Silas Macharia

Visualization of Financial Data

Study Based on Kenyan Government Annual Budgets

Helsinki Metropolia University of Applied Sciences

Bachelor of Engineering

Degree Programme

Thesis

27 November 2015

Helsinki Metropolia University of Applied Sciences

Author(s) Title Number of Pages Date	Silas Macharia Visualization of budget data: Study based on Kenyan government annual budgets 38 pages 4 December 2015
Degree	Bachelor of Engineering
Degree Programme	Media Engineering
Specialization option	Hybrid Media and Printing Technology
Instructor	Aarne Klemetti, Senior Lecturer

The purpose of this thesis is to provide insight into the importance of visualization of spending and budget data and to find the best techniques and methods that can be used to effectively visualize financial data. The Kenyan government budget data has been used in the study to illustrate the benefits of visualization of financial data in a way that makes it easy to understand.

The data used in the study focuses only on five government ministries showing budget allocation to the sectors in six financial budget years. Different methods have been used in the study in order to demonstrate different perspectives of visualizing budget data. Also examples of applications and projects done by different organizations and people about spending and budget data have been discussed.

The radial bubble tree, the tree map and combination charts have been used in the study to illustrate the different ways in which budget and spending data can be visualized to bring out differences, to show trends in spending over the years or to enable users to easily see where too much or very little has been spent. This makes it interesting and easy for a/the target audience to understand the meaning of financial budget reports.

Keywords	budget data visualization, Kenyan budget visualization

Contents

List of Abbreviations

SVG	Scalable Vector Graphics
HTML	Hypertext Markup Language
SQL	Structured Query Language
API	Application Programming Interface
VML	Vector Markup Language
DOM	Document Object Model
JSON	JavaScript Object Notation
CSS	Cascading Style Sheets
PDF	Portable Document Format
JSAPI	Java Servlet Application Programming Interface

1 Introduction

Pictures were used as a tool of communication for a long time before the inception of written language. A single picture can communicate a lot of information and can be more easily understood than comparable text based data. Pictures are also independent of any particular language since people with no common language may all understand what a picture represents.

The world is full of data and everywhere that people go they are surrounded by it. This information can be mined easily and shaped to provide different perspectives about the data and new insights. It might take substantial time and effort in order to understand patterns, trends and correlations that are usually related with text-based data. However, it is easy to distinguish a change in shape, color or line length without significant processing effort.

Effective visual communication of data helps people to understand large volumes of data in an easy and effective way. This thesis aims to demonstrate the importance of visualization of financial data basing the study on the Kenyan government annual financial budget for a five-year period between the years 2010/2011 and 2015/2016.

2 What Is Data Visualization?

Data visualization is communicating information clearly and effectively to users by the use of information graphics in order to help them easily understand and analyze data which otherwise might be hard to comprehend. This is important since visualized data helps to communicate complex information in an easy and effective way. [1, 1-3.]

It is often hard for professionals when met with large amount of data that is dull, abstract and hard to understand. These data can be meaningless without a way to organize and present important findings. In order to bring data live to people so that they can explore it and use it, it has to be visualized so that it can be understood better. [2.]

Newspapers and magazines use visualization to enrich their stories, perhaps to show the link between poverty and lack of education, where sales growing are and what the driving growth is? Another thing to illustrate is the characteristics of customers using different services. Bloggers also use it to start discussions with data all across the web. [2.]

Visualization is very useful when trying to extract and communicate important information from a big amount of data. The use of graphics instead of text to indicate trends and patterns from a massive amount of data is very effective. It is a challenge for designers when choosing a visualization tool to be used to represent data accurately. Often a number of factors have to be considered before choosing a particular tool to visualize data. The cost of visualization, type of data to be visualized and effectiveness of the visualization tool are among a few factors to be taken into consideration. [2.]

3 Data Visualization Applications

Data presentation can be beautiful, elegant and descriptive and there are many conventional tools used every day to visualize data. In order to get the best visual representation of data, a person needs a tool that can capture the attention of the intended user, deliver the message clearly and also be flexible enough to tell the story [2].

Visualization tools allow a person one to organize and display data in a way that is easy for users to understand. A lot of tools exist today that can be used in visualizing data. For some of these tools a people have to pay to use them while others are open source and are free for use by everyone. Majority of the open source tools are developed using JavaScript libraries and are easy to use in visualizing data for both skilled developers and people with less coding skill. In this study five popular JavaScript based tools will be discussed in brief.

