuality of deliveries to clients and calculating the percentage of late deliveries.

Through this detailed analysis, I aim to provide a data-driven foundation for strategic decision-making to optimize our supply chain processes and improve overall service delivery.

# METHODS

### DATA COLLECTION

Data was sourced from Supply_Chain_Shipment_Pricing_Dataset.csv file available on shipment records covering the period from June 2006 to August 2015.

The dataset included fields such as id, project cod, pq #, po / so #, asn/dn #, country, managed by, fulfill via, vendor inco term, shipment mode, pq first sent to client date, po sent to vendor date, scheduled delivery date, delivered to client date, delivery recorded date, product group, sub classification, vendor, item description, molecule/test type, brand, dosage, dosage form, unit of measure (per pack), line item quantity, line item value, pack price, unit price, manufacturing site, first line designation, weight (kilograms), freight cost (usd) and line item insurance (usd).

### DATA PREPARATION

Data cleaning was involved in the weights(kilograms) column. Every non-numerical value was considered to be garbage value. Garbage values were then filtered to maintain integrity of dataset.

***Handling Garbage Data***

- **Identification:** I identified garbage data as non-numeric values in the "weight (kilograms)" column and other nonsensical entries.

- **Strategy:**

- Marked garbage data entries using a helper column.

- Filtered out rows marked as garbage to ensure the quality and integrity of the dataset.

Formula: =IF(ISNUMBER(X), "No", "Yes")

X is respective cell.

Data integrity was maintained by re-naming column titles for better understanding and uniformity. Data cleaning was done by removing irrelvant/not required columns. After analysis, only fields such as ID, Country, Shipment Mode, Scheduled Delivery Date, Delivered To Client Date, Product Group, Vendor, Brand, Dosage Form, Unit Of Measure, Pack Price, Unit Price, and Treated Weight were maintained as relevant data.

Data was transformed to ensure consistency in date formats and categorical values.

Missing data handling were performed on all the relevant fields. The method I chose was to delete record for all the fields except "Treated Weight(kilograms)". For "Treated Weight" column, I performed missing data handling but chose the mode of "mean" to fill the missing records.

Reason: After analysis, the garbage values consisted the lowest subset values across all categories. So after considerations and comparisons, concluded that most of the top indicators

remain same in all categories. Hence, removed the garbage values from the columns "Shipment Mode", "Treated Weights" and "Dosage Form".

Data integrity was checked by comparing two columns, "Delivered To Client Data" and "Delivery Recorded Date" to understand if the values are same.

Formula: =IF(ISNA(VLOOKUP(X, Yx:Yy, 1, FALSE)), "Difference", "Match")

X being value in first column, Yx:Yy being the range

The values were different, so after analysis I concluded that "Delivery Recorded Date" will be relevant data.

- Column Changes:

In the Country Column, replaced "CÃ´te d'Ivoire" with "Ivory Coast" for aesthetics and uniformity.

In the Shipment Mode Column, replaced "Air" value with "Air Freight", replaced "Air Charter" with "Air Freight" and included both in Air Freight mode. Also replaced "Ocean" value with "Ocean Freight" to maintain uniformity and standardisation.

Sorted the data sheet by ascending order of IDs

## TOOLS AND SOFTWARES

Data preparation was done in Excel, XLMiner.

Data Analysis and Visualization was done in Tableau.

## DATA ANALYSIS TECHNIQUES

Imported CSV in Tableau

Created calculated fields:

Shipment Delay:

DATEDIFF('day', [Scheduled Delivery Date], [Delivered To Client Date])

This will help in analyzing delays in shipments.

Total Pack Price:

[Unit Price] * [Unit Of Measure (per pack)]

To calculate the total price per pack.

Delivery Efficiency:

IF [Delivered To Client Date] <= [Scheduled Delivery Date] THEN "On Time" ELSE "Late" END

This will categorize shipments as "On Time" or "Late".

