INFO 430 DATABASE DESIGN AND MANAGEMENT

Class Demo Project: International Financial Aid

By Kirby Vo, Sree Pratti

February 13, 2026

Abstract

This project builds a data warehouse to analyze international aid transactions across countries, organizations, and development sectors. The dataset shows who is providing funding, which countries receive it, what sectors the money supports, and how funding changes over time, and how these flows relate to country context. Using SQL analytical queries, we explored patterns in how aid is distributed, identified the largest donors and recipient countries, measured year-to-year changes in funding, and looked at longer-term trends using moving averages.

The results show that aid is not distributed evenly across countries or sectors, and funding levels often fluctuate from year to year. Some types of organizations regularly fund specific sectors, while certain countries receive a much larger share of assistance than others. By turning raw transaction data into summaries and visualizations, this project shows how database analysis can help better understand global funding behavior and support planning decisions.

I.   Introduction

International aid plays an important role in development around the world. Governments, international organizations, and other donors provide funding to support areas like infrastructure, healthcare, education, and economic growth. However, it can be difficult to understand how this aid is actually distributed because the data exists as large, complex transaction records.

In this project, we built a data warehouse to analyze international aid transactions. The system brings together information about donors, recipient countries, sectors, and time periods into a structured model so the data can be queried more easily. Organizing the data this way makes it possible to run large-scale analytical queries efficiently.

The goal of the project is to explore funding patterns and answer questions such as which donors fund which sectors, which countries receive the most aid, how funding changes over time, and whether long-term trends exist. Using SQL analytical functions like aggregation, ranking, and time-series analysis, we turn raw aid records into insights that are easier to understand.

II.   Data Sources, Dimensional Model, and Data Architecture

IATI Aid Transaction Data

The primary dataset comes from the IATI Country Development Financial Data portal (https://countrydata.iatistandard.org/), which contains detailed transaction-level aid data for recipient countries. The dataset includes fields such as country, sector type, flow type, finance type, transaction year/quarter, and value in USD. This dataset captures who provides funding to whom and can be categorized by sector type and more.

Column	Description
IATI Identifier	Identifier for the activity in the IATI system
Title	Name or title of the aid activity
Reporting Organisation Group	Broad category of the reporting organisation (e.g., Multilateral, NGO, Government)
Reporting Organisation	Name of the organisation reporting the activity
Reporting Organisation Type	Type/category of the reporting organisation
Aid Type	Form of aid (scholarships, debt relief, etc.)
Finance Type	Financing mechanism (standard loan, standard grant, etc.)
Flow Type	Classification of the financial flow (ODA, private market, etc.)
Provider Organisation	Organisation providing the funds for the transaction
Provider Organisation Type	Type/category of the provider organisation
Receiver Organisation	Organisation receiving the funds
Receiver Organisation Type	Type/category of the receiving organisation
Transaction Type	Type of financial transaction (commitment, disbursement, expenditure, etc.)
Recipient Country or Region	Country or region receiving the funds
Multi Country	Indicator (0/1) for whether the activity spans multiple countries
Sector Category	Broad sector classification (e.g., Health, Education)
Sector	Specific sector code/name
Humanitarian	Indicator (0/1) for humanitarian activities
Calendar Year	Year of the transaction
Calendar Quarter	Quarter of the transaction (Q1–Q4)
Value (USD)	Transaction value in USD

WORLD DEVELOPMENT INDICATORS (WDI)

To provide some country context and measure recipient needs, we had the IATI dataset enriched using World Bank World Development Indicators. The dataset covers population count, GDP per capita, primary school completion rate, mortality, and more, with these metrics being contained under a "Series code column". These country-level metrics enable analysis of aid flows relative to population size, and development indicators.

Column	Description
Country Name	Name of the country
Country Code	ISO country code
Series Name	Name of the indicator (e.g., GDP per capita, life expectancy)
Series Code	Indicator code (e.g., NY.GDP.PCAP.CD)
1970 [YR1970]	Indicator value for year 1970
1971 [YR1971]	Indicator value for year 1971
…. [YR ….]	Indicator value for year ….
2023 [YR2023]	Indicator value for year 2023
2024 [YR2024]	Indicator value for year 2024

Data Integration

By combining IATI transaction data with WDI country context, we can not only analyze raw aid patterns but aid effectiveness relative to recipient country needs and development profiles.