D3.js

D3 refers to Data-Driven Documents. It is a JavaScript library used for visualization of data. D3 was developed using functional style JavaScript and allows reuse of code through a varied group of plugins and components. It uses HTML, CSS and SVG graphics to visualize data and works well across all modern browsers with full capabilities. It works fast, supports huge datasets and dynamic behaviors for animation and interaction. [3.]

D3.js code is simple and it has a well-written API and a very responsive community. It supports large data sets and it has a fast performance [3]. There are other libraries that allow the creation of graphic representation of data easily like charts. However, D3 gives someone the flexibility of representing data from many different angles enabling visualization of different kinds of data and has a huge variety of responsive and interesting chart types. [4, 1-7.]

To demonstrate well how to create graphs with D3.js, the following example gives a clear picture. First include a D3 library in the document head of the file as shown by the following code example.

```
<script type="text/javascript" src="d3/d3.v3.js"</script>
```

Listing 1: Reference to D3 library. Copied from Data-Driven Documents 2014 [3].

D3 can be used to create a new DOM element and also SVG objects for representation of data. The illustration below shows how this can be done.

```
<script type="text/javascript">
d3.select("body").append("p").text("New paragraph!");
</script>
```

Listing 2: D3 DOM element creation. Copied from Data-Driven Documents 2014 [3].

In the above code, the select method selects a single DOM element (body) using a CSS syntax selector. It then creates a new p element (paragraph) and appends it to the selection. Finally it inserts the text “New paragraph” to the newly created paragraph. The D3 chaining method allows several actions to be executed in a single line of code as demonstrated below. [3, 5.]

```
d3.select("body")
        .append("p")
        .data(dataset)
        .enter()
        .append("p")
        .text("New paragraph!");
```

Listing 3: D3 chaining method. Copied from Data-Driven Documents 2014 [3].

The element (***.data())*** joins an array of data to the current selection. (***.enter()***) looks at the DOM and then adds the data being handed over to it. If there is no more data than corresponding DOM elements, D3 creates placeholder for missing elements. It also hands off reference to this new placeholder to the next step in chain. (***.append()***) binds element ***(“p”)*** to the place holder selection created by ***enter()*** in the DOM. Data can then be loaded to the created element in the DOM and is then ready to be used. [3.]

D3 accepts data as an array of numbers, strings or objects. It can also take CVS and JSON files, once data is loaded and bound to the created DOM elements; it is then visualized in various ways that bring out the important facts of the data to the intended audience. [3.]

3.1 JavaScript InfoVis Toolkit

JavaScript InfoVis Toolkit is an open source tool that allows developers to build rich interactive visualization. One can visualize data in many advanced graphical forms such as bar charts, RGraphs, ForceDirected visualization, TreeMaps, SpaceTree and Sunburst visualization. There is an example figure below of a Sunburst of a Directory Tree visualization made using InfoVis ToolKit. [6.]

