Mzwakhe Sefo

Maji Ndogo Project

Improving water access and safety for people in Maji Ndogo

Figure 1: Maji Ndogo water source.

Data Extraction & Transformation (ETL)

Data was imported from Md_water_services_data.csv into Power BI. The following Power Query steps were executed to ensure a clean dataset:

- **Removing Duplicates** – Identified and removed repeated records.
- **Handling Missing Values** – Addressed null and incomplete entries using appropriate imputation and reference lookups.
- **Standardizing Formats** – Unified data formats to ensure consistency across the dataset.

Data Modeling (Star Schema)

The model utilizes a Multi-Star Schema to ensure high performance and DAX accuracy.

Data model: Organized into 2 Fact Tables (visits & water_source) and 6 Dimension Tables.

Relationship Logic:

- **Cardinality:** Enforced 1-to-many relationships; many-to-many relationships were strictly avoided.
- **Bridge Tables**: Utilized location_id as a unique key to maintain referential integrity.
- **Filter Flow:** Set to Single Direction by default. Bi-directional filtering was only applied where functionally necessary to prevent ambiguous paths.
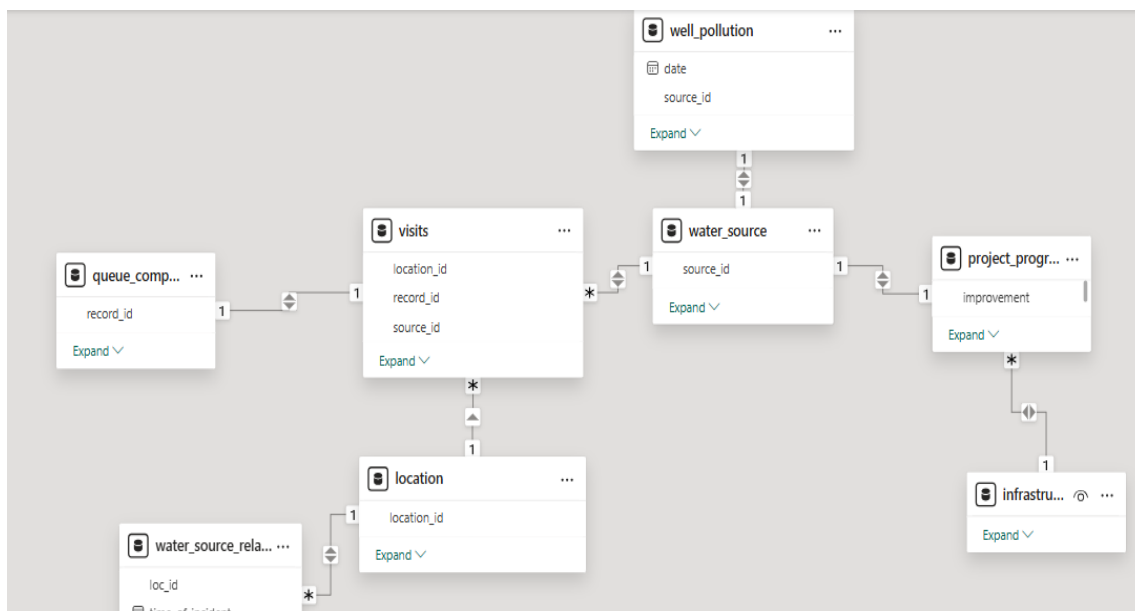
☞ **Aim for Star Schema**

Figure 2: Multi-star schema model.

Mzwakhe Sefo

DAX Transformations (Business Logic Layer)

To translate raw operational data into strategic insights, business rules were embedded directly into, the semantic model using DAX. These transformations ensured financial projections and service impact metrics were both accurate and decision focused.

- **Aggregated Improvements** – Standardized reporting by consolidating multiple "Install tap" categories into a unified group (*Install public tap(s)) to reduce categorical noise and improve readability.
- **Rural Cost Adjustment** – Incorporated contextual cost modeling by introducing a *Rural_adjusted_cost* column, increasing infrastructure costs by 50% for rural sources to reflect logistical and operational complexity.
- **Queue Time Aggregation** – Calculated *Average_queue_time* per source using controlled aggregation logic, enabling a good performance evaluation of shared water sources across multiple visits.
- **Water Access** – Classification aligned with UN service-level definitions to determine whether each source qualifies as "Basic Access" or "Below Basic Access." This created a measurable benchmark for project success.
- **Impact & Financial KPI Measures** – Developed measures to quantify both cost and societal impact:
    - **Basic Access %**
    - **No Basic Access %**
    - **Total Budget**
    - **Provincial Budget % Allocation**

Collectively, these transformations elevate the model from descriptive reporting to decision support analytics, directly linking infrastructure investment to measurable improvements in public water access.