Dimensional Model

The data warehouse follows a star schema design centered around the **`FactAidTransaction`** table. This fact table stores individual aid transactions, including the amount of funding and references to the related descriptive attributes.

Surrounding the fact table are several dimension tables that provide context for each transaction. The `DimCountry` table identifies the recipient country, while `DimTime` records the year and quarter of the transaction. The `DimSector` table describes the

development sector the funding supports, and `DimOrganization` identifies the donor organization providing the aid.

Additional dimensions further classify the funding details. `DimAidType`, `DimFinanceType`, `DimFlowType`, and `DimTransactionType` describe the type of assistance, financing method, flow classification, and transaction category. These tables allow transactions to be analyzed from multiple perspectives without duplicating data.

The warehouse also includes a `FactCountryContext` table containing country-level indicators such as population, GDP per capita, life expectancy, and education completion rates. This table shares the country and time dimensions, enabling analysis of how funding relates to country conditions.

Overall, the star schema organizes transactional aid data and country indicators into a structure optimized for analytical queries, making it possible to efficiently study funding patterns across countries, sectors, and time.
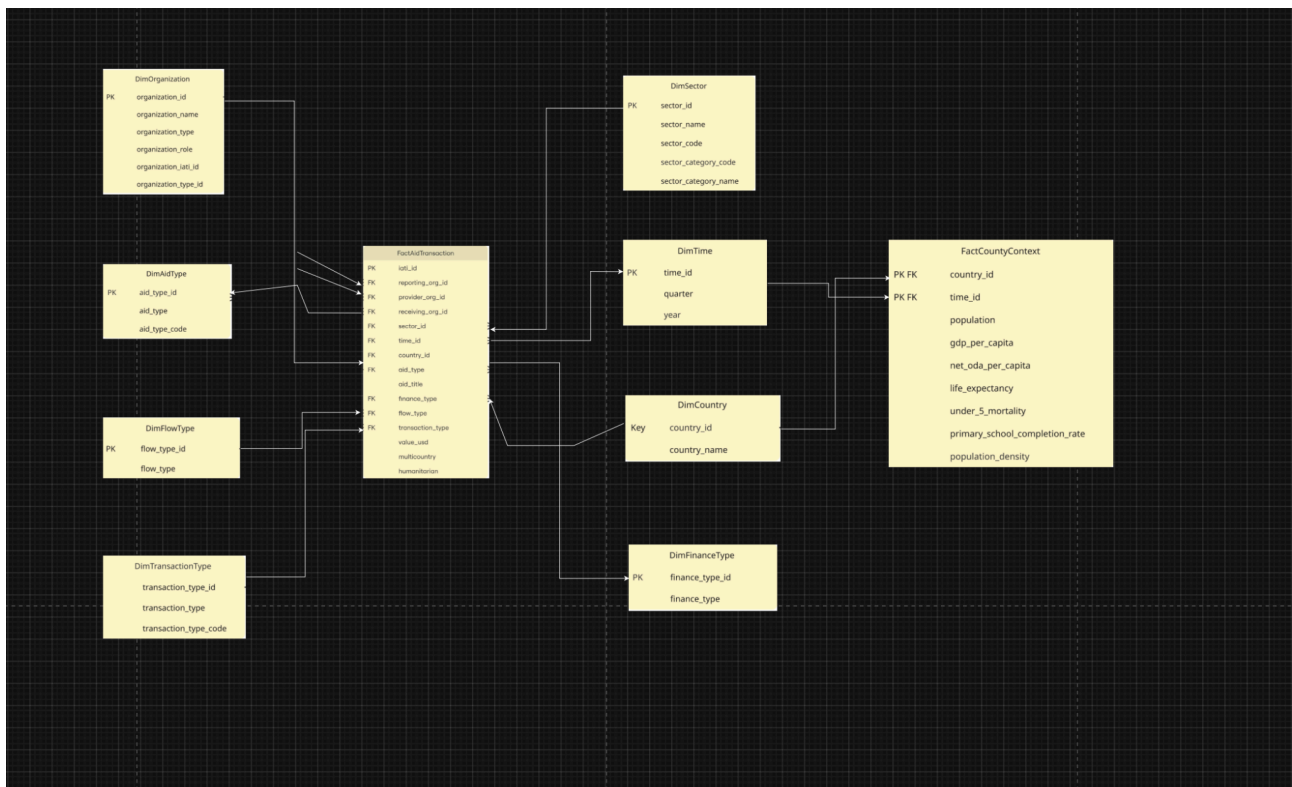