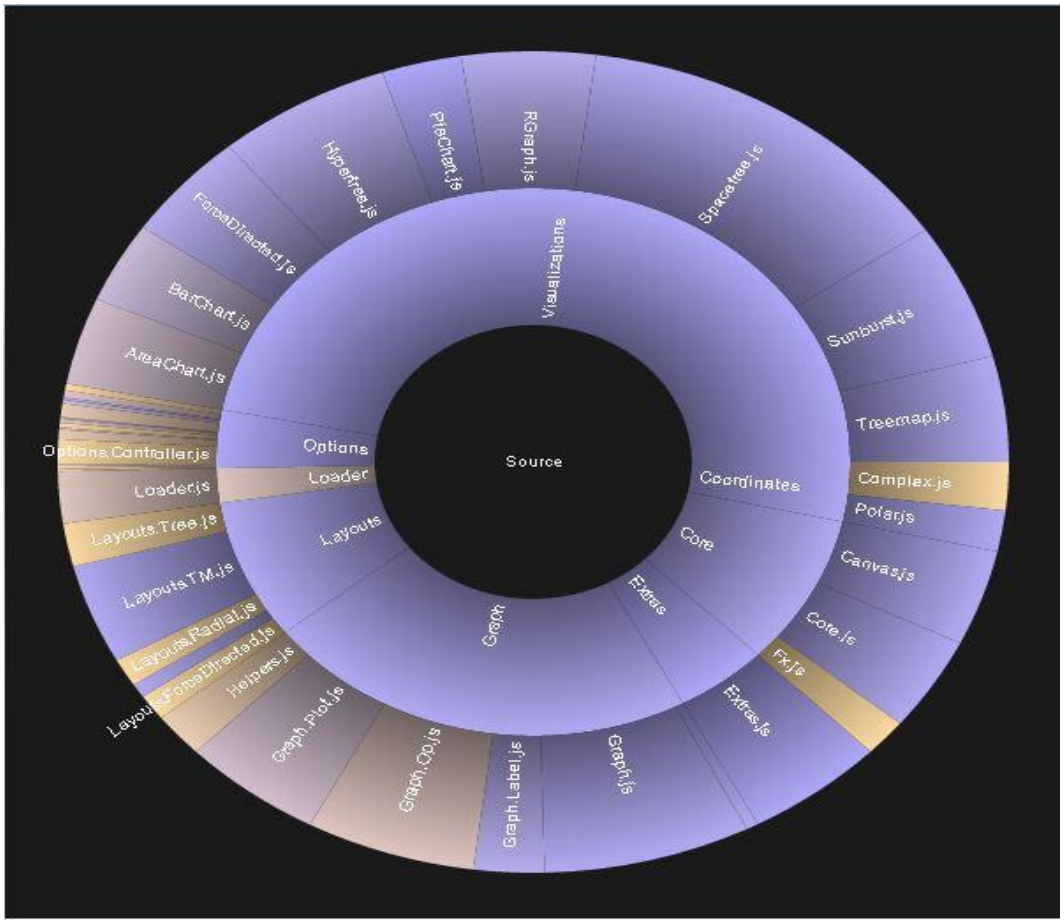

Figure 1. File system visualization. Copied from Belmonte (2015) [6].

An easy way to learn using the tool is to look at code of example visualization made using this platform. It is easy to create visualization using this tool since a person has to define the parameters of the graph and source of data to create a simple chart. In the year 2010 it was however acquired by Sencha Labs foundations but it is still an open source tool. [6.]

3.2 JQuery Sparklines

JQuery Sparklines is a plug-in based on jQuery and it is used in generating Sparkline, area charts and line charts. It is used mainly to display variation in measurement over time visualization like a change of temperature over time or variation of stock prices and has support for older browsers. To add Sparklines into a webpage, it is important to define four elements: jQuery JavaScript library, JQuery Sparkline library, inline tag to display Sparkline within the page and Sparkline function call to display the chart. [7.]

```
//jQuery JavaScript library

<script type="text/javascript" src="jquery-1.7.2.js"></script>

//JQuery.Sparkline.JS

<script type="text/javascript" src="jquery.sparkline.js">

//inline tag to display Sparkline within the page

Inline Sparkline: <span class="inlinesparkline">1,4,4,7,5,9,10
</span>.

//Sparkline function call to display the chart

$(function() {

            $('.inlinesparkline').sparkline();

                        }
```

Listing 4: JQuery Sparklines important elements. Copied from Garcia Belmonte (2015) [7].

Sparklines are designed in a way that they fits alongside a line of text to indicate a pattern or trend and may lack an attribute associated with other kind of charts [7].

3.3 Google Chart Tools

Google Chart Tools can be used to generate different types of charts from simple line charts to more complex shapes of data representation. It is an advanced version of image chart API [6]. It has many types of chart types available to choose from and it uses HTML5 and SVG to render charts in order to ensure cross browser compatibility and also cross platform compatibility to iPhones and Android. Older versions of Internet Explorer are supported using VML. [8.]

In order to understand how visualization with google charts tools works, a study of example pie chart shown below will be used to demonstrate how simple it is to visualize using google chart tools.

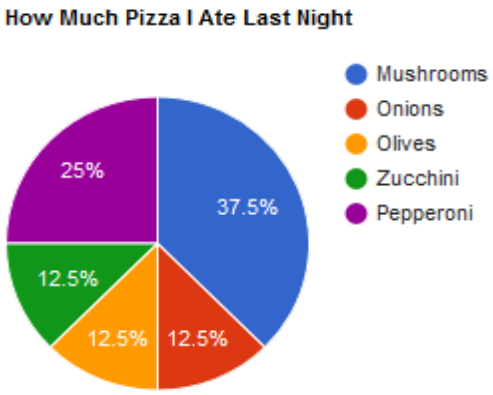

Figure 2: Example of a simple pie chat. Copied from Google Developers (2015) [8].

The first stage is to load Google JSAPI library loader and use it to load the visualization library as shown below.

```
<!--Load the AJAX API-->
<script type="text/javascript" src="https://www.google.com/jsapi">
</script>
<script type="text/javascript">

// Load the Visualization API and the piechart package.
google.load('visualization', '1.0', {'packages':['corechart']});

// Set a callback to run when the Google Visualization API is loaded.
google.setOnLoadCallback(drawChart);
```

Listing 5: Google JSAPI loader code. Copied from Google Developers 2015 [8].