Late Delivery:

IF [Delivered To Client Date] > [Scheduled Delivery Date] THEN 1 ELSE 0 END

Delivery Status:

IF [Delivered To Client Date] > [Scheduled Delivery Date] THEN "Late" ELSE "On-Time" END

Total Shipments:

WINDOW_SUM(COUNT([ID]))

Late Shipments:

IF [Delivery Status] = "Late" THEN 1 ELSE 0 END

Proportion of Late Deliveries:

SUM([Late Delivery]) / COUNT([ID])

Market Value:

[Pack Price] * [Unit of Measure (per pack)]

Segmented the data by country, vendor, and dosage form to identify key trends and patterns.

Performed sorting and filtering to identify the top 10 vendors based on shipment volume.

## VISUALIZATION TECHNIQUES

Shipment Mode

Bubble Chart representing Top Shipment Modes.

COUNT ID vs Shipment Mode, Shipment Mode in Color Mark

Common Dosage Form

Bar Chart representing Most Common Dosage Form

COUNT ID vs Dosage Form, Dosage Form in Color Mark

Average Weight of Products

Bar Chart representing Average Treated Weight(Kilograms)

COUNT ID vs AVG Treated Weight

Top 10 Vendors

Bar Chart representing Top 10 Vendors

COUNT ID vs Vendor, Vendor Filtered to Top 10 values with regards to COUNT ID and sorted in descending order.

Popular Brands

Bubble Chart representing Popular Brands

COUNT ID vs Brand, Brand Filtered to Top 3 values with regards to COUNT ID.

