Kosovo Mosaic Visualization Documentation

|  |  |
| --- | --- |
| **Document** | Kosovo Mosaic Visualization Documentation |
| **Version** | 0.1 |
| **Last Updated** | 27 November 2015 |
| **Updated by** | Brett Romero |

# Introduction

The Kosovo Mosaic visualization is a tool designed to help users understand and explore data collected as part of the Kosovo Mosaic project. This data contains a variety of data points reflecting the level of satisfaction or dissatisfaction of citizens for a range of public services and processes. The data also covers all municipalities and 5 (non-consecutive years) – all of which can be viewed.

The visualization is designed to provide access to this data in a user-friendly way while also allowing even basic users the ability to drill down to the detailed statistics relevant to their municipality.

# Structure

There are three different dimensions along which the user can interact with the data. These are:

1. The survey year – 2003, 2006, 2009, 2012 and 2015 (at the time of writing).
2. Satisfied/dissatisfied – The possible responses for respondents were ‘Very Satisfied’, ‘Somewhat Satisfied’, ‘Don’t Know’, ‘Somewhat Dissatisfied’ and ‘Very Dissatisfied’. The ‘Satisfied percentage’ reflects the percentage of people who were ‘Somewhat’ or ‘Very Satisfied’. The ‘Dissatisfied percentage’ reflects the percentage of people who were ‘Somewhat’ or ‘Very Dissatisfied’. They are typically the inverse of each other, with the difference resulting from the people who responded ‘Don’t know”.
3. The indicator – The indicators are a list of public services and processes about which citizens were asked to provide their level of satisfaction or dissatisfaction.

## Landing Page

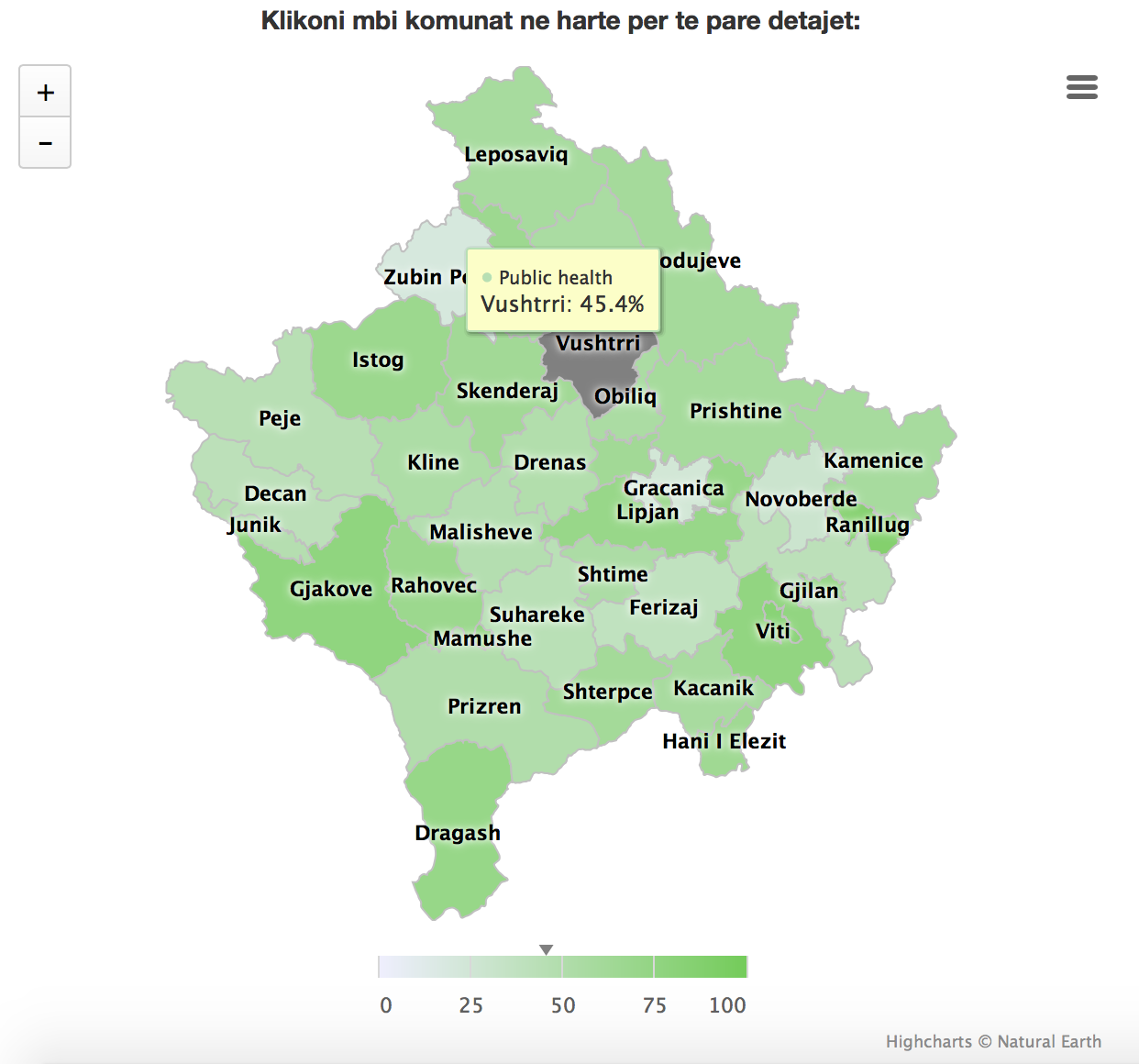
The landing page provides the user with options to change the selections for all of the dimensions mentioned above:

1. The survey year can be selected using the radio buttons in the top right hand corner of the page. Only one year can be selected at a time.
2. To switch between the satisfied (the default) and dissatisfied percentages, the option is provided with two face buttons, one happy and one angry, at the top center of the page. Clicking these buttons will switch between the two views.
3. To modify the indicator, the user is provided with a drop down box at the top right hand corner of the page.

To display this data, the visualization incorporates 5 individual charts (4 of them distinct). On the Landing page, two charts are displayed:

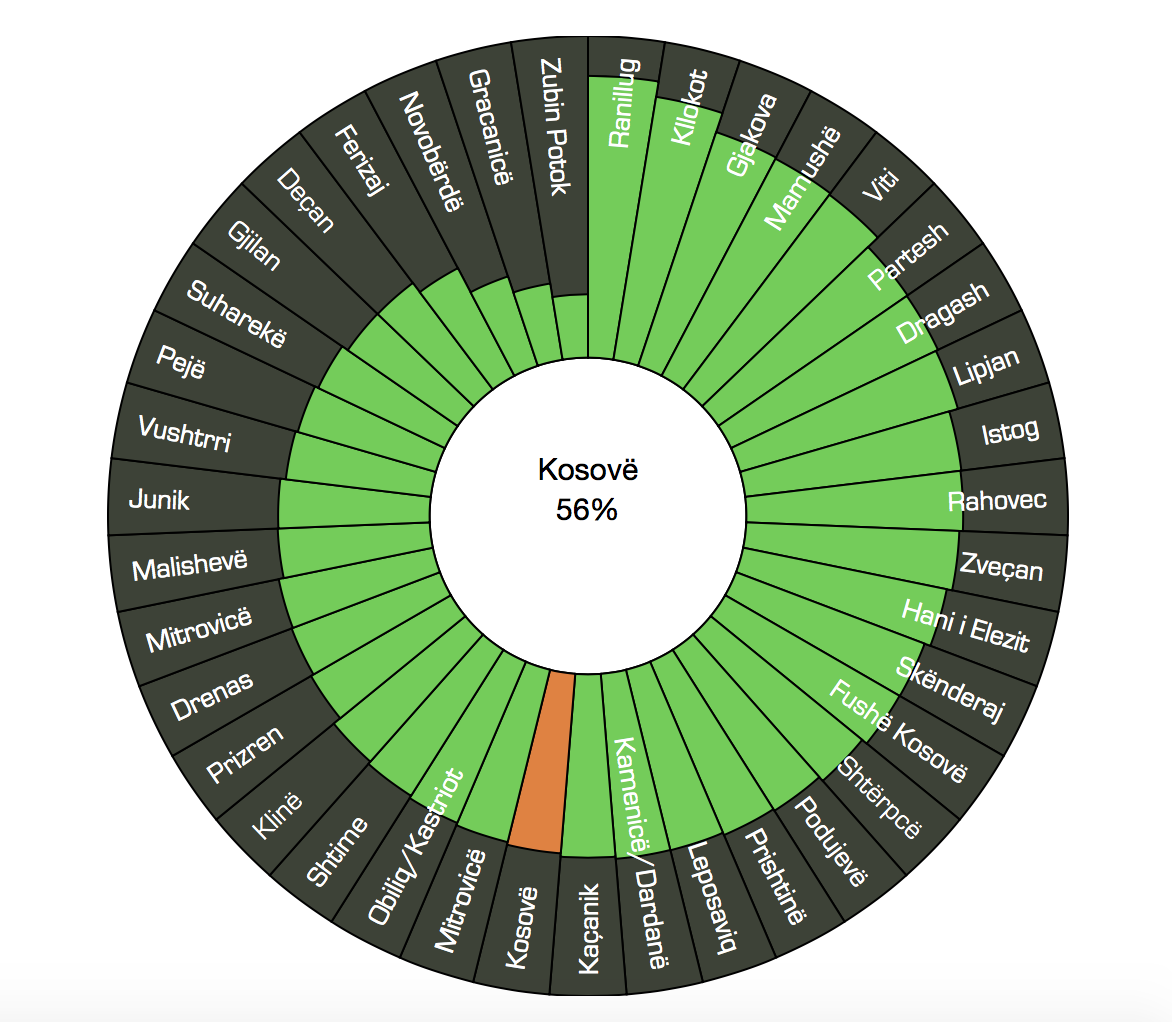
1. An interactive Map of Kosovo (see Figure 1) that displays, for the selected indicator and year, the percentage of satisfied or dissatisfied citizens for each municipality through varying shades of green or red respectively.