Figure 1: Dimensional Model for International Aid and Country Context Data Warehoes

Conceptual Data Architecture

The conceptual data architecture is presented in figure 3. Our architecture depicts the key technologies that will be used for ETL (extract, transform and load) steps. Specifically, we will use R for converting .xlsx files to csv, transform our data using Python and save this processed data as .csv files as well. After this, we will upload our data to our SQL server for data warehouse implementation and querying. While we could've loaded the data directly from our ETL script using sqlachemy, it was not within our time scope to completely automate the ETL process. In the future, given more time, we would definitely implement direct database upload for a more automated end-to-end ETL pipeline.
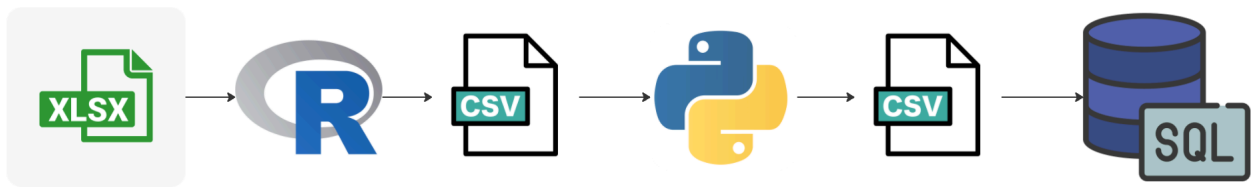

Figure 2: Data Architecture pipeline for Data Warehouse

III.   ETL Process

Our ETL Process was implemented in a Jupyter notebook (iata_ETL.ipynb) using python to transform our raw IATA and WDI indicators data into the star schema shown above.

1. Extract Data

   IATA transactions: contains 4,710,485 rows of aid activity records from 1900 to 2107

   ● Downloaded in XLSX so used R to convert to CSV

   WDI Indicators: pre-downloaded CSVs with country-year metrics for selected countries from 1960 to 2024.

2. Transform Data

   Dimension Tables:

   ● dim_time
   ● dim_organization
   ● dim_country
   ● dim_aid_type
   ● dim_sector
   ● dim_finance_type

- dim_flow_type
- dim_transaction_type

Fact Tables:

- fact_aid_transaction
- fact_country_context

Key Transformations Applied:

- Filtered years (1970-2024)
- Null handling: convert 'No Data' to None
- String cleaning: Mapping country names to a standardized format
- Regex Parsing: ex. Splitting '120 - Health' into '120' and 'Health'
- Surrogate keys applied to dimensions
3. Load
    - Exported each dataframe to CSVs
    - Imported manually into SQL server.

IV.   Queries, Visualizations, and Telling a Data Story

Visualization 1: Dominant Organization Types by Top Sectors

This query ranks organization types within top 10 sectors using DENSE_RANK()

The stacked bars reveal that multilateral organizations control a great percentage of aid funding across all major sectors except security system management, which is dominated by government organizations, which makes sense given security is a government matter.

From a managerial perspective, this query could be used to find relationships between organization types and specific sectors that would allow donors to match with relevant sectors.
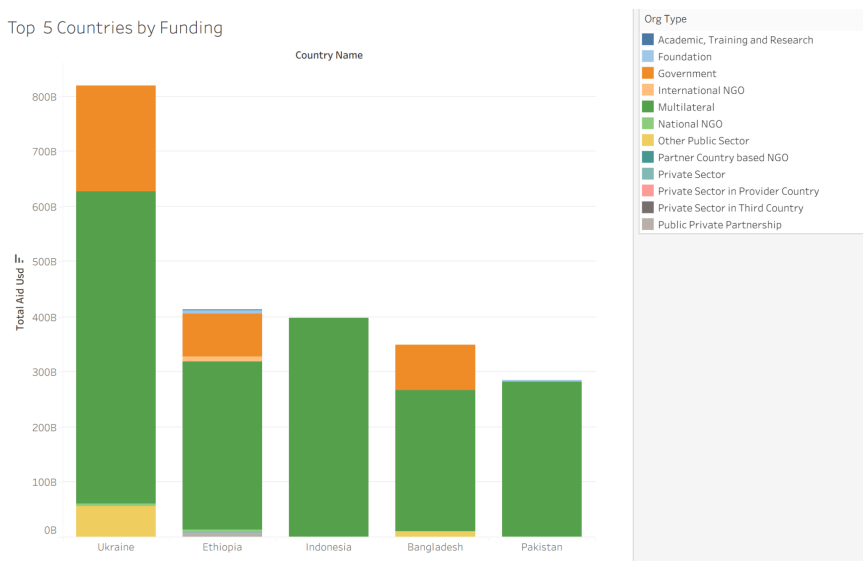

