A Web-based CBMS Dataset Visualization and Simulation Tool

Nelson Marcos[1,*], Gerardo Largoza[2], Briane Paul Samson[3],
John Jelvin S. Base[4], Lawrence Patrick C. Calulo[5], Bervyn S. Co[6], Charles Geoffrey R. Lo[7]

[1] *College of Computer Studies, De La Salle University*
[2] *School of Economics, De La Salle University*
[3] *College of Computer Studies, De La Salle University*
[4,5,6,7] *College of Computer Studies, De La Salle University*
*nelson.marcos@dlsu.edu.ph

Abstract: A visualization and simulation tool for community-based datasets, from Angelo King Institute (AKI) of De La Salle University, called Community-Based Monitoring System (CBMS) is designed. Currently, there is no local visualization and simulation tool for CBMS data, which were gathered from individual households from various barangays in the country. The data in the CBMS dataset is mainly categorized into eight key indicators considered to be the basic needs of a local government unit and these are health, nutrition, housing, water and sanitation, education, income, employment, and peace and order. This research is useful not only for government officials, and local communities in creating or formulating policies, but also for economists. This tool is developed to give policymakers and program implementors necessary information to study the effect of macroeconomic reforms and different kinds of policy shocks. The application is designed to be web-based. It allows users to upload and extract data from the CBMS dataset through the data management module, to visualize the data with the use of maps (over Google Maps) and charts to provide a more accessible summary of the conditions inside a community, and simulate new indicators by providing simple formulas and view the results of effects adjusting the indicator reference values. Through the results based on economic computations performed on the data, this can help the government units in deciding what policies to implement in a certain community. Several prototypes were presented to CBMS AKI and various local government units in order to consider their needs and preferences. The final application user-interface and features were deemed very acceptable to the intended users.

Key Words: community data, visualization, simulation, maps, economics

1. INTRODUCTION

Visualization and simulation is important especially with regards to data analytics and processing datasets. Considering the capabilities of current technologies, visualizing and simulating big data has evolved from the traditional Command Line Interface (CLI) to the modern Graphical User Interface (GUI). Visualizing data and converting it into graphics while being able to manipulate and simulate it can aid data analytics in a fast and productive way.

There are ways to visualize and simulate data such as using software like IBM Cognos, a Web-based suite by IBM (2012). Another demonstration of visualizing and simulating big data in the field of economics is Outline (Bier et al., 2013a). It provides a platform for open policymaking by allowing users to interact with a simulation of the US economy and observe the effects of creating and modifying policies (Bier et al., 2013b).

This paper presents a web-based tool that can be used to visualize and simulate datasets based on the needs of researchers or end-users, primarily using Community-Based Monitoring System (CBMS). CBMS is the first way of collecting and processing information about the development and current status of the households (Reyes and Valencia, 2009) in the Philippines. In cities or provinces, CBMS is used as a tool to gauge the poverty level (Reyes and Valencia, 2009).

CBMS provides assistance to its members in order for them to develop. This tool would give policymakers the necessary information to study the effect of macroeconomic reforms and different kinds of policy shocks (DILG Region VI, 2010).

The data in CBMS is categorized into 8 key indicators, which are considered to be the basic needs of a local government unit. These criteria are: health, nutrition, housing, water and sanitation, education, income, employment, and peace and order. Other information may include, but are not limited to, migration of families, community/political participation and access to government programs.

With the availability of the CBMS, it is no surprise that other research efforts have been conducted in the field of poverty reduction. One such example is Tan (2010), which discusses the use of regression modeling and spatial analysis in order to simulate strategies on reducing poverty in a community. CBMS and the Statistical Package for the Social Sciences (SPSS) were used. SPSS was used to generate poverty models. CBMS data were then fed to these models.

CBMS is only a method for collecting and processing current household data and its development (Reyes and Valencia, 2009). Because of the unavailability of a modern way of processing and viewing of these data, the motive of the research is to provide a visualization and simulation application of this dataset. Users are provided with features to select data to be visualized in table form, chart form, or on maps. Users are also allowed to define indicators and input the corresponding formulas for the indicators and simulate what these indicators show as reference values are adjusted. This allows policy makers to see how good or bad a local government unit is based on an indicator.