#### Figure 1 – Interactive Map of Kosovo



1. An Aster Chart that has two different views:
   1. The default view (see figure 2) displays a ranking of the municipalities for the selected indicator and year, based on the level of satisfaction/dissatisfaction.
   2. The other view displays the Kosovo average level of satisfaction/dissatisfaction across all indicators for the selected year. It should be noted that in this view, selecting an indicator will update the selection as would occur if the user selected it from the dropdown box.

#### Figure 2 – Aster Chart Default Municipality Ranking View



## Municipality Profile Page

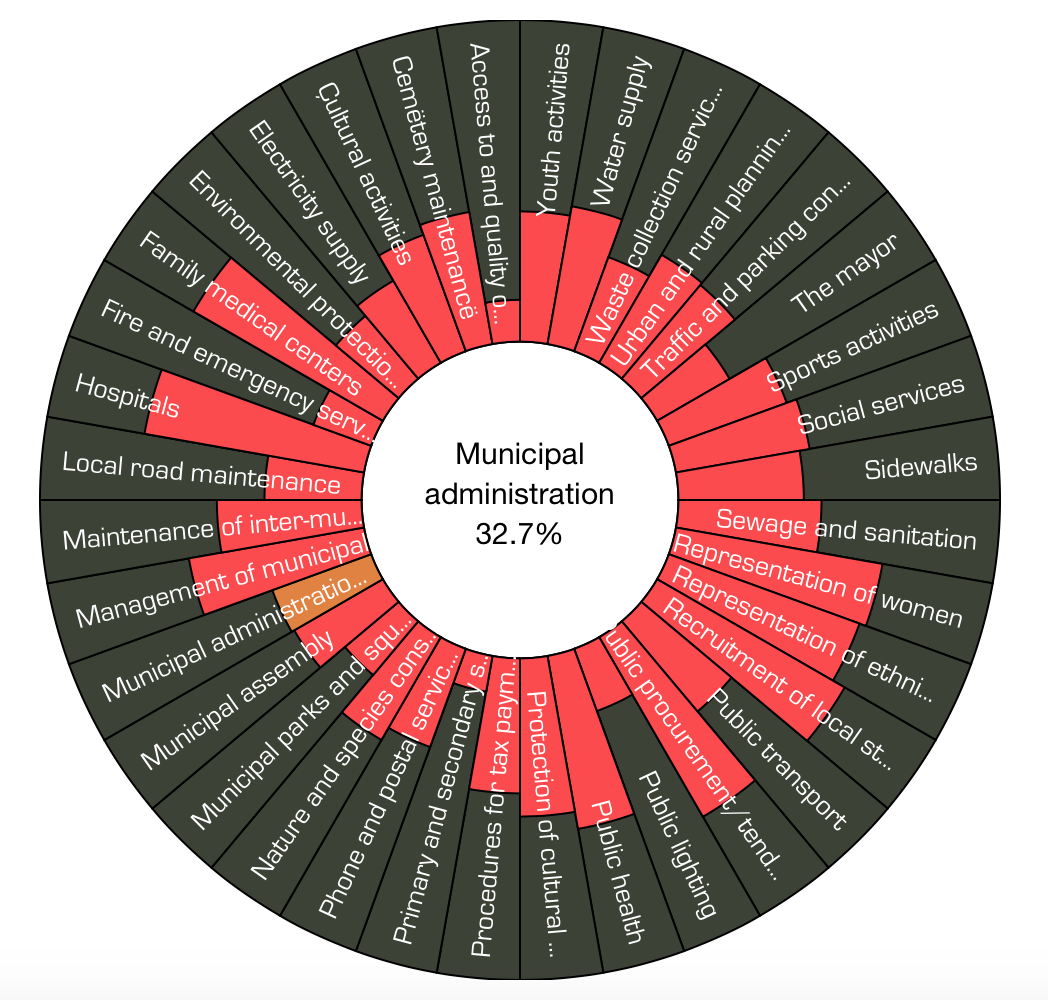
The user can access a second ‘Municipal Profile’ page by clicking on a municipality in the Map. On this page the user has access to similar options as the Landing page for changing their selections. The year selector buttons and satisfaction/dissatisfaction options are similarly located on the Municipality profile page.

