INFO 430 DATABASE DESIGN AND MANAGEMENT

Class Demo Project: Sustainable Goals Financing and Achievement Scores

Arya Karki & Muhamed Hashi

May 19, 2025

Abstract

This project investigates global patterns of international aid distribution in 2023 by analyzing data from the International Aid Transparency Initiative (IATI). The primary goal is to assess whether aid disbursements align with development needs across countries and sectors. To support this analysis, we designed and implemented a dimensional data warehouse, using a star schema to organize aid transaction data and related contextual dimensions such as organization, sector, geography, and finance type.

To enrich our analysis, we integrated Human Development Index (HDI) data from the United Nations Development Programme, allowing us to assess whether aid flows are responsive to levels of human development. Using PySpark for data cleaning and transformation, and SQL for analysis, we generated four complex analytical queries. These included summary cubes, ranking and value-based window functions, and time-series moving averages to uncover trends and outliers.

Through multidimensional analysis and interactive visualizations, this project reveals key insights into disbursement trends, geographic distribution, financial mechanisms, and sectoral imbalances in global aid flows.

I. Introduction

International aid plays a critical role in addressing global disparities and supporting development across health, infrastructure, education, and humanitarian sectors. However, understanding how this aid is distributed, to whom, and whether it's aligned with developmental needs is crucial for accountability and planning. This project investigates global aid disbursement patterns in 2023 using data from the International Aid Transparency Initiative (IATI), enriched with Human Development Index (HDI) scores. Our objective is to provide actionable insights into where funds are flowing and whether they align with indicators of need and capacity.

II. Data Sources, Dimensional Model, and Data Architecture

Main Data Source

The main dataset for this project was sourced from the IATI database, capturing aid-related project information, with preselected columns by Professor Otim, including:

Column Name	Description
Activity IATI Identifier	A globally unique identifier assigned to each aid activity to ensure traceability and consistency.
Activity Title	A brief, descriptive title summarizing the aid project or initiative.
Activity Description	An explanation outlining the goals, scope, and context of the activity.
Activity Start Date	The date on which the aid activity officially began.
Activity End Date	The scheduled or actual end date of the aid activity.
Reporting Organization	The entity submitting the project data to IATI (e.g., donor agency, NGO, multilateral).
Reporting Organization Type	Classification of the reporting organization (e.g., bilateral, multilateral, NGO).

Aid Type	The form of aid provided, such as grants, loans, or technical assistance.
Finance Type	Describes the financial arrangement (e.g., loan, equity, reimbursable grant).
Flow Type	Specifies the nature of the funding flow (e.g., ODA, other official flows, private).
Provider Organization	The organization funding or implementing the project.
Provider Organization Type	The classification of the funding organization (e.g., government, NGO, foundation).
Receiver Organization	The organization receiving and managing the aid on the ground.
Receiver Organization Type	The type of the receiving body (e.g., local government, civil society, international org).
Recipient Country or Region	The country or region that is the target of the aid activity.
Sector Category	The broad category of the sector the project belongs to (e.g., infrastructure, health).
Sector	A more specific classification within the sector category (e.g., education, water supply).
Humanitarian/Development	Indicates whether the project is humanitarian (emergency relief) or development-focused.
Transaction Type	The kind of financial event recorded (e.g., commitment, disbursement, expenditure).
Value (US $)	The amount of money involved in the transaction, expressed in U.S. dollars.

Data Enrichment

To enrich our analysis of international aid distribution, we incorporated Human Development Index (HDI) data from the United Nations Development Programme (UNDP) for the year 2023. The HDI is a composite indicator measuring a country's overall achievement in key dimensions of human development: health (life expectancy), education (mean and expected years of

schooling), and standard of living (GNI per capita). By integrating HDI scores with our aid disbursement data, we were able to examine whether funding levels correlate with indicators of developmental need. This addition allows us to assess equity in resource allocation and identify potential mismatches, such as low-HDI countries receiving disproportionately low levels of aid, providing critical insight for donors, policymakers, and development agencies. (https://hdr.undp.org/data-center/human-development-index#/indicies/HDI)

Here are all the columns part of the HDI dataset. I have bolded the ones we used for our data enrichment.

Column Name	Description
HDI rank	**The rank of the country based on its Human Development Index (HDI) score for 2023.**
Country	**The name of the country, categorized by level of human development (e.g., Very High, High, etc.).**
Value	**The Human Development Index (HDI) score for 2023, ranging from 0 (lowest) to 1 (highest).**
Life expectancy at birth (years)	Average number of years a newborn is expected to live if current mortality rates continue.
Expected years of schooling (years)	The number of years a child entering school is expected to study, including primary to tertiary education.
Mean years of schooling (years)	Average number of years of education received by people aged 25 and older.
Gross national income (GNI) per capita (2021 PPP $)	Average income per person, adjusted for purchasing power parity.
GNI per capita rank minus HDI rank	Indicates whether income rank is higher or lower than overall HDI rank; positive values suggest overperformance on income relative to HDI.
HDI rank (2022)	The country's HDI rank from the previous year for comparison.

Dimensional Model

From the above two sources of data, we created the following dimensional model. We chose to use the star schema to represent the dimensional model because it provides a flexible, high-performing, and user-friendly approach to analyzing complex aid transaction data. The dimensional model is depicted in the figure below.