## Late Delivery Percentage

Bar Chart representing Late Delivery Percentage

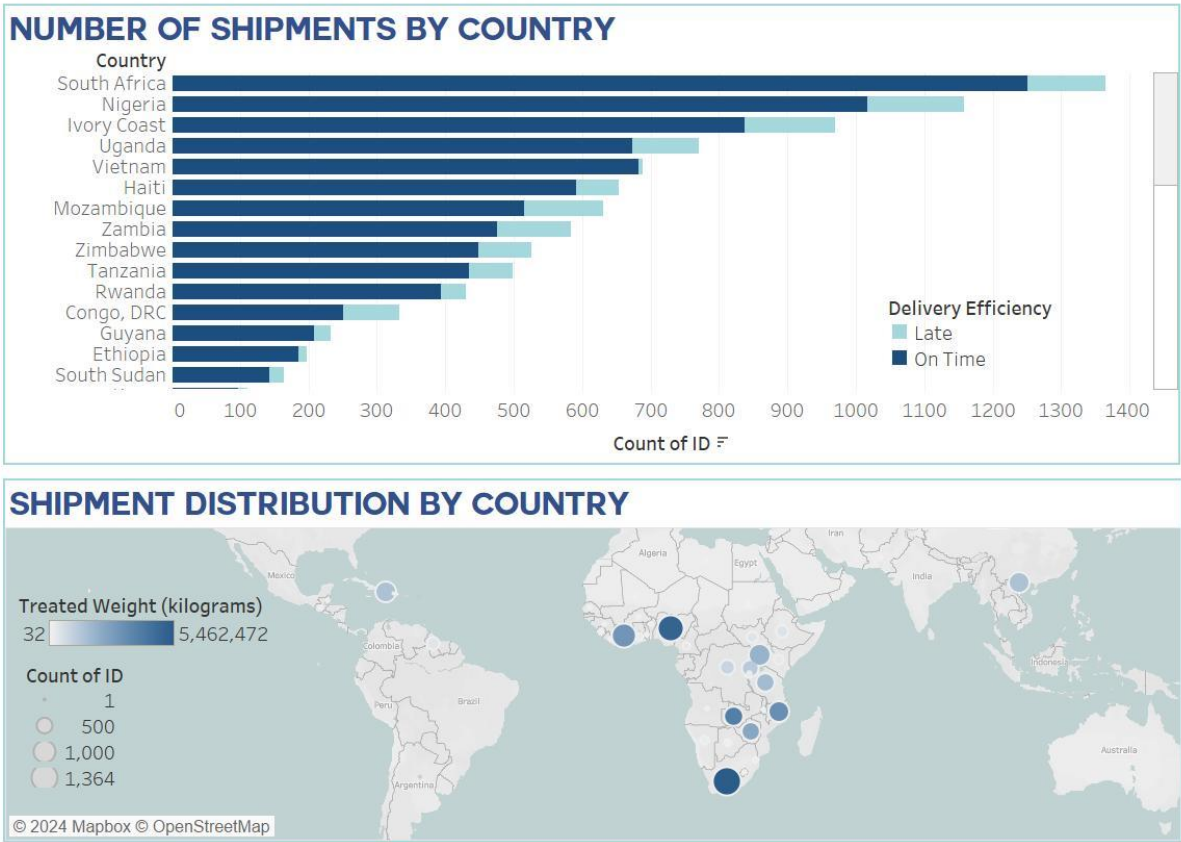% TOTAL COUNT ID vs Delivery Status

## Total Market Value

Bar Chart representing Total Market Value

SUM Market Value vs Country

# DASHBOARDS

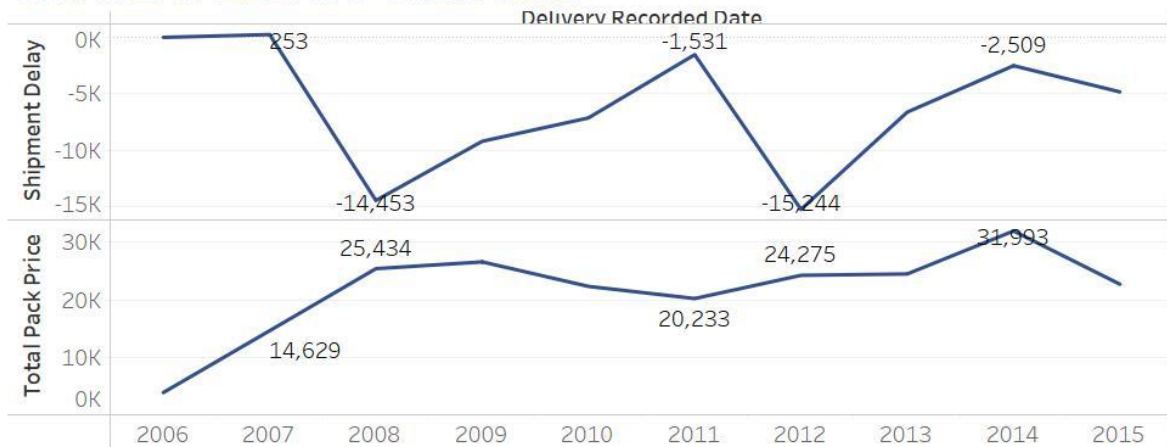## SHIPMENT PERFORMANCE AND DISTRIBUTION ANALYSIS

### NUMBER OF SHIPMENTS BY COUNTRY



### SHIPMENT DISTRIBUTION BY COUNTRY



- Chart 1: Number of Shipments by Country: This bar chart shows the total number of shipments per country and highlights delivery efficiency. It helps identify countries with the highest shipment volumes and areas where delivery performance can be improved.
- Chart 2: Shipment Distribution by Country: This map visualizes the distribution and weight of shipments across different countries. It provides insights into the geographical spread of shipments and identifies key regions with high shipment activity.
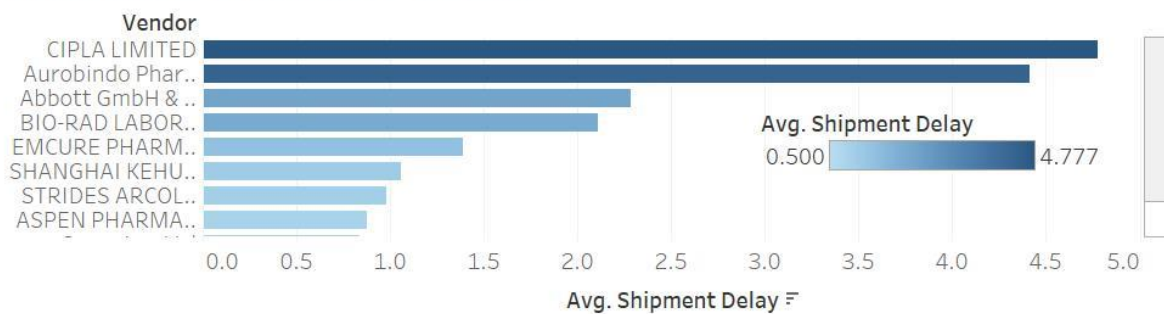
By putting these two visualisations together, the dashboard gives a clear picture of how shipments are doing and where they are being sent. This helps supply chain leaders make better strategic decisions that will improve operations and delivery times.

