Proposed Problem Statements:

World Bank

IDHack

Tufts University

February 2015

1. **Visualizing Tanzania’s population boom**
   1. Summary:
      1. Develop a series of visualizations that showcase Tanzania’s population growth and the effects it will have on water and education resources.
   2. Potential tasks:
      1. Localize the existing API with a taxonomy for sub-national Tanzanian terms (e.g. ward, district)
      2. Spin up an instance of *population.io* specific to Tanzanian data
      3. Merge *population.io* information with water and education data to extrapolate population growth’s impact on those resources in the future
   3. Context:
      1. Tanzania is among the twenty fastest-growing countries in the world. Its population is young (nearly half of all Tanzanians are under 15) and fertile (women have on average 5 children). The UN and World Bank project its population to grow from its current 45mn to 130mn by 2050 and over 300mn by the next century.
      2. Tanzania’s population growth will present new challenges to governments and citizens alike. Constrained by budgets and finite resources, already-fragile service delivery infrastructure could be stretched thinner as demand grows with the population.
      3. The country is undertaking renewed efforts to leverage data to make more informed decisions in the face of these emerging challenges. Under a new open data initiative, the government is keen to collect and publish information about water and education service delivery[[1]](#footnote-1) so that citizens can be more informed, contribute to more accurate data, and track changing conditions.
      4. *Population.io* is a new tool built on open frameworks that empowers citizens to understand age and population in their countries and worldwide. While this is an interesting tool at the macro level, it could be applied locally to the sub-country level to help decisionmakers and the public be more informed about population growth and its effect on many aspects of local life including education and water access.
   4. Details about *population.io*:
      1. Data: Currently *Population.io* runs on cleaned data sourced from the UN. These data help to compare countries against each other, but they are not granular enough for use within a specific country (e.g., comparing districts). Tanzania’s 2012 census data[[2]](#footnote-2) can be used to provide more specific data.
      2. Algorithm: To conduct its estimations of population at specific times and locations, *population.io* uses various interpolation techniques.[[3]](#footnote-3)
      3. API: *Population.io* leverages a RESTful API[[4]](#footnote-4) whose taxonomy runs down to the country level. The API should be improved to allow for sub-country units, including region, district, and ward.
   5. Supporting material:
      1. Original code: All code is released on Github[[5]](#footnote-5)
      2. Data: The World Bank will provide several datasets:
         1. School locations and performance countrywide
         2. Water point locations countrywide[[6]](#footnote-6)
   6. Impact:
      1. The visualizations and API created by participants of ID Hack could be a) incorporated into current World Bank open data operations in Tanzania; and b) shared openly with other donor and development agencies in the region.
2. **Making open data dashboards more relevant**
   1. Summary:
      1. Develop an iconographic schema / design theme to make existing open data dashboards more relevant and readable by Tanzanians.
   2. Potential tasks:
      1. Develop a visual identity guidebook for open data dashboards
      2. Rebuild existing open data dashboards using new graphical iconography
   3. Context:
      1. Tanzania is undertaking renewed efforts to leverage data to make more informed decisions in the face of these emerging challenges. Under a new open data initiative, the government is keen to collect and publish information about water and education service delivery[[7]](#footnote-7) so that citizens can be more informed, contribute to more accurate data, and track changing conditions.
      2. A prototype education open data dashboard[[8]](#footnote-8) was built in 2014 and continues to draw feedback from users (Tanzanian government and citizens). This dashboard tracks primary and secondary school performance against key indicators, but information is still presented graphically in conventional tables/charts. End-users who may have low literacy levels need a more graphical representation of the information.
   4. Details about the prototype education open data dashboard:
      1. A prototype education open data dashboard[[9]](#footnote-9) was built in 2014 and continues to surface feedback from users (Tanzanian government and citizens). This dashboard tracks primary and secondary school performance against key indicators.
   5. Supporting material:
      1. OCHA iconography[[10]](#footnote-10), Noun Project, and Iconmonstr may provide useful visual cues.
   6. Impact:
      1. The iconographic schema / design theme created by participants of ID Hack could be incorporated into a countrywide roll-out of a suite of national open data dashboards within the next 12 months.
3. **Digitizing budget data**
   1. Summary:
      1. Digitize Tanzania’s budget reports to support open data and civic innovation.
         1. Digitize into what format? **JSON? If so, how should be it organized?**
   2. Potential tasks:
      1. Develop a scraper to find and download budget PDFs from the Parliament of Tanzania’s public expenditure database[[11]](#footnote-11)
         1. All links?
         2. Keyword search on each document?
            1. What keywords?
      2. Develop a scraper to find and extract key line information from these PDFs into machine-readable format.
   3. Context:
      1. Open data continues to drive better service delivery and civic engagement. Unfortunately, critical service delivery information is often not published in machine readable format, making it difficult for governments and citizens alike to access, search, manipulate, visualize and analyze that data.
      2. Donor agencies, civil society organizations, government agencies and other stakeholders hold a common interest in making data, including budgets, more accessible. Doing so increases accountability and transparency while fostering civic innovation and engagement opportunities for civil society to interrogate, analyze, and publish meaningful stories.
      3. Supporting material:
         1. All budget data for 2014-2015 is available online in PDF format.
      4. Impact:
         1. The data will feed directly into planning and operationalizing the ongoing Dar Es Salaam Metropolitan Project, a major $600 million World Bank program to improve the quality of urban service delivery and institutional capacity in Tanzania’s capital.
      5. Questions
         1. What are the different documents?
            1. National Irrigation Comm Board?
            2. Katavi Region
            3. Ministerial v. regional

1. See http://www.pmoralg.go.tz/quick-menu/brn/ [↑](#footnote-ref-1)
2. See http://www.nbs.go.tz/ [↑](#footnote-ref-2)
3. See https://github.com/worldpopulation/population.io-api/tree/master/modeling [↑](#footnote-ref-3)
4. See http://api.population.io/ [↑](#footnote-ref-4)
5. See https://github.com/worldpopulation/ [↑](#footnote-ref-5)
6. An API for water points has been created and could be leveraged to merge data sources in visualizations. See https://github.com/taarifa/TaarifaAPI [↑](#footnote-ref-6)
7. See http://www.pmoralg.go.tz/quick-menu/brn/ [↑](#footnote-ref-7)
8. https://tz-govstat.demo.socrata.com/en/stat/goals/default [↑](#footnote-ref-8)
9. https://tz-govstat.demo.socrata.com/en/stat/goals/default [↑](#footnote-ref-9)
10. http://reliefweb.int/report/world/world-humanitarian-and-country-icons-2012 [↑](#footnote-ref-10)
11. <http://www.parliament.go.tz/index.php/home/pages/44> and <http://www.mof.go.tz/index.php?option=com_content&view=category&layout=blog&id=20&Itemid=560> (for 2014-2015) [↑](#footnote-ref-11)