2. METHODOLOGY AND DESIGN

The project was developed with the help of DLSU School of Economics and the DLSU Angelo King Institute Community-Based Monitoring System (AKI CBMS). Data used in the system were gathered from DLSU AKI CBMS. Prototypes were presented to the DLSU School of Economics, DLSU AKI CBMS, and local government units to meet the user needs.

The system is designed to be composed of three major modules, each providing one category of functionality required by the system. This is shown in Fig. 1. The three major modules are: dataset management module, visualization module, and simulation module.

The dataset management module handles the import, storage, and querying of datasets to be used in the system. This module specifically allows users to:

- upload datasets
- request access for datasets
- retrieve desired datasets
- select subsets of datasets (like specific fields only or number of entries)
- download filtered datasets
- modify datasets

- specify units of measurement for fields of a dataset
- perform simple manipulations on datasets
- upload metadata
- upload shapefiles needed for maps
- accept or reject other users requesting access for dataset
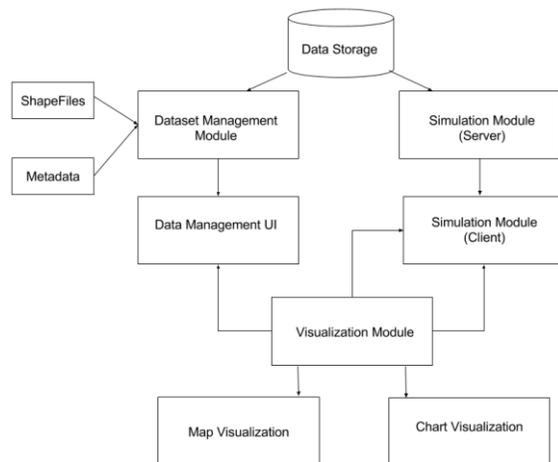
Fig. 1. System Modules

Sample screenshots for this module are presented next. Datasets can be downloadable in whole or specifying number of rows and variables to be downloaded. A sample is shown in Fig. 2.
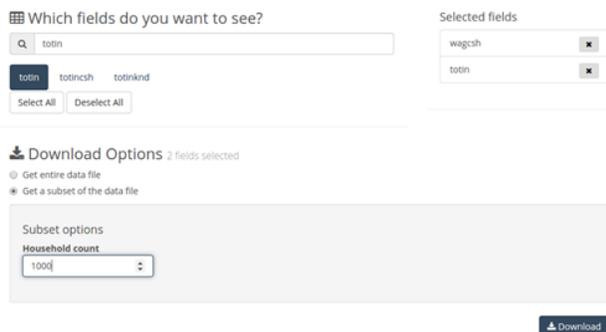
Fig. 2. Downloading Dataset

Shapefiles are uploaded in order to view the boundaries per "barangay". These are needed in order to view barangay-level visualization on map visualization module. A sample is shown in Fig. 3.

Fig. 3. Uploading Shapefiles

Users can request for access to some datasets they are interested in. A sample screen shot of a request is shown in Fig. 4.
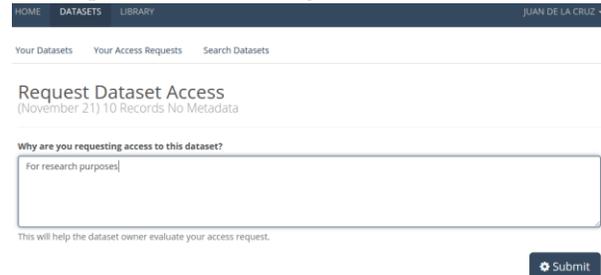
Fig. 4. Requesting Access for Datasets

The visualization module is responsible for the display of datasets as visualizations. Through this module, the user can select from a list of datasets, visualize data in chart form, and visualize data in map form, among others. In this module, a user specifically can:

- choose from a list of datasets for visualization
- transfer necessary data on demand
- visualize data using charts (pie chart, line graph, bar graph, histogram, scatter chart, etc.)
- edit or customize charts to suit user needs for their research and analysis
- download charts
- visualize data using maps (over Google maps in satellite or non-satellite mode)
- view household data per barangay

- adjust sliders to determine future status of household and barangays
- view markers and boundaries within a barangay or municipality

The succeeding screenshots presents sample chart visualizations. Fig. 5 shows a sample line chart and Fig. 6 shows a sample bar chart that can automatically be generated by the system given a user selected field.