Google JSAPI library is loaded using the first script and the second script loads google visualization and chart libraries. It basically contains the chart code. In this line a person also defines which version of Google charts to load by defining it as shown by the following code example. [8.]

google.load('visualization', '1.0', {'packages':['corechart']});

Listing 6: Google Chart version definition. Copied from Google Developers 2015 [8].

Data is contained in a JavaScript class called google.visualization.DataTable, the following example code illustrates how it can be done.

```
// Create the data table.
var data = new google.visualization.DataTable();
data.addColumn('string', 'Topping');
data.addColumn('number', 'Slices');
data.addRows([
     ['Mushrooms', 3],
     ['Onions', 1],
     ['Olives', 1],
     ['Zucchini', 1],
     ['Pepperoni', 2]
     ]);
```

Listing 7: Google visualization data table. Copied from Google Developers (2015) [8].

The data source can be from a web page, database or any data source supporting the chart tools data source protocol that includes SQL type of query language and can be implemented by Google Spreadsheets or Google Fusion Tables. Google chart tool provides an easy way to switch between different chart types in order to find the ideal visualization since data is loaded using the DataTable class that offers techniques of sorting through the data. [8.] For instance in order to convert from a pie chart to a bar chart a simple change of code shown below will make the changes.

```
// creates a pie chart.
google.visualization.PieChart

// creates a bar chart.
google.visualization.BarChart
```

Listing 8: Google Chart type definition. Copied from Google Developers (2015) [8].

The chart can be customized by defining the parameters like colors and background fill as show by below code

```
// Set chart options
var options = {'title':'How Much Pizza I Ate Last Night',
               'width':400,
               'height':300};
```

Listing 9: Google Chart parameters definition. Copied from Google Developers (2015) [8].

The last step is to draw the chart by first instantiating an instance of the chart class that a person wants to use and then define ***draw()*** at the end of the line [8]. A complete code compiling all the above illustration is shown below.

```
<html>
<head>
<!--Load the AJAX API. Do this only once per web page! -->
<script type=”text/javascript” src=”https://www.google.com/jsapi”>
</script>
<script type=”text/javascript”>

// Load the Visualization API and the piechart package.
google.load(‘visualization’, ‘1’, {‘packages’:[‘corechart’]});

// Set a callback to run when the Google Visualization API is loaded.
google.setOnLoadCallback(drawChart);

// Callback that creates and populates a data table,
// instantiates the pie chart, passes in the data and
// draws it.
function drawChart() {

// create the data table.
var data = new google.visualization.DataTable();
    data.addColumn(‘string’, ‘Topping’);
    data.addColumn(‘number’, ‘Slices’);
    data.addRows([
        [‘Mushrooms’, 3],
        [‘Onions’, 1],
        [‘Olives’, 1],
        [‘Zucchini’, 1],
        [‘Pepperoni’, 2]
      ]);
// Set chart options
var options = {‘title’:’How Much Pizza I Ate Last Night’,
               ‘width’:400,
               ‘height’:300};

      // Instantiate and draw our chart, passing in some options.
var   chart   =   new   google.visualization.PieChart(document.getEle-
mentById(‘chart_div’));
chart.draw(data, options);
  }
</script>
</head>

<body>
    <!--Div that will hold the pie chart-->
    <div id=”chart_div” style=”width:400; height:300”></div>
  </body>
</html>
```

Listing 10: Google Chart Complete code for a simple pie chart. Copied from Google Developers (2015) [8].

3.4 Highcharts JS

Highcharts JS is a charting tool developed using JavaScript. It supports many chart types and has its own available themes that can be used for design [6]. It uses an external JavaScript framework for animation, DOM manipulation and event handling and has a built-in support for jQuery [9].

Even though Highcharts is developed using a JavaScript framework library, its application does not totally depend on one particular framework. It can be integrated into Prototype, MooTools and jQuery JavaScript frameworks. Highcharts requires two files to run, highcharts.js and either jQuery, MooTools or Prototype. Below is illustration of how it can be defined [9.]

```
/ Highcharts with jQuery
<script                                          src=”http://ajax.goog-
leapis.com/ajax/libs/jquery/1.8.2/jquery.min.js”></script>
<script src=”http://code.highcharts.com/highcharts.js”></script>

// Highcharts Standalone Framework

<script src=”http://code.highcharts.com/adapters/standalone-frame-
work.js”></script>
```

Listing 11: Highcharts JS file library definition. Copied from Highcharts developers (2015) [9].