Dashboard Results & Discussion

National Budget Overview shows, the total projected national upgrade budget is $146,737,375. Sokoto and Kilimani combined account for more than half of the total national budget. As a results, Sokoto and Kilimani are the highest-funded provinces. Amanzi has the lowest cost per citizen, making it the most cost-efficient province for infrastructure upgrades. Sokoto has the highest cost per citizen at $6.95 per citizen, indicating higher infrastructure intensity. These insights support funding prioritization discussions.
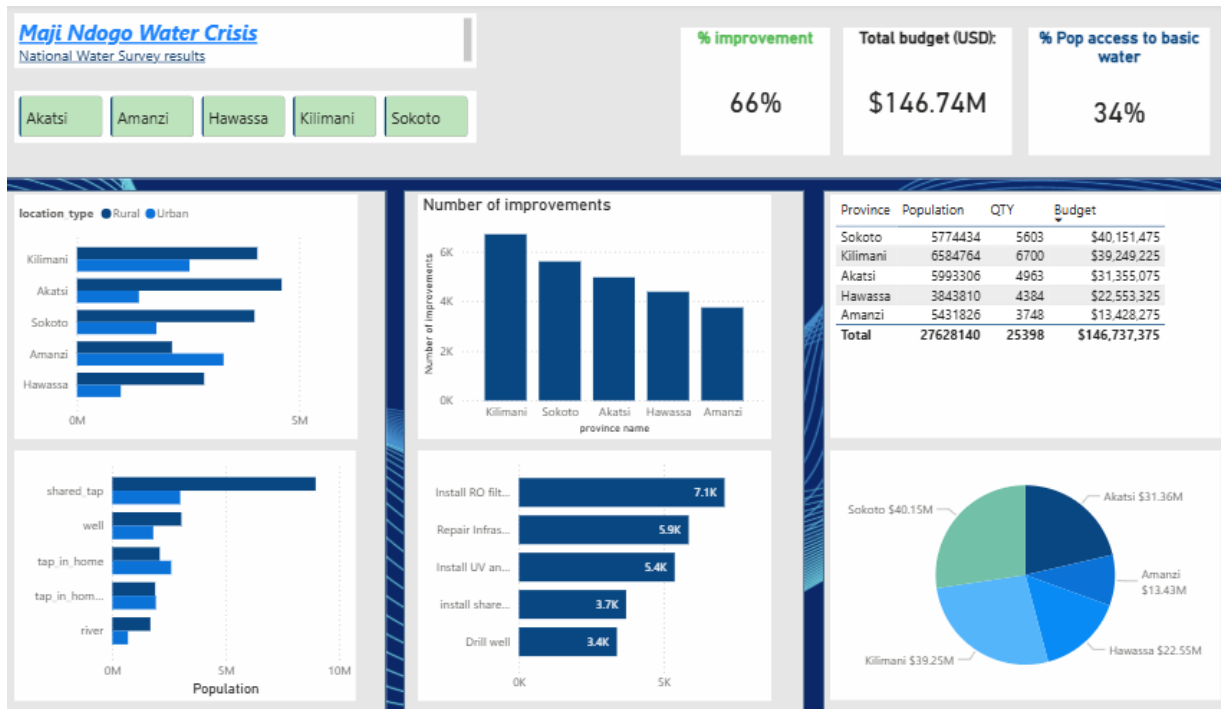

Figure 3: National report.

Crime Insights

Town results show (Zuri, Harare, Amina, Dahabu). Zuri leads in shared tap sites, Harare in wells, and Amina in broken taps. Crime patterns align with these infrastructure distributions, suggesting these hotspots are linked to access points.

There are 3,182 wells serving rural commiunities (Amanzi, Akatsi and Hawassa). Because many people depend on these wells, they can sometimes become targets for small thefts or damage. This shows that improving security should be an important part of planning and managing rural water systems.
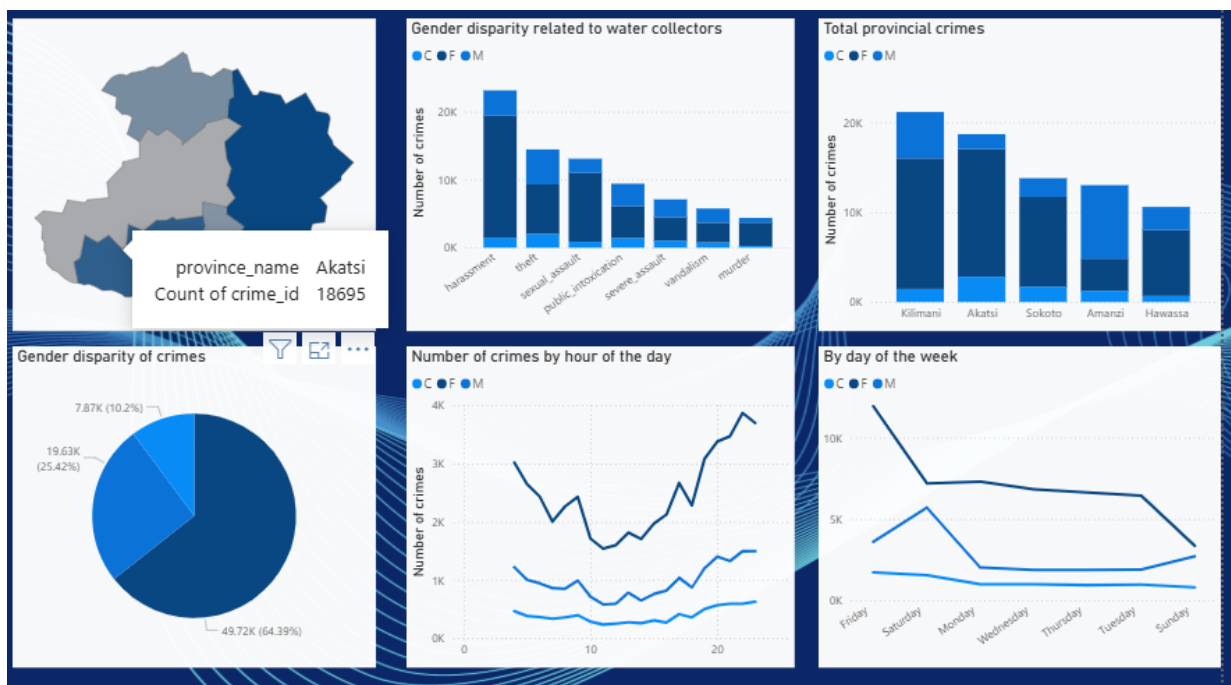