## SHIPMENT DELAYS: PRICE IMPACT & TIMELINESS

### SHIPMENT DELAYS OVER TIME

Delivery Recorded Date

Shipment Delay:
- 253
- -14,453
- -1,531
- -15,244
- -2,509

Total Pack Price:
- 14,629
- 25,434
- 24,275
- 20,233
- 31,993

Years: 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

### AVERAGE SHIPMENT DELAY

Vendor:
- CIPLA LIMITED
- Aurobindo Phar..
- Abbott GmbH & ..
- BIO-RAD LABOR..
- EMCURE PHARM..
- SHANGHAI KEHU..
- STRIDES ARCOL..
- ASPEN PHARMA..

Avg. Shipment Delay
0.500 — 4.777

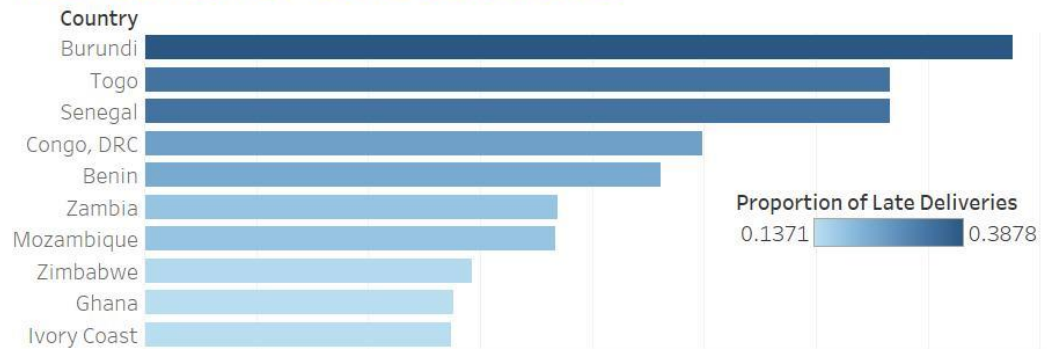Axis: 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0

Avg. Shipment Delay =

In this dashboard, I have analyzed two charts.

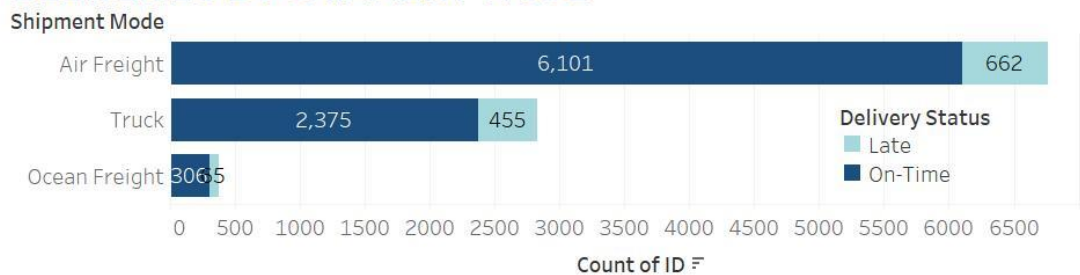- Chart 1: Shipment Delays Over Time: This line chart highlights the shipment delays across the timeline. This could be crucial for future forecasting and can be a factor to improve on the delivery mode, delays and shipments. This chart also talks about total pack price affected by late deliveries which can be crucial for future forecasting in losses and also acts as a lag indicator.
- Chart 2: Top 10 Vendors with Highest Shipment Delays: This chart focuses on the top 10 vendors having the most delays in shipments. This insight can help to reduce delays for top-performing vendors. This can lead to prioritizing improvement efforts on vendors with the highest delays.