Figure 1: Dimensional model for International Financial Aid Data Warehouse

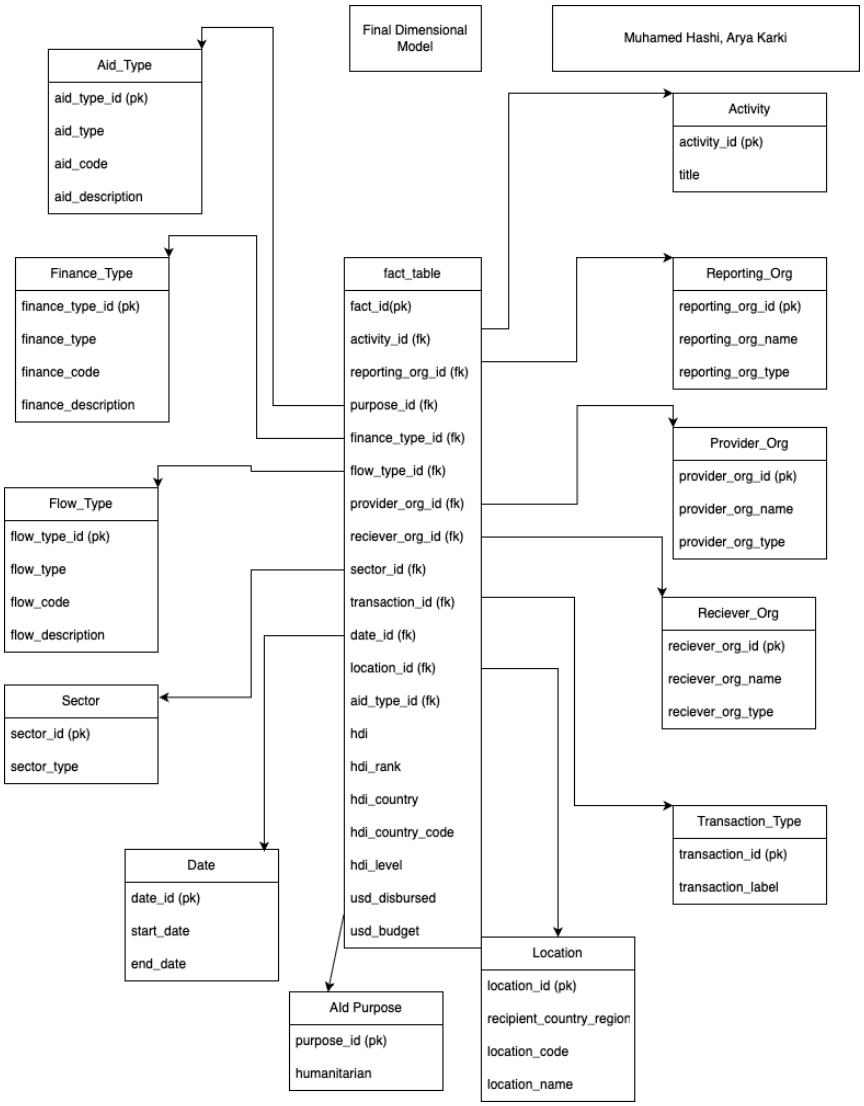
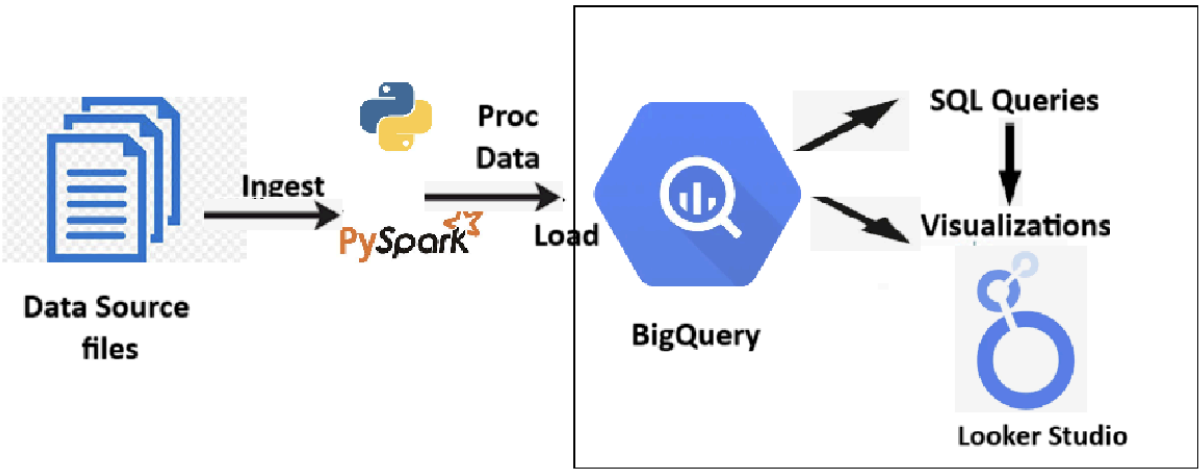


Figure 2: Conceptual Data Architecture



III. Data Processing

Our data processing pipeline included the following key steps:

- Renaming and aligning columns across three input files.
- Merging all parts into a unified fact table using the unique IATI Identifier.
- Computing disbursement values across partial columns.
- Renaming key fields for clarity and ease of integration with dimension tables.
- Broadcast joins between the core fact table and dimensions for activity, sector, date, and organizations.
- Enriched the dataset with Human Development Index (HDI) scores from UNDP, trimming country names to ensure accurate joins.
- Final selection of analytical fields, including all surrogate IDs and financial values.
- Exported the enriched fact table to Google BigQuery, enabling efficient cloud-based SQL querying and dashboard integration.

The pipeline allowed for the creation of a clean, enriched data warehouse ready for analytical querying.