Fig. 5. Sample Line Chart
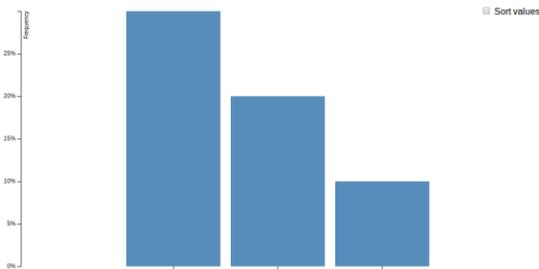
Fig. 6. Sample Bar Chart

Editing of charts can be done through chart editor by clicking the edit button while downloading the chart can be done by clicking the download button as shown in Fig. 7.
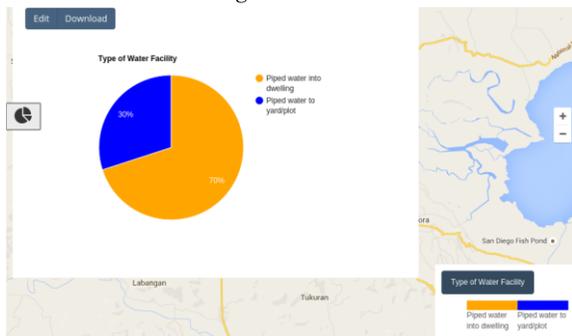

Fig. 7. Editing and Downloading Charts

For map-based visualizations, Google Maps is used for the geographic map, allowing users to zoom in or out, and pan around the map, as one would expect with Google's mapping tool, view maps in satellite or non-satellite mode. Visualizations can be household-level or barangay-level.

Fig. 8 shows the visualization of households in a barangay. In this type of visualization, the households are grouped per barangay and classify colors according to how poor or non-poor the barangays are.
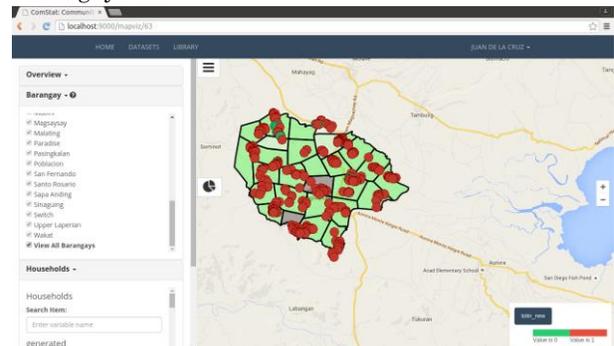
Fig. 8. Barangay-Level Visualization

The simulation module handles the creation of new simulations given a dataset. The simulation module works alongside the data management module and visualization module to perform calculations on the records of a dataset. This module specifically allows users to:

- view a list of available or previously saved simulations
- define indicators or new variables for simulation
- define formulas for indicators or new variables
- run simulations
- see simulation results as visualizations
- save indicators or new variables as a new part of a dataset
- save simulations

Fig. 9 shows a screen for simulation where the user enters information needed to run simulation such as simulation name, variable name, and the formula.

Fig. 9. Creating Simulations

After running the simulation, the user can visualize the simulated variable as shown in Fig. 10. Visualizing simulated variable will render green and red colors compared to blue and orange for non-simulated variable visualization.

Fig. 10. Visualization of Simulated Variable

Basic statistics are also displayed after the user selects an item. As shown in Fig. 11, the user can see the mean, median, mode, minimum and maximum if the item selected is a numerical variable.

Fig. 11. Viewing of Statistics of Numerical Item

3. RESULTS AND DISCUSSIONS

One CBMS dataset used in testing the system is that of Barangay Ramon Magsaysay in Zamboanga containing more than 3000 households. Results of the tests were presented to DLSU AKI CBMS, DLSU School of Economics, and local government units. Some examples of results, which were all validated correct, are presented next.