Visualization 2: Top 5 Recipient Countries by Total Aid

This query ranks the top 5 countries per organization type by total aid, using the RANK() function.

The stacked bar graph shows that Ukraine is very dominant in terms of overall funding and aid at $800B, which dwarfs Ethiopia's $400B second-place total. All of these countries are mostly funded by multilateral and government organizations.

Ukraine's overwhelming aid dominance likely highlights crisis funding from geopolitical conflicts between Ukraine and Russia, but will need more investigation to prove this.
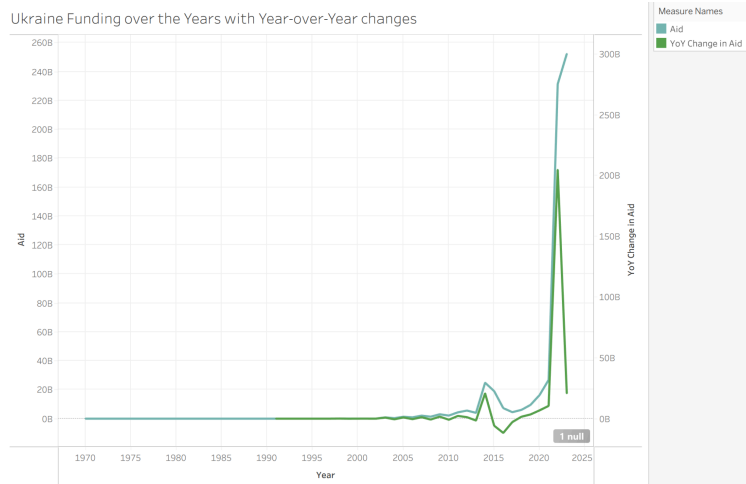

Visualization 3: Ukraine Aid Funding with YoY changes

Calculates Ukraine's year-over-year changes in funding using the LAG() function.

The $800B total from visualization 2 can be traced to 2022's large spike in a single year, which confirms our hypothesis regarding funding from Ukraine's geopolitical crisis. However, we can also see the aid flow beginning to drop off at around 2024, suggesting decreases in global attention towards Ukraine's crisis.

This sort of query helps establish crisis response benchmarks, being how much money one should expect to be sent in terms of future crisises and conflicts.

Ukraine Funding over the Years with Year-over-Year changes

V.　Conclusion

This project is a data warehouse to analyze how international financial aid flows between donors, sectors, and recipient countries over time. By cleaning and transforming raw IATI transaction data and enriching it with World Development Indicator metrics, we organized millions of records into a star schema centered on the FactAidTransaction table and supporting dimensions like country, organization, sector, and time.

Using SQL analytical queries, we identified major donor and recipient patterns, examined year-to-year funding changes, and detected longer-term trends. The analysis showed that aid is concentrated among certain countries and sectors, and that funding levels fluctuate significantly across years. Linking transactions to country context also allowed funding to be interpreted relative to development conditions rather than only total amounts.

Overall, the project demonstrates how proper data modeling and ETL preparation make complex transactional data easier to analyze. The resulting warehouse enables meaningful insights into global funding behavior and could support planning or policy evaluation, with future improvements focusing on automating the ETL pipeline and expanding dashboards.

References

International Aid Transparency Initiative. (n.d.). *Country Development Finance Data* [Data set]. IATI. https://countrydata.iatistandard.org/

World Bank. (2024). *World Development Indicators* [Data set]. https://databank.worldbank.org/source/world-development-indicatorss)