Highcharts can be used in visualizing data in various different chart forms, below is an example chart made using Highcharts JS tool

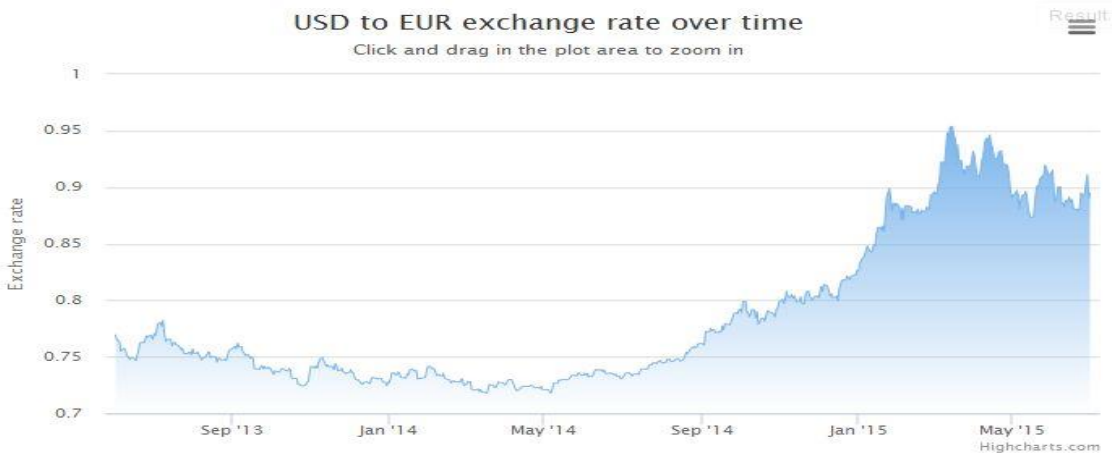

Figure 3: US Dollars to Euro exchange rate over time. Copied from Highcharts Developers (2015) [9].

Highcharts has been developed using native browser technologies and does not require server side plugins in order to run. It can also be run locally from a file system because rendering is done locally in a browser. The use of Highcharts for non-commercial purposes is free. However, it cost money to use it for commercial visualization development.

4 Data Visualization Techniques

4.1 Tables

Data representation utilizing tables takes the form of columns and rows. Most commonly it is used in accumulating and organizing raw data and also for data representation. It enables users to check through the data and make comparison of the values. This method of data representation is the most commonly utilized in data management and it is also flexible in representing various types of data. [10.]

Microsoft Excel uses table as its primary way of representing data. This data can be visualized in various charts tools available in the Excel program [10]. The example below is a data table about the Kenya budget allocation to five sectors of the economy from the year 2010/2011 to the year 2013/2014.

Sectors	2010/11 Total Budget Estimates	2011/12 Total Budget Estimates	2012/13 Total Budget Estimates	2013/14 Total Budget Estimates
Infrastructure + Energy	135,520,923,694	183,987,366,088	268,100,000,000	220,800,000,000
Agriculture	23,230,799,372	19,565,914,807	22,099,000,000	38,100,000,000
Health	31,564,229,463	31,684,477,684	85,028,900,000	34,750,000,000
Security	113,278,607,371.00	119,056,679,790.00	83,500,000,000	74,420,000,000
Education	193,108,581,065	202,570,176,743	233,100,000,000	273,700,000,000
Total	496,703,140,965	556,864,615,112	691,827,900,000	641,770,000,000

Figure 4: Kenyan budget estimates to 5 sectors of the economy between the years 2010/2011 and 2013/2014, amounts in Kenyan shilling. Data from the Institute of Economic Affairs in Kenya [11].

When using a table for data representation care should be taken to limit data in the table to only what is required for the study. Avoid having many rows or columns that might confuse the reader, ensure that the table can be understood easily without much further explanations and design it neatly for it to be appealing to the readers. [10.]

4.2 Bar Chart

Bar Chart is a method of representing a set of definite data. Data is visualized in a number of arranged bars each representing a particular category. The height of every bar represents a specific value indicated by the vertical length of the bar chart. [10]

Color-coding the bars makes it easy for users to distinguish between the various bars and also enables them to understand the data representation with ease [10]. The simple bar chart below is a representation of the Kenyan government financial budget estimates indicating allocation to five sectors of the economy for the years 2015-2016.