IV. Queries, Visualizations, and Telling a Data Story

CUBE queries

1. This query evaluates whether international aid disbursements are aligned with development needs by incorporating Human Development Index (HDI) scores. By

enriching a cube-style summary of aid flows with HDI values, this query enables visual correlation of aid allocation with human development levels.

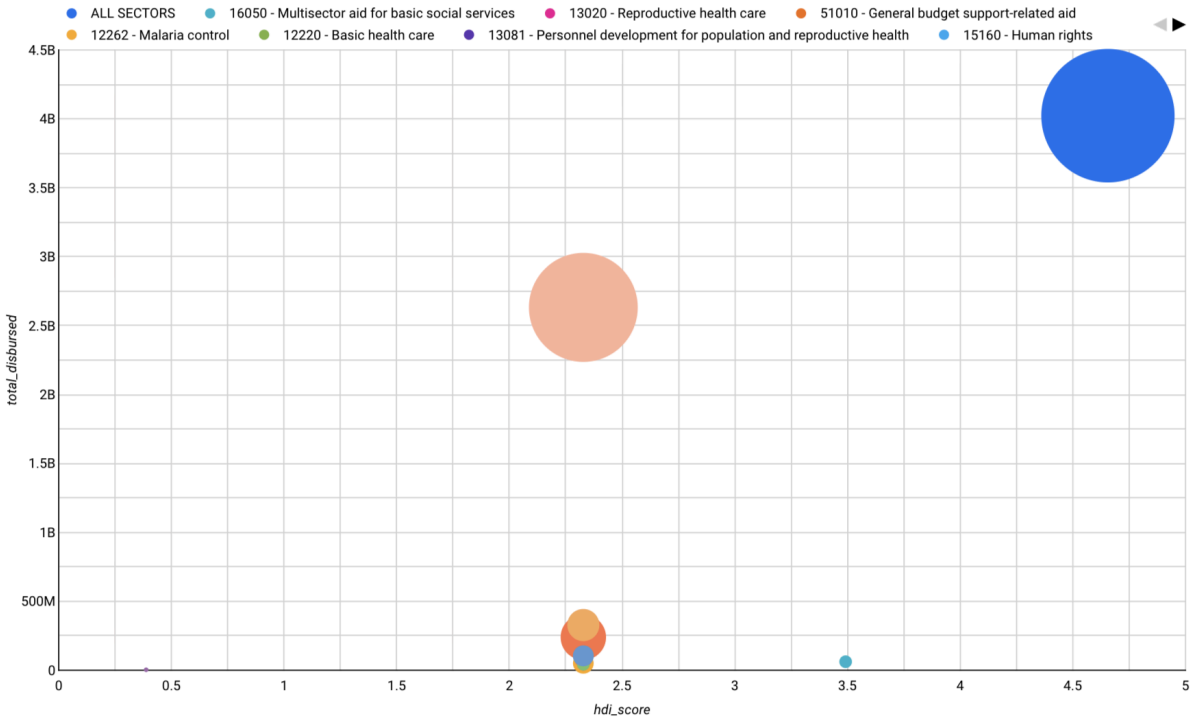


Figure 3: Aid Disbursement by HDI Score and Sector
This bubble chart visualizes the relationship between a country's HDI score and the total aid disbursed to sectors within that country. Each bubble represents a sector, with the X-axis showing the HDI score, the Y-axis showing total disbursements, and the bubble size representing the total budgeted amount. The chart highlights potential mismatches between developmental needs and funding levels.

Narrative & Insight

- The chart reveals that some lower-HDI sectors receive relatively low disbursement, while certain higher-HDI regions/sectors receive disproportionately larger funds. For example, general budget support-related aid (sector 51010) is heavily funded even at higher HDI scores, while reproductive and health-focused sectors with lower HDI scores appear underfunded in comparison. This suggests potential misalignment between developmental need and resource allocation.

Managerial Implications

- Aid Alignment: Visual discrepancies between HDI and disbursement signal the need to reassess whether funding aligns with human development needs.
- Sector Rebalancing: Overfunded high-HDI sectors may reduce the impact of aid; underfunded low-HDI areas may need targeted investment.

2. This query analyzes how different finance types (e.g., grants, loans) are distributed across flow types (e.g., ODA, private finance) by aggregating both total disbursements and transaction volume using a CUBE function.