# SHIPMENT MODE EFFICIENCY AND DELIVERY TIMELINESS

## PROPORTION OF LATE DELIVERIES

Country

| | |
|---|---|
| Burundi | |
| Togo | |
| Senegal | |
| Congo, DRC | |
| Benin | |
| Zambia | |
| Mozambique | |
| Zimbabwe | |
| Ghana | |
| Ivory Coast | |

Proportion of Late Deliveries

0.1371     0.3878

## DELIVERIES BY SHIPMENT MODE

Shipment Mode

| | | |
|---|---|---|
| Air Freight | 6,101 | 662 |
| Truck | 2,375 | 455 |
| Ocean Freight | 306 | 5 |

Delivery Status
- Late
- On-Time

0   500   1000   1500   2000   2500   3000   3500   4000   4500   5000   5500   6000   6500

Count of ID ₌

---

In this dashboard, I have analyzed two charts.

- **Chart 1:** Countries with Late Deliveries: This bar chart highlights which countries most frequently receive late deliveries, providing insight into potential geographic challenges.

This method ensures that the visualization focuses on the rate of late deliveries rather than just the raw count, providing a more accurate picture of which countries have the most significant issues with late deliveries.

- **Chart 2:** Shipment Mode and Late Deliveries: This stacked bar chart shows the distribution of late deliveries across different shipment modes, helping identify which methods contribute most to delivery delays.

This dashboard provides a clear view of delivery performance, allowing us to pinpoint problematic regions and shipment methods. By highlighting these areas, we can make more informed strategic decisions to enhance supply chain efficiency.

# REPORT

- **What are the target markets represented in the dataset?**

The dataset mainly represents African countries, along with a few others like Afghanistan, which are the target markets for our supply chain operations. These countries primarily receive ARV and HRDT drugs. The products are provided in various dosage forms, specifically designed for pediatric needs and HIV testing. A comprehensive analysis suggests the top markets being the African countries who recieves shipments and have a high market value for the shipment.
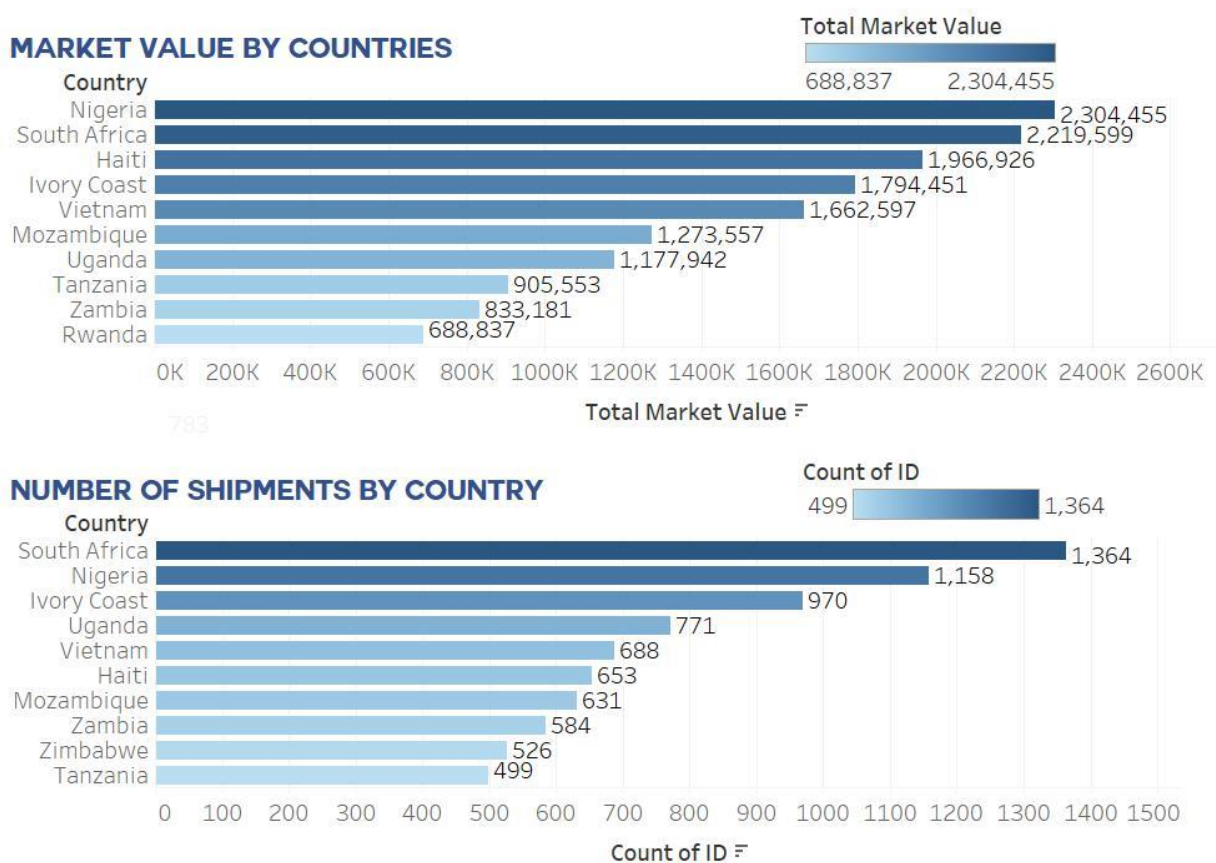