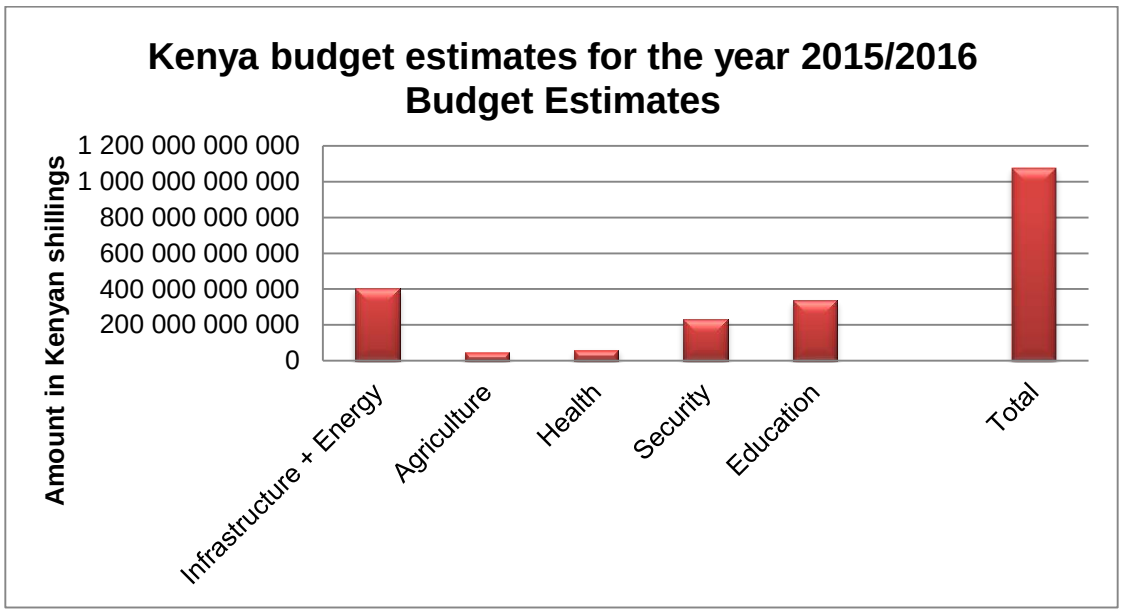

Figure 5: Kenyan budget estimates to 5 sectors of the economy for the years 2015-2016. Data from the Institute of Economic Affairs in Kenya [11].

In the above simple bar chart, each individual bar represents a unique category. There are also other types of bar charts namely horizontal bar chart, stacked bar chart and grouped charts. A histogram is similar to a bar chart in look but differs by the fact that it represents a continuous set of data rather than an independent data category. [10.]

4.3 Line Chart

Line chart is a good visualization tool where the intention is to show trend over a time period. A good example would be to show how the sales of a particular product change over time or fluctuation of stock prices in the market over a specific period of time [10].

It is most commonly used to show more of rate of change of a value over a period of time than the measure that has changed [10]. The line chart below is an example comparison of expenditure in infrastructure and security by the Kenyan government for a six year period.

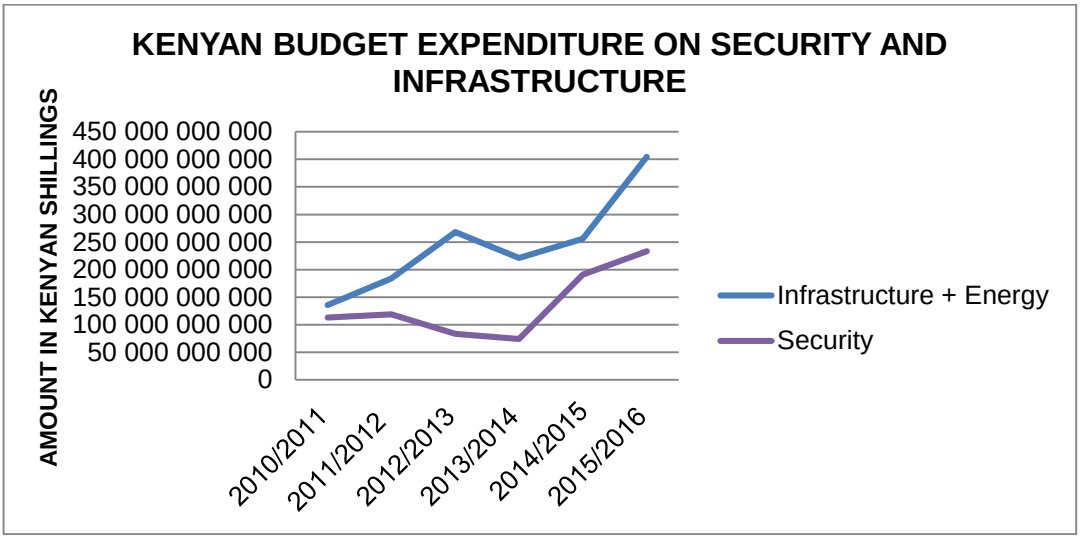

Figure 6: Example line chart indicating change of price for two items. Data from the Institute of Economic Affairs in Kenya [11].