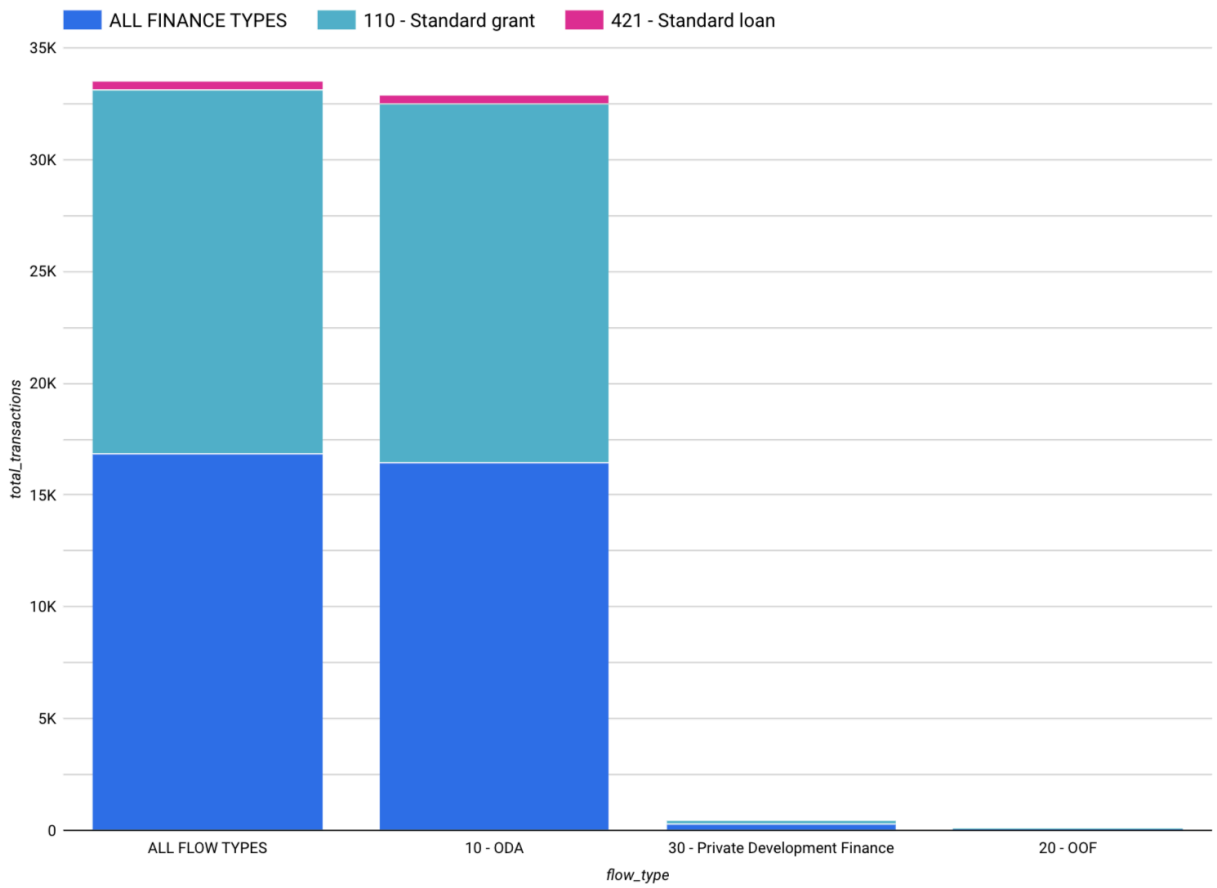


Figure 4: Total aid transaction counts by flow type, segmented by finance type.
The stacked bar chart displays the volume of aid transactions by flow type, segmented by finance type. Each bar shows the total transaction count for a flow type (e.g., ODA, private development finance), with color-coded segments distinguishing finance mechanisms like standard grants and standard loans.

Narrative & Insight

- The chart reveals that the vast majority of transactions occur through Official Development Assistance (ODA), predominantly using standard grants (type 110). Standard loans (type 421) appear far less frequently and are concentrated in a small subset of flow types. Private development finance and OOF represent only a tiny fraction of transaction activity, suggesting limited engagement from alternative finance channels.

Managerial Implications

- Strategic Design: The dominance of standard grants within ODA suggests a preference for low-risk, non-repayable aid, which is useful for emergency relief or fragile-state development.
- Opportunity Areas: The minimal use of private finance and OOF highlights underutilized financing models that could be explored to expand development impact.

Ranking Window Queries

1. This query uses a ranking window function to identify the top 3 recipient organizations by total disbursements in each country, helping surface which partners are receiving the most aid.

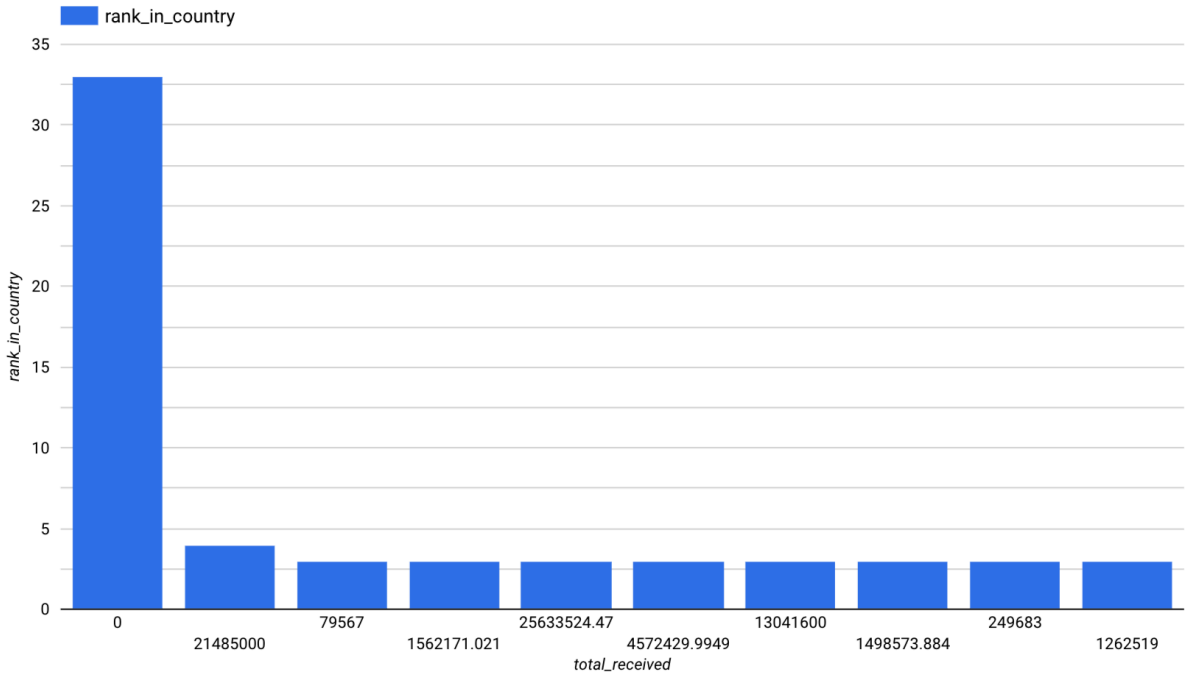


Figure 5: Ranked Recipient Organizations by Disbursements across Countries
The bar chart visualizes ranked recipient organizations based on their total disbursements, with rank_in_country on the y-axis and total_received on the x-axis.