Fig.1. Top Markets

- **What is the time period represented in this dataset?**

The dataset encompasses a specific time period during which shipments were scheduled, delivered, and recorded. Upon examining the Scheduled Delivery Date, Delivered To Client Date, and Delivery Recorded Date fields, it was determined that the dataset spans from [Scheduled Delivery Date] to [Delivery Recorded Date]. This time frame provides a comprehensive view of our supply chain operations and the delivery performance over this period.

- **How is most product transported?**

The most commonly used shipment mode is Air Freight. Nearly 6,763 deliveries from the data analysed have used Air Freight as their shipment mode.



Fig.2. Mode of Transport

- **What is the most common "dosage form"?**

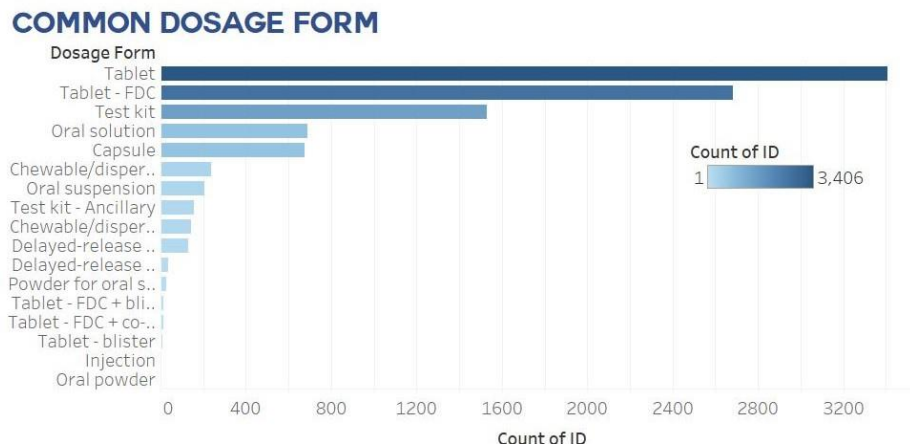The most common dosage form is Tablet.