Line charts and scatter plots are very similar in representation. However, a line chart has the various points connected with a continuous line. Line charts are used in establishing a connection between two sets of independent values. If a trend is established for an item in a dataset, then prediction of future results can be made without having real data at hand. [10.]

4.4 Combination Chart

Combination Chart is a type of visualization where two types of charts have been combined in a single visualization. An example of this would be combination of a bar chart and a line chart. This kind of visualization is good when making a comparison of data values in two different categories as illustrated by the below figure. [10.]

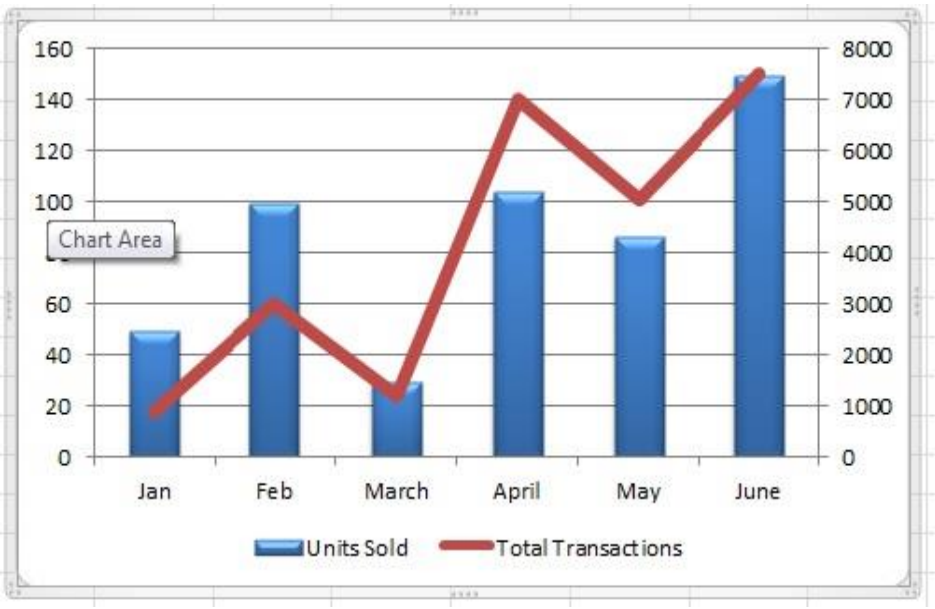

Figure 7: Combined chart showing the number of product units sold by a company and the total revenue generated [12].

The above way of data representation makes it possible to visualize different kinds of datasets at the same time without the need of having separate charts.

4.5 Pie Chart

A pie chart is a type of graph that has a circle divided into sectors. Each sector represents a certain percentage of the whole graph. The size of a particular sector is proportional to the amount it represents of the whole amount on the pie chart. A pie chart is mostly used to show comparison of values. [10.] One example area where it is used is to show financial budget allocation to different departments as indicated by the figure below.