Narrative & Insight

- One organization in a specific country stands out with significantly higher disbursements, while most others received comparatively modest funding. This indicates a centralized flow of aid toward key partners in some regions.

Managerial Implications

- Aid agencies can use this insight to audit concentration of funds, evaluate recipient performance, and consider whether diversifying aid distribution might improve local impact or reduce dependency.

2. This query identifies the top 3 aid types (e.g., project-type interventions, budget support) by total disbursement within each finance type (e.g., grants, loans). It uses a window function (RANK) to highlight the most financially significant aid categories under different funding models.

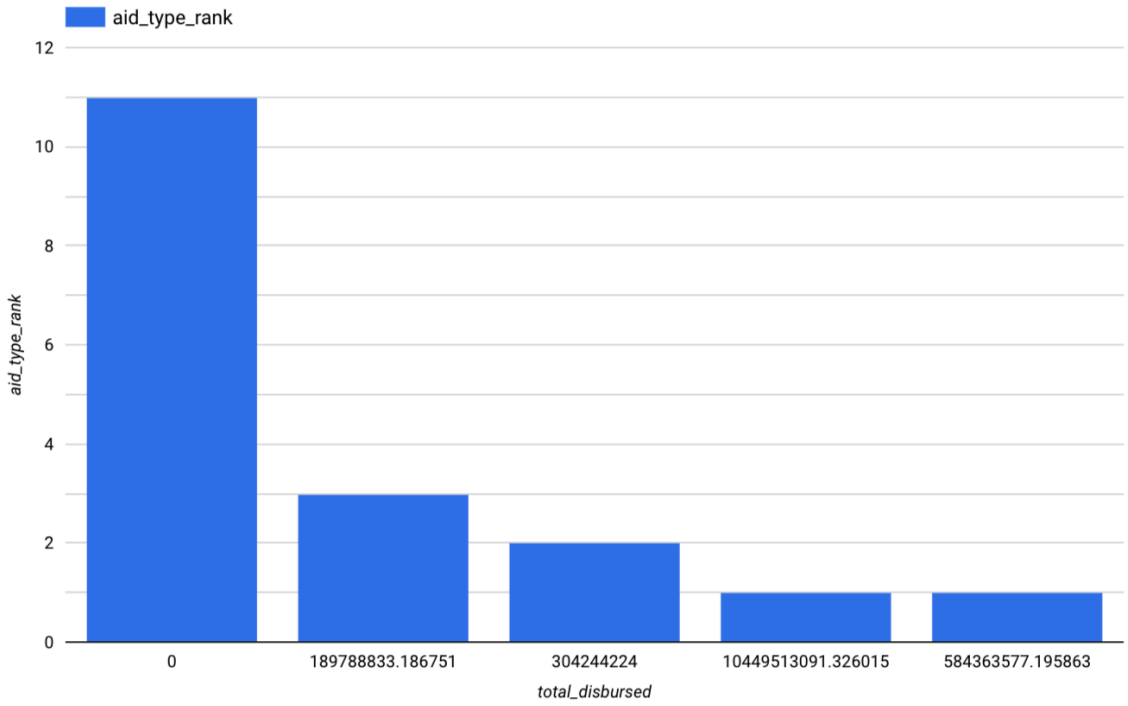


Figure 6: Ranked aid types by total disbursement within each finance type
The bar chart shows the ranking of aid types by total disbursed funds, with each bar representing the disbursement total and its corresponding rank. The chart is filtered to include only the top three aid types per finance type, simplifying comparison of which aid types dominate each financial channel.

Narrative & Insight

- The visualization reveals a steep drop-off in disbursements between the top-ranked aid type and subsequent ones, suggesting a heavy concentration of funds in one or two dominant aid strategies per finance type. This highlights a clear prioritization within funding categories, with one aid type often receiving the majority share of disbursements.

Managerial Implications

- Funding Diversification: The ranking exposes imbalances in aid distribution, suggesting opportunities to diversify funding across underutilized aid types.

- Resource Allocation: Policymakers can use this insight to refine allocation strategies, ensuring that dominant aid types are both effective and equitably supported.

Value Window Function Queries

1. This query analyzes month-to-month changes in aid disbursement by country using a value window function (LAG). It highlights trends over time, calculates disbursement percent changes, and smooths irregular fluctuations using a 3-month rolling average.