One option that is slightly different is that the indicator dropdown is not available on this page. To select a different indicator on the Municipality Profile page, the user selects the indicator on the Aster chart that will then update the Line chart (see below for details) for the selected indicator.

Overall, the Municipality Profile page displays 3 charts:

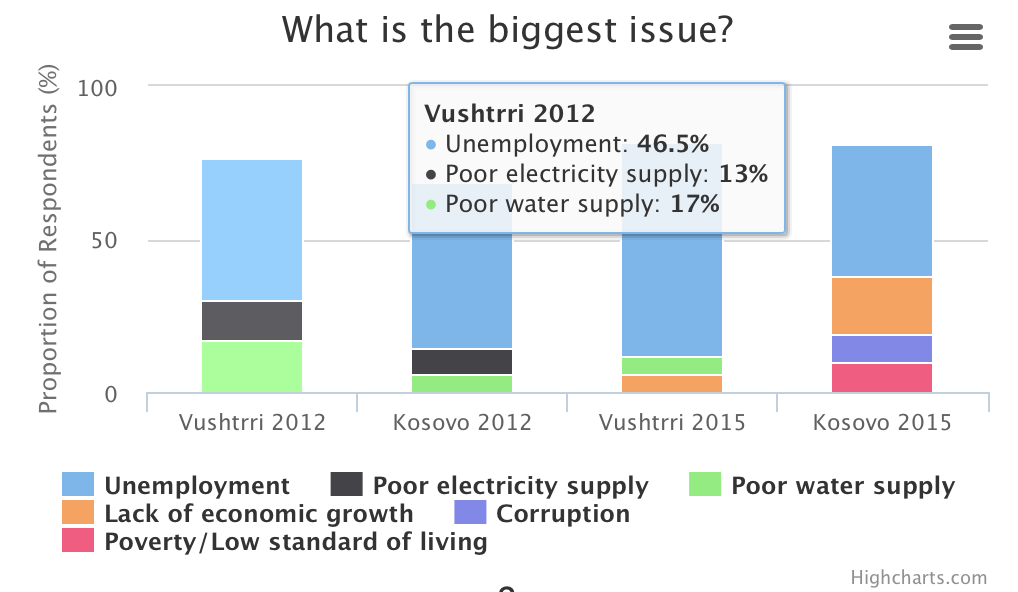
1. A new version of the Aster Chart from the Landing page. This chart now displays the levels of satisfaction/dissatisfaction across all indicators for the selected year. As mentioned above, clicking an indicator on this chart will update