Fig. 12 shows visualization of a single variable (total income per household) over a map in non-satellite mode. Households within each barangay are shown. Red and green are used to indicate whether households are below or above the desired income level. Users are allowed to select specific barangays only.
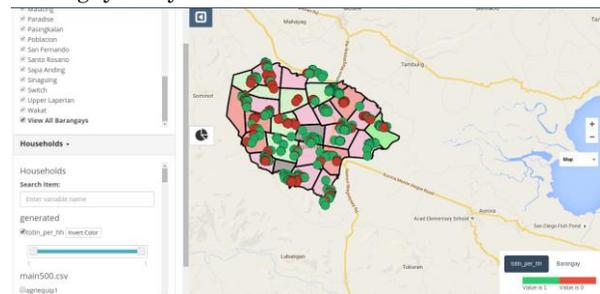

Fig. 12. Visualizing households and barangays

Fig. 13 shows an example of a test on visualizing two variables on a map in satellite mode. The variables compared are total income in cash versus in kind.
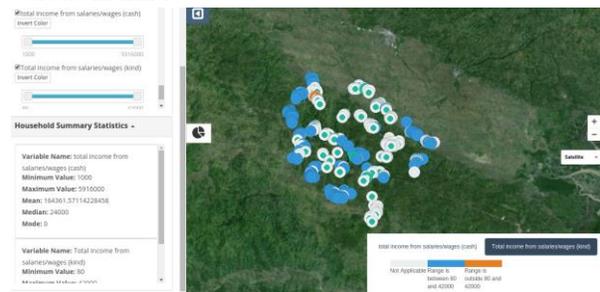
Fig. 13. Visualizing two variables on satellite map

Fig. 14 shows visualization of water facility per household, a non-numeric variable. The households are color-coded depending on the type of water facility.
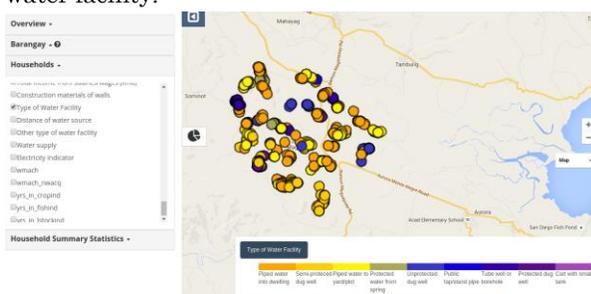

Fig. 14. Visualizing non-numeric variables

Fig. 15 shows a chart visualization of the households in Fig. 14. These visualizations allow users to analyze and evaluate status of households and barangays which can help local government units in planning and implementing programs be it for health, for the environment, or for sustainability.
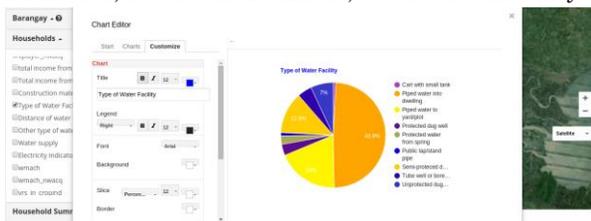
Fig. 15. Chart Visualization

4. CONCLUSION

The newly developed web-based tool provides a lot of features that can help users in extracting data and making analysis and evaluation of what the data shows. Through the chart and map visualizations, they can formulate policies, plan program to implement, identify barangays needing attention, among others.

Some features that are recommended for future development include: linking the system to a data gathering tool so that newly acquired household data can be immediately used in the system, developing a metadata creator to allow users to specify data formats to avoid data reading errors, and accepting more complex formulas for simulation.

5. REFERENCES

Bier et al. (2013a). Outline | The Platform for Digital Government. Retrieved June 25, 2013 from http://outline.com/.

Bier et al. (2013b). Outline: The Platform for Digital Government has Arrived. Speech given at Boston, MA.

DILG Region VI (2010). Community-Based Monitoring System (CBMS). Retrieved June 25, 2013 from http://www.lgrrc6.org.ph/index.php?option=com content&view=article&id=113&Itemid=2.

IBM (2012). IBM { Cognos Software. Retrieved October 3, 2013 from http://www-01.ibm.com/software/ph/analytics/cognos/.

Kosslyn, S. (1994). Elements of Graph Design (2nd Ed). New york City, USA: W. H. Freeman and Company.

Krug, S. (2000). Don't Make Me Think: A Common Sense Approach to Web Usability (2nd Ed). USA: New Riders Publishing.

Reyes, C. M. and Valencia, L. E. (2009). Poverty Reduction Strategy and Poverty Monitoring: Philippine Case Study. Retrieved February 14, 2010 from http://siteresources.worldbank.org/INTPAME/Re sources/Country studies/philippinespovmonitoringcasestudy.pdf.

Tan, K. (2010). Simulating Poverty Reduction Strategy with the use of Regression Modeling and Spatial Analysis (Master's thesis).