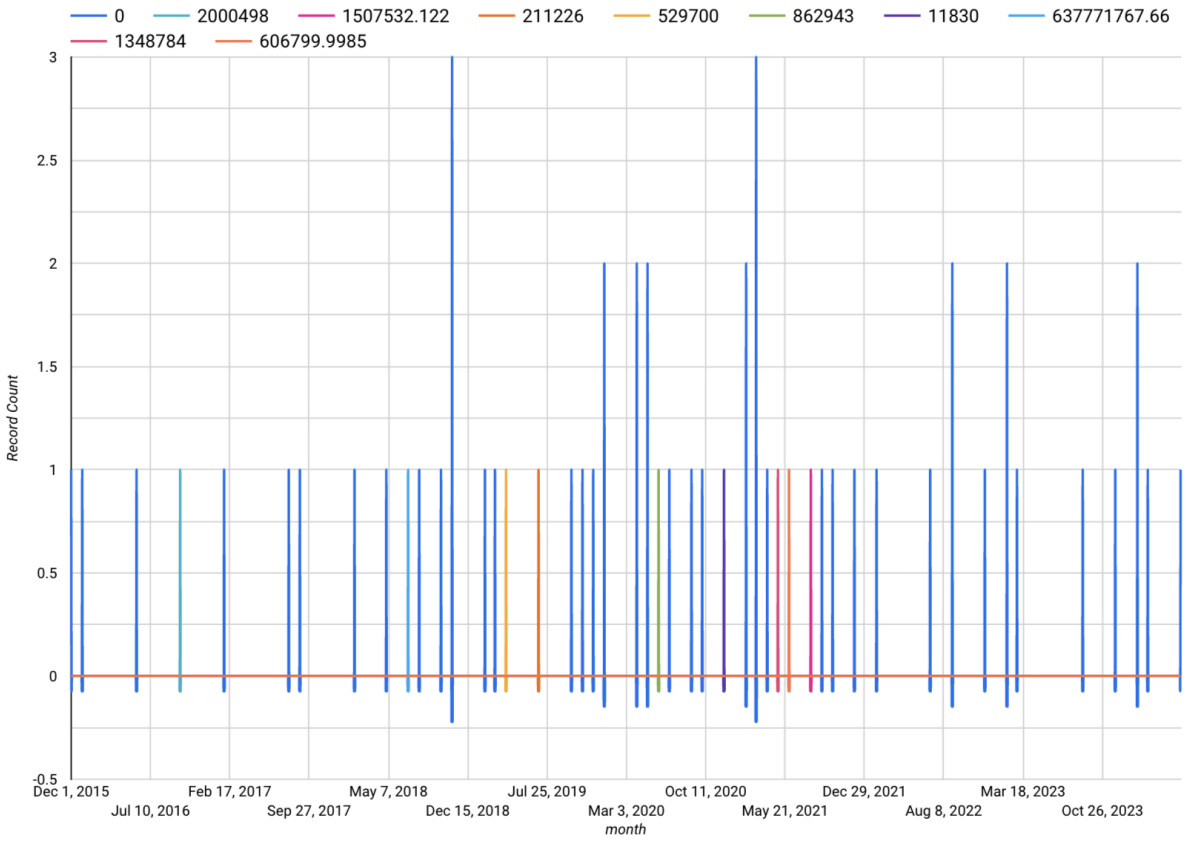


Figure 7: Monthly disbursement trends for the top 5 recipient countries
The line chart plots disbursement activity by month for the top 5 recipient countries from 2015 onward. Each line represents one country's aid flow, showing when disbursements increased, decreased, or plateaued.

Narrative & Insight

- The visualization reveals that aid distribution is irregular, with several months of inactivity or minimal disbursement, followed by sharp spikes. This suggests funding is often delivered in large, irregular tranches, rather than as consistent monthly flows. Some countries received far larger disbursements than others, as reflected in line separations.

Managerial Implications

- Timing matters: Policymakers should ensure that aid is not just delivered, but timed effectively for local impact.
- Predictability & Planning: Volatile month-to-month disbursements can hinder program continuity and local planning. Rolling average metrics can help stabilize funding expectations.

2. This query analyzes year-over-year disbursement trends by country, using lag and lead window functions to classify each year as having increasing, decreasing, stable, or no clear trend in aid disbursement. It helps reveal overall funding patterns and shifts in financial priorities over time.