#### Figure 3 – Aster Chart, Municipality Profile page



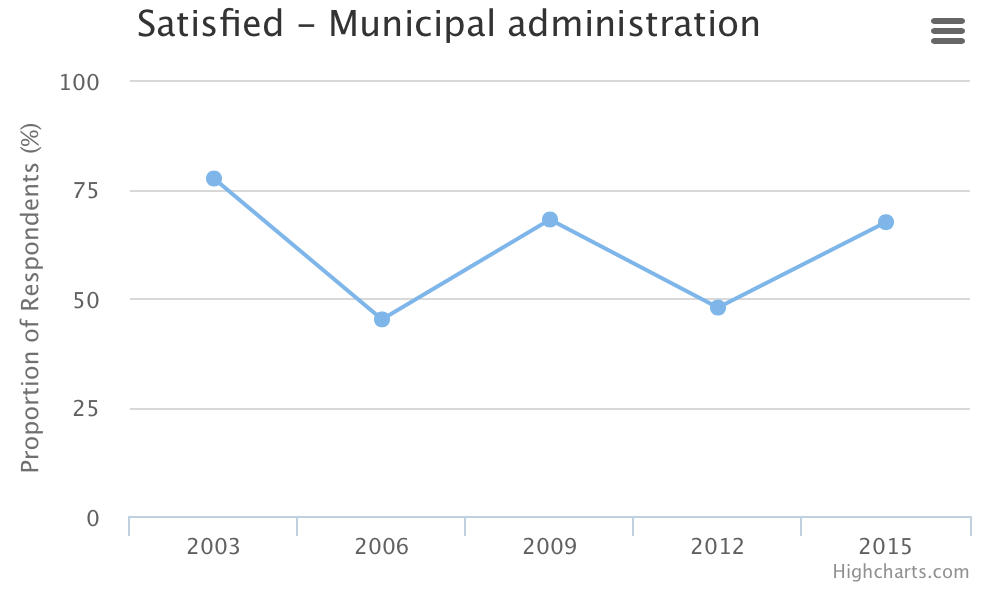
1. A stacked column chart. This chart displays data from a different dataset that details the three main things that people perceived to problems in their municipality. The column chart represents these problems and the percentage of people who listed them. This is then compared to the Kosovo average for all years available.

#### Figure 4 – Stacked Column Chart, Main Problems



1. A line chart. This chart displays the satisfaction/dissatisfaction level for the selected indicator across all years available. This allows users to see trends in how that indicator has changed over time.

#### Figure 5 – Line Chart, Indicator Satisfied Percent over Time



# Information Pages

## Indicators Page

## Problems Page

# Technology

Visualization runs on HTML and javascript.

Repository also includes a Python class and script that is used for processing data.

# Updating

The Kosovo Mosaic visualization has been designed in such a way that it is as dynamic as possible. In practical terms, this means that by updating the underlying data files, the visualization should also update and render correctly, provided it is in the correct structure and is processed correctly. The following section provides a guide on how to update the visualization.

## Introduction to Github

Currently, the entire visualization is being hosted, free of charge using GitHub. Github is a service that allows users to create public ‘repositories’ to store code, free of charge. As of 2015, GitHub added a new service named ‘gh-pages’. This service provides users with free hosting if the repository is a web site/app and the repository is structured in a certain way.

Assemblio, in creating this visualization, has utilized this service as a way of guaranteeing free hosting and easy maintainability.

### Accessing the Repository

### Cloning the Repository

## Updating the Data

Generally, the current files should serve as templates for the format of any new updated data added. However, there are certain additions that can be made to the files that will update the visualization in the expected format. Broadly, there are three files that need to be updated to update the visualization.

### Satisfied/Dissatisfied Data

The satisfied/dissatisfied datasets are saved in the data/raw\_data file location and are saved, separately, as CSVs. The file name (except the extension) is flexible – they will be handled in the processing step.

Using the current files as templates (data/raw\_data /satisfied.csv and data/raw\_data /dissatisfied.csv), the files can be effectively added to or modified by:

1. Adding new rows to the data. These rows can have new indicators and/or new years, and do not need to have values for all municipalities. If new indicators are added, they also need to be added to the standard lists file.
2. Modifying existing rows. The values, indicators and years can be modified. Again, if indicators are being modified, corresponding changes will also need to be made to the standard lists file.

##### Important Notes

1. It is **strongly** recommended that you modify a copy of the original satisfied.csv and dissatisfied.csv files. This will ensure you can check the format and structure of the originals if your updates are not working as expected.
2. Changing the names of the municipalities from the names currently in the file will cause the data for the changed municipalities to not display (the names are required to match names used to generate the Map visualization).
3. Do not use merged cells. Merged cells will not work in a CSV file, but please ensure that you simply repeat the value in the column for all the rows you would typically merge.
4. Format the numbers as decimals rather than as percentages. For example, format the numbers as 0.157, not 15.7%.
5. Additional indicators also need to be added to the standard list files (see below). If new indicators are added to the satisfied/dissatisfied data *without* adding them to the standard lists, the visualization will not update correctly.
6. Adding “satisfaction with” or “dissatisfaction with” to the name of the indicator is unnecessary. These are added automatically by the visualization.

