Data Handling – Pump it up: Analyzing the water table in Tanzania
Stuart King – June 2016

Cleaning:

- Installer: Used trim and proper functions; reviewed filtered results and normalized values; all null values were categorized as "unknown" (much more could be done with this variable to group categorically similar values such as "Government" and "Central Government")
- Region code: edited records that had been mis-coded by evaluating coding of other records (e.g. a record coded as 60 was actually region 6) – to do this I created a pivot table to display the minimum value in each region and then manually verified that the minimum value did indeed represent the majority of records for each region
- Management: combined similar categories ("company" with "private operator"; "water user group" with "water user association"; etc.)
- Created a new variable to group water points into buckets of fully functional or not (needs repair + non-functional)
- Created a new variable – Sea Level: if the GPS height was above 0 the pump is considered to be above sea level, if GPS height is below 0 it is considered to be below sea level, and if the GPS height is 0 then the water pump is at sea level
- Created a new variable – Payment Scheme or No Payment Scheme: this was created after observing that water points with no payment scheme have a greater proportion of points that are non-functional. To simplify and enhance my analysis I thought a cross-section of whether or not a water point has a payment scheme or not would be insightful.
- Calculated the number of years since a water point was installed and created a new variable of the results – used to segment data into age of water point
- Pulled in status for each water pump by adding the status data as a separate tab and then used an index/match function combination to assign status values to each pump matching pump id

Identified initial questions to answer:

- How many pumps are functional? Functional, but need repair? Non-functional?
    - Functional: 32,259
    - Functional, but need repair: 4,317
    - Non-functional: 22,824
- In what size community are pumps installed?
    - Vast majority of water pumps are installed in communities/villages of less than 500 people
- Where are the pumps located? (use lat/long to plot on map in Tableau)
    - Fill maps helped visually identify where the majority of pumps by function status were located
- What regions have the greatest number of broken pumps?
    - Mbeya, Morogoro, and Shinyanga have the highest proportion of total non-functional pumps, with 8%, 7%, and 6.8% of all broken pumps, respectfully
- What's the status of pumps by altitude?
    - 45% of all pumps below sea level are non-functional
- What's the status of pumps by water basin?
    - Lake Victoria and Pangani water basins have 18% and 14% of all non-functional water points
    - Ruvuma/Southern Coast and Lake Rukwa are the two regions with a higher proportion of non-functional water points than functional
- What's the status of pumps by extraction type?
    - Of water points using a mono extraction method, 58% of these water points are non-functional – nira/tanira water points have the highest proportion of functional water points (66%)
- What's the status of pumps by management type?

    - Village water committees (VWCs) manage the greatest quantity of water pumps (70% of all water points), but 43% of water points managed by VWCs are non-functional
    - Water boards and private operators have the highest proportion of water points that are fully functional (74% and 66% respectfully), but combine for less than 10% of all water points
- Does water quality seem to impact functionality of pumps?
    - No, the overwhelming majority of water points are soft water
- What about water source or water point type?
    - Sources of shallow wells and machine dbh have the highest proportion of water points that are non-functional (45% and 47% respectfully), and 53% of multiple communal standpipes water points are non-functional (but only represent 12% of all water points)
- Do certain installers have a higher amount of broken pumps than other installers? What proportion of installed pumps are broken for each installer? (pivot)
    - Top 10 installers constitute ~55% of all installed water points (30% have been installed by the Department of Water Engineer, DWE)
    - Finwater is the installer with the highest proportion of non-functional water points (80%), with private installers have the lowest (20%)
        - This is in part due to the majority of water points having been installed by Finwater over 20 years ago
    - 37% of DWE-installed water points are non-functional

Tableau manipulations:

- Created a new measure using the Create a Calculated Field option to categorized population size into categories/buckets of 500 increments (<500, 500-1,000, 1,000-1,500, etc.). This new measure was used to create a distribution of community size for the dataset.
- Created a new measure using the Create a Calculated Field option to segment the number of years since construction (<10 years, 10-20 years, etc.), which was then used to pool the number of water points per installer into the defined year categories.

Interpretation and relevance of findings:

- The Mbeya, Shinyanga, and Morogoro regions have a high number of non-functional and semi-functional water points; the Ruvuma/Southern Coast and Lake Rukwa river basins have a higher proportion of non-functional water points, which begs the question why.
- DWE installs the greatest proportion of water points in Tanzania by sheer volume. Improving installation methods, techniques, and/or materials used by DWE increases the odds of better water point functionality and sustainability.
- 70% of water points are managed by village water committees. Of these, 43% are non-functional. By targeting VWCs to improve management practices and maintenance capacity the number of functioning water points will increase.
- 58% of water points using the mono extraction method are non-functional. In contrast, 68% of water points using the afridev method, 66% using the nira/tanira method, and 65% using the rope pump method are fully functional. Focusing maintenance and upgrades on water points using extraction methods that have a statistically higher greater proportion of non-functional pumps will improve water point functionality and access to clean water.
- The majority of water points do not have a payment structure. Of water points without a payment structure, 48% of water points are non-functional. Conversely, only 25% of water points with a payment scheme are currently non-functional. Assisting water point managers/groups to devise and institute payment schemes will likely improve the functionality of water points and ensure clean water for users.
- While descriptively informative, the findings can be enriched with deeper analysis into the age of water points. For example, are more mono extraction water points non-functional because they are mechanically flawed, or the majority of water points installed in the 1980s use the mono extraction method?