Fig.3. Common Dosage Form

- **What is the average "weight" (in kilograms) of the products provided?**

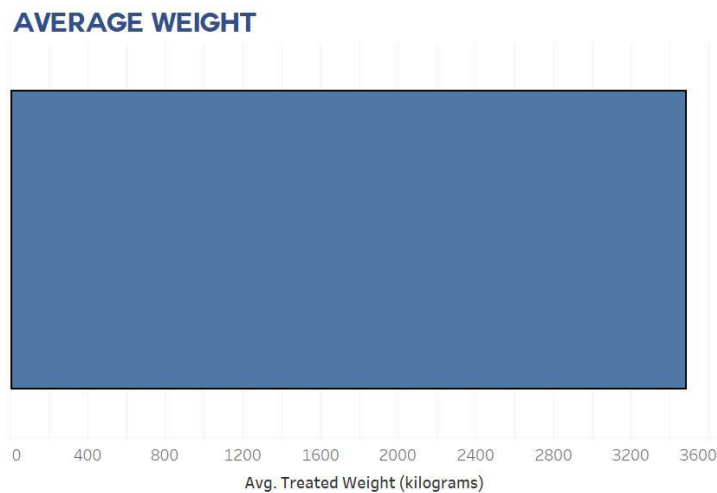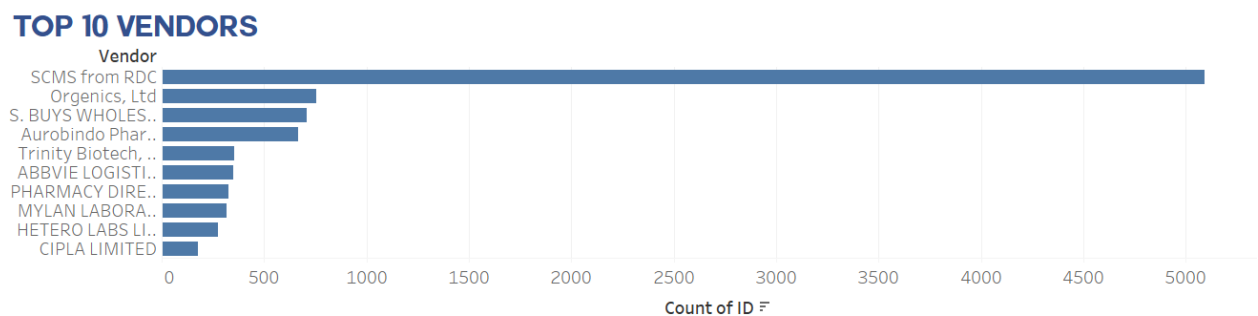Average weight of the products is 3479kg



**AVERAGE WEIGHT**

Fig.4. Average weight of the products provided

- **What are the top 10 most popular "vendors"?**

The following are the Top 10 Most popuar vendors from the data analyzed.



**TOP 10 VENDORS**

Count of ID for each Vendor. The view is filtered on Vendor, which keeps 10 of 72 members.

Fig.5. Popular Vendors

- **What are the top 3 most popular "brands"?**

The below are the top 3 brands popular from the data collected.

**TOP 3 BRANDS**

Brand
- Determine
- Generic
- Uni-Gold
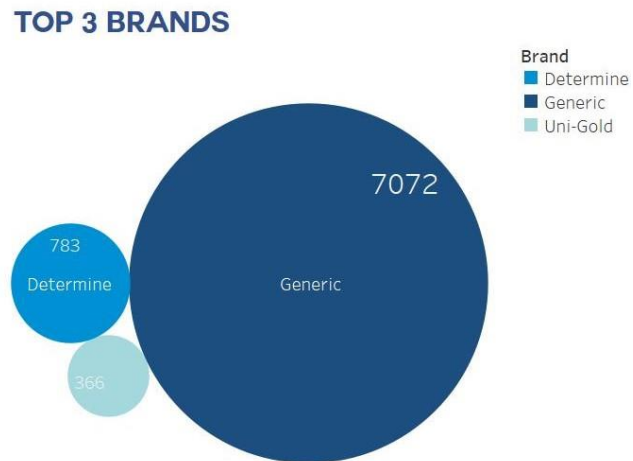
7072