Figure 4: Crime dashboard.

Queue times:

- Men constitute **24% of queues** on Saturdays across Maji Ndogo, reflecting gendered patterns of water collection.
- **Weekday queues:** Wednesdays have the longest queues during weekdays, guiding scheduling and staffing for improved service.
- Optimal visit times for reduced wait times: **Thursday at 15:00**, with **2% children** in the queue, indicating low congestion periods for vulnerable populations.

Water Contamination & Safety:

- **Most chemically contaminated wells:** Akatsi (2,056 wells), highlighting public health risk clusters.
- **Crimes against males:** Akatsi reports the lowest incidents, suggesting gender-specific risk variations across provinces.
- **Crimes against children:** Lowest reported incidents occur at **Sunday, 12:00**, valuable for planning safety interventions.
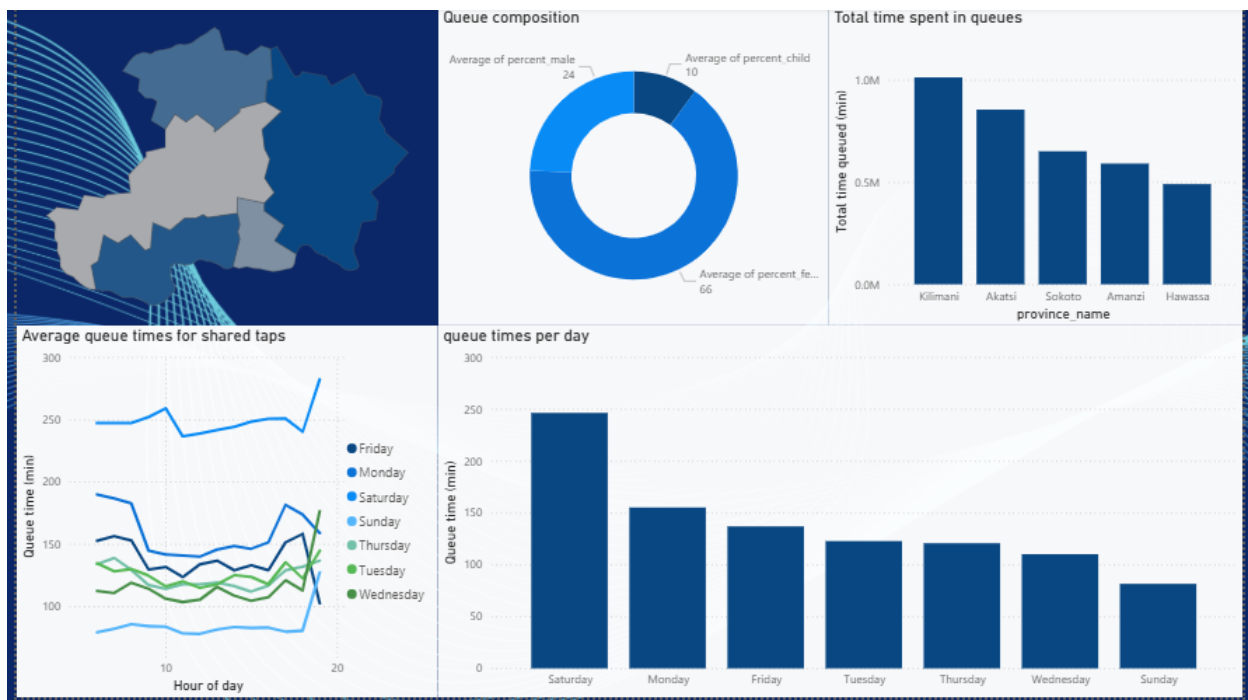- **Sexual assault gender disparity:** Highest difference between female and male victims occurs on **Friday**.

Figure 5: Queue times various water sources.