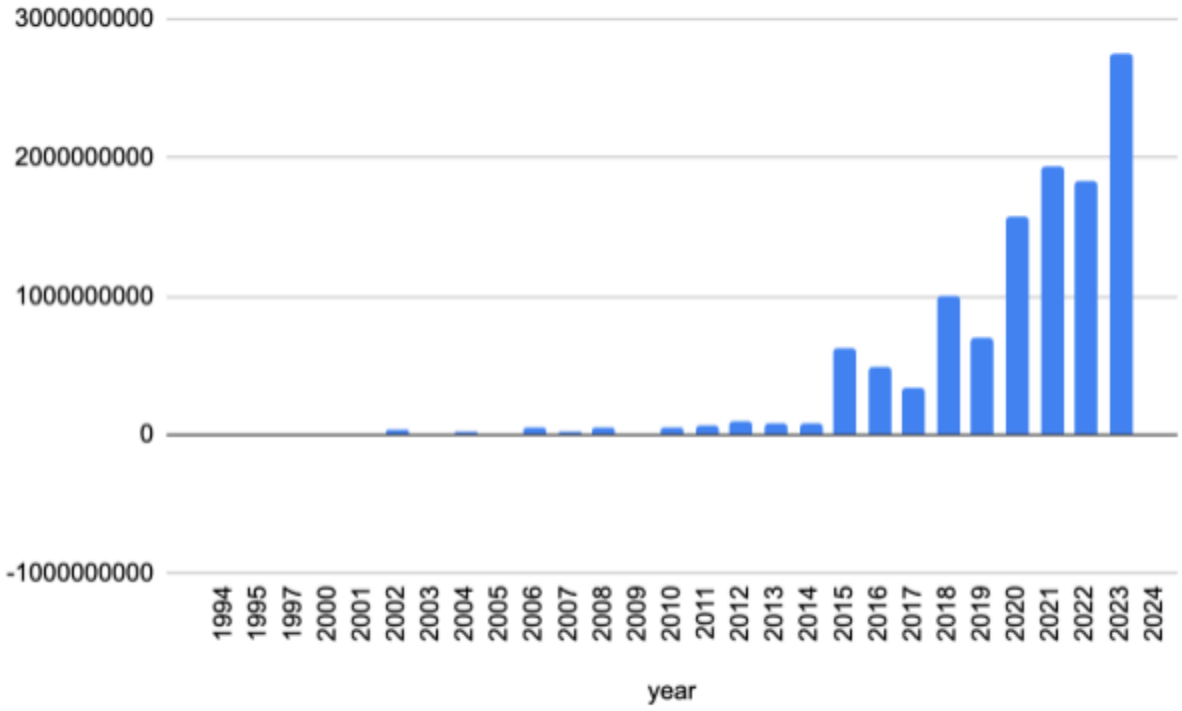


Figure 8: Total Global Disbursement by Year
The bar chart displays total disbursements by year, revealing a clear upward trajectory in aid volume—especially from 2015 onward. The data spans from the early 1990s through 2024, with particularly sharp increases in the most recent years.

Narrative & Insight

- The chart reveals a dramatic growth in global aid disbursements, particularly post-2015. Funding spikes in 2020–2022 suggest a strong response to global crises (COVID-19), with continued investment in 2023 and 2024. Earlier years show lower, more volatile disbursement levels, but recent trends reflect scaling and institutionalization of aid flows.

Managerial Implications

- Forecasting & Budgeting: A strong upward trend enables more confident forecasting and capacity planning for recipient countries.
- Program Sustainability: Organizations can capitalize on consistent growth by developing long-term strategies, rather than short-term project-based responses.
- Policy Timing: Spikes in funding during crises highlight the need for flexible policy frameworks that adapt to surge funding without bottlenecks.

Time Series / Moving Average Queries

1. This query calculates a 3-month moving average of disbursements for each country, helping to smooth short-term spikes or dips in funding. This makes it easier to evaluate long-term trends in aid distribution and support effective planning.

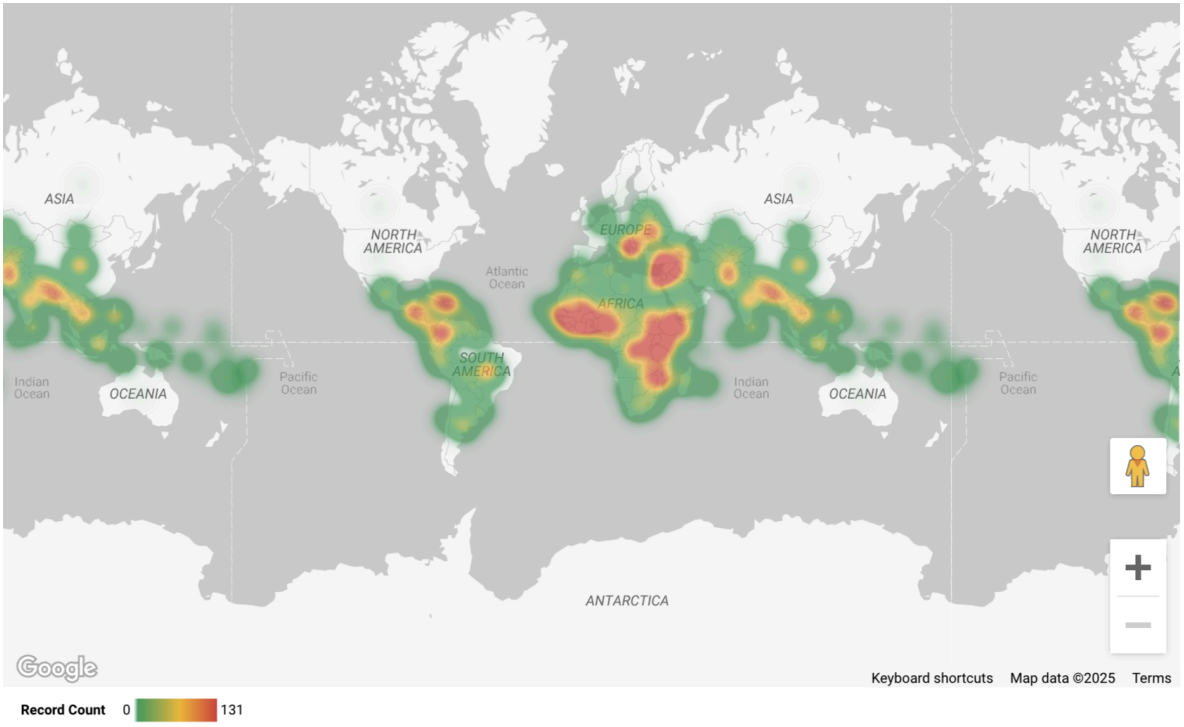


Figure 9: Global heatmap of aid disbursement activity
The heatmap highlights the geographic concentration of aid disbursement activity using the record count of transactions across global coordinates. Areas in red represent regions with higher frequency of aid activity, while green and yellow indicate lower density.