### Problems Data

The process for updating/modifying the problems data is very similar to the process for updating the satisfied/dissatisfied data. The problems dataset is saved in the data/raw\_data file location and is saved as a CSV. The file name (except the extension) is flexible – it will be handled in the processing step.

Using the current file as a template (data/raw\_data /problems.csv), the file can be effectively added to or modified by:

1. Adding new rows to the data. These rows can have new problems and/or new years, and do not need to have values for all municipalities. If new problems are added, they also need to be added to the standard lists file.
2. Modifying existing rows. The values, indicators and years can be modified. Again, if indicators are being modified, corresponding changes will also need to be made to the standard lists file.

##### Important Notes

1. It is **strongly** recommended that you modify a copy of the original problem.csv file. This will ensure you can check the format and structure of the original if your updates are not working as expected.
2. Changing the names of the municipalities from the names currently in the file will cause the data for the changed municipalities to not display (the names are required to match names used to generate the Map visualization).
3. Do not use merged cells. Merged cells will not work in a CSV file, but please ensure that you simply repeat the value in the column for all the rows you would typically merge.
4. Format the numbers as decimals rather than as percentages. For example, format the numbers as 0.157, not 15.7%.
5. Additional problems also need to be added to the standard list files (see below). If new problems are added to the problems data *without* adding them to the standard lists, the visualization will not update correctly.

### Standard Lists

The second step required for updating the visualization is only required if new indicators/problems are added to their respective datasets, or existing indicators/problems have been changed. *If no changes have been made to the indicators or problems, this step is not required.*

This file that needs to be modified in this section is in the following location in the file structure: data/standard\_lists/whitelist.xlsx

##### New/Modified Indicators

If new indicators have been added or existing indicators modified, then those changes will need to be reflected in the lists file used by the visualization.

In this file you will find the three standard lists used by the application. For this section the tab of concern is the ‘indicators’ tab.

The data in this tab reflects the list of all indicators in the visualization and contains the following columns:

1. Index – this column contains the *internal* name of the indicator. This name MUST match the name of the indicator used in the satisfied/dissatisfied files.
2. Name\_albanian – this is the name of the indicator that will display in the visualization when the user has the Albanian language selected.
3. Name\_serbian – this is the name of the indicator that will display in the visualization when the user has the Serbian language selected.
4. Description\_albanian – This is the description of the indicator that will display on the Indicators page when viewed in Albanian.
5. Description\_serbian – This is the description of the indicator that will display on the Indicators page when viewed in Serbian.

*If you are modifying an existing indicator, please remember to update the existing row in the list rather than creating a new row.*

##### New/Modified Problems

If new problems have been added or existing problems modified, then those changes will need to be reflected in the lists file used by the visualization.

In this file you will find the three standard lists used by the application. For this section the tab of concern is the ‘problems’ tab.

The data in this tab reflects the list of all indicators in the visualization and contains the following columns:

1. Index – this column contains the *internal* name of the problem. This name MUST match the name of the problem used in the satisfied/dissatisfied files.
2. Name\_albanian – this is the name of the problem that will display in the visualization when the user has the Albanian language selected.
3. Name\_serbian – this is the name of the problem that will display in the visualization when the user has the Serbian language selected.
4. Description\_albanian – This is the description of the problem that will display on the Problems page when viewed in Albanian.
5. Description\_serbian – This is the description of the problem that will display on the Problems page when viewed in Serbian.

*If you are modifying an existing problem, please remember to update the existing row in the list rather than creating a new row.*

## Converting the Files

After all the required files have been updated as described above – the next step is to run a command from the command line that will convert all those files to the formats required for the visualization. In order to do this, you will need Python 3.1 or higher installed on your machine.

The command that will be run is calling a script that will in turn run a custom built Python class called MosaicData. This command will specify for the script

Run the python script which will recreate all the files as needed for the application.

### Pushing the Changes Online

Push to master

Push to gh-pages