783
Determine

Generic

366

Fig.6. Popular Brands

- **What percentage of the total products were delivered late to clients?**

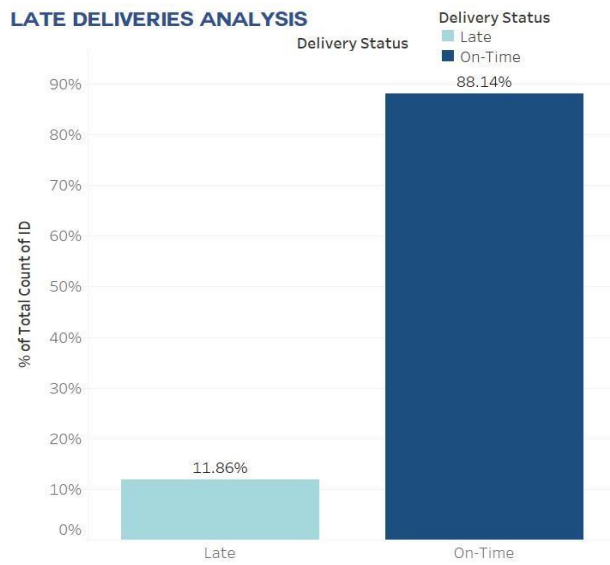11.86% of deliveries were delivered late to clients.

Fig.7. Late Deliveries Analysis

- **The relationship between "unit of measure (per pack)", "pack price" and "unit price".**

The "Unit of Measure (per pack)" indicates the number of individual units within a pack, the "Pack Price" represents the total cost of the pack, and the "Unit Price" is the cost of each individual unit. These metrics are interconnected, with the unit price being directly affected by both the pack price and the unit of measure. Understanding this relationship is crucial for analyzing and comparing product pricing strategies.

# RESULTS AND CONCLUSIONS

Our analysis of the ARV and HIV lab shipment data yielded valuable insights to optimize our supply chain. Air freight emerged as the most efficient mode of transport, demonstrably reducing delivery delays compared to other methods. This finding suggests a potential shift in strategy towards prioritizing air freight for time-sensitive shipments.

Furthermore, we identified the top 10 vendors based on shipment volume. However, Cipla Limited exhibited the highest delay rate, hindering overall delivery efficiency. Addressing these delays through targeted interventions, such as renegotiating contracts or exploring alternative partnerships, can significantly improve on-time deliveries.

The comprehensive dashboard provides actionable insights for strategic decision-making. By focusing on improving performance with top vendors like Cipla Limited and prioritizing timely deliveries in regions experiencing high delays, we can achieve several key benefits. This data-driven approach can enhance customer satisfaction with improved access to critical medications, potentially reduce costs associated with delayed deliveries, and strengthen our brand reputation as a reliable supplier.

As of 2022, South Africa was the country with the highest number of people living with HIV in Africa. At that time, around 7.6 million people in South Africa were HIV positive. In Mozambique, the country with the second highest number of HIV-positive people in Africa, around 2.4 million people were living with HIV. (Statista.com)

This is concluded with our analysis as South Africa being the country with highest numbers of shipments received.

Looking ahead, further analysis incorporating data on inventory levels and demand forecasting can be instrumental. This would allow us to optimize stock management, proactively address potential shortages, and ensure we meet regional needs more effectively. Ultimately, this data-driven approach will empower us to build a more resilient and efficient supply chain for delivering life-saving ARV and HIV lab test kits.

# References

World Health Organization. (2023, October 6). HIV/AIDS [RETRACTED June 7, 2024]. https://staging.afro.who.int/health-topics/hivaids

Statista.com. (2023, November 19). Number of people with HIV in African countries 2021 [RETRACTED June 7, 2024]. https://www.statista.com/statistics/1305217/number-people-with-hiv-african-countries/