Narrative & Insight

- The data reveals that aid is most frequently dispersed in parts of Africa, South Asia, and parts of Latin America, with noticeable hot spots in central and eastern Africa. These regions likely represent countries facing the most urgent development needs or ongoing humanitarian crises. The widespread green patterns show that while aid is globally distributed, a few core regions dominate transaction volumes.

Managerial Implications

- Strategic Focus: Identifying geographic hot spots allows funders to confirm whether resource allocation aligns with priority regions.
- Equity in Aid Distribution: The visualization helps detect potential neglected regions where aid activity may be disproportionately low.

2. This query measures the three-period rolling standard deviation of aid budgets for each country as an indicator of volatility in aid allocation over time. Countries with higher rolling standard deviations experience greater short-term budget swings, while lower values indicate more stable and predictable funding patterns.

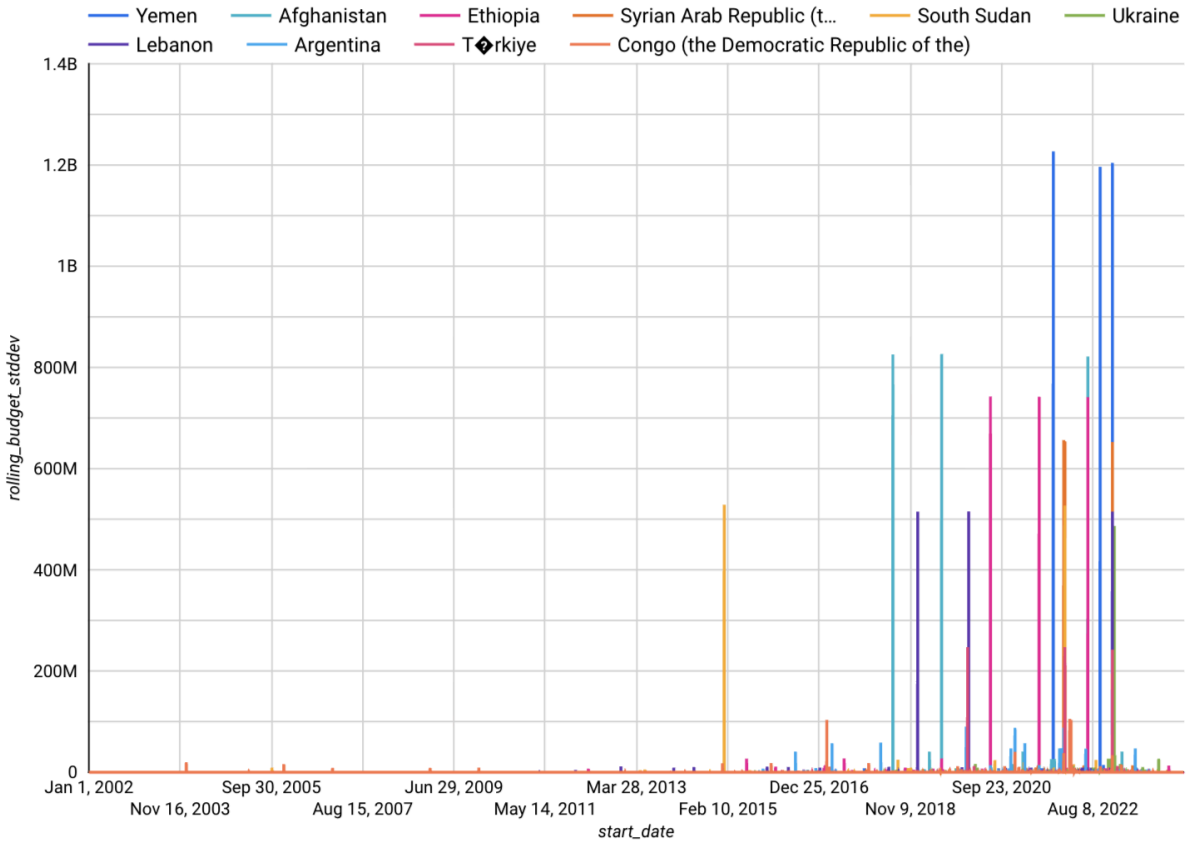


Figure 10: Rolling 3-period standard deviation of aid budgets per country
The multi-line chart shows the rolling volatility of aid budgets for selected countries over time. Each line represents one country, and the Y-axis reflects the rolling standard deviation of monthly aid budgets. Peaks in the chart indicate periods of high funding variability, while flat or low areas suggest stable support.

Narrative & Insight

- The chart reveals that countries such as Yemen, Afghanistan, and Ukraine have experienced periods of extreme budget volatility, particularly in recent years. This likely reflects surge funding in response to crises, such as conflict or natural disasters. In contrast, some countries maintain relatively low variance, suggesting a more consistent funding structure over time.

Managerial Implications

- Risk Planning: Countries with high volatility may face disruption in program delivery, requiring contingency planning and flexible budget execution strategies.
- Aid Effectiveness: Stable budgets support long-term development projects, which rely on predictable funding to build infrastructure, institutions, and capacity.

V. Conclusion

This project demonstrates how integrating transactional aid data with human development indicators can uncover both alignment and misalignment in global funding strategies. The use of dimensional modeling and advanced SQL (including CUBE, window, and time-series functions) enabled multidimensional insights that can support policy planning, transparency efforts, and performance evaluation. Future extensions could include integration of project impact metrics, qualitative outcome data, or additional socioeconomic indicators to further contextualize aid effectiveness.

References

United Nations Development Programme. (n.d.). Human Development Index (HDI). UNDP Human Development Reports. Retrieved May 16, 2025, from https://hdr.undp.org/data-center/human-development-index#/indicies/HDI