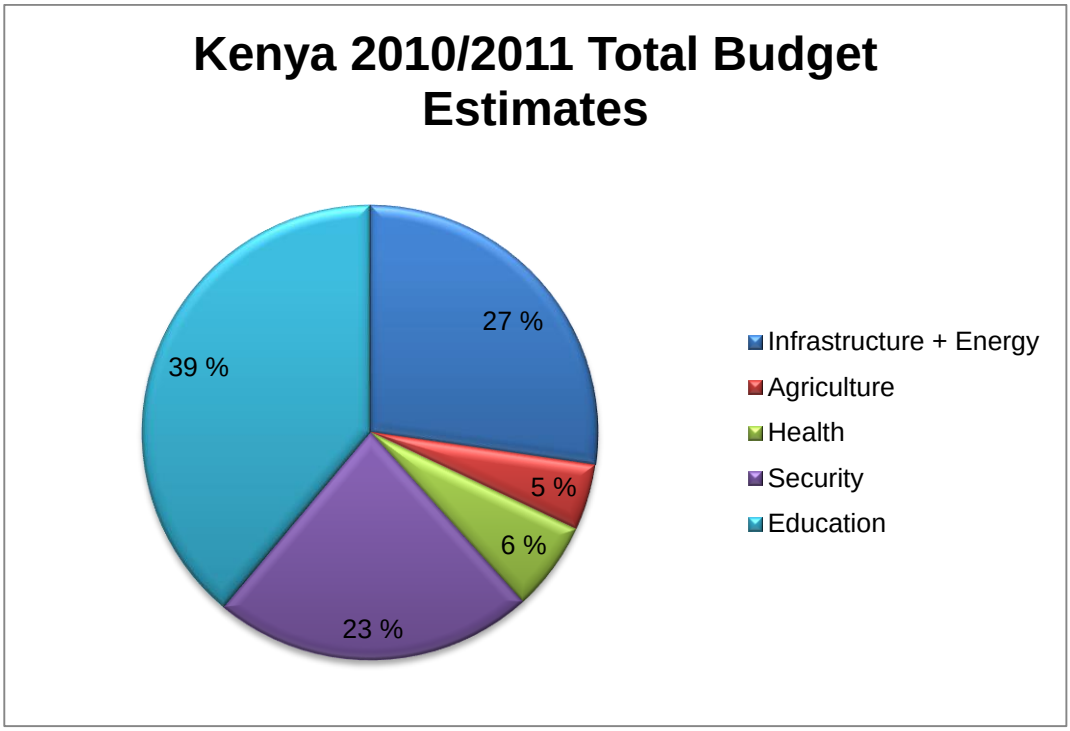

Figure 8: Kenyan financial budget allocation to a number of departments for the years 2010-2011. Data from the Institute of Economic Affairs in Kenya [11].

It is often easy to read information represented in the pie chart because of its simplicity in design. Color-coding the various sectors makes the illustration clear, and it is easy to distinguish the various sectors. However, when representing a large amount of data a pie chart many not be the best suitable method because many small segments might confuse the reader of the pie chart. [10.]

4.6 Scatter Plot

Scatter Plot is a type of graph where a variable is plotted along two axes with the aim of showing a relationship between two sets of data that share the same event. Data is graphed as a number of points in the graph with each point representing a value. The use of a scatter plot is demonstrated below. Figure 9 shows how the prices and quality of a product compare and affect one another other. [10.]

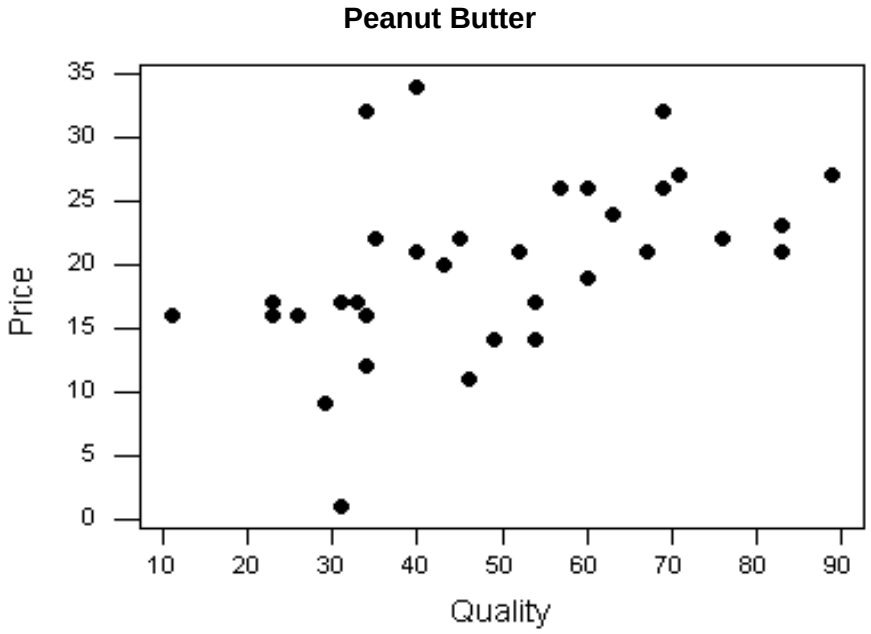

Figure 9: Peanut butter comparison study of price versus quality. Copied from Inter-Math (2015) [13].

The points can be colour coded to increase contrast and make viewing and interpretation easy. The relationship between the two variables is known as their correlation. A three-dimensional scatter plot can be used to visualize the relationship between three variables. [10.]

4.7 Tree map

Tree map is a method of visualizing hierarchical data in the form of rectangles. It enables users to look at the whole picture of data representation as a whole and to make comparison of the nodes and sub-nodes at various levels. Tree Maps can be created in more than one way but the interface of the Tree Map visualizations looks similar. However, some are interactive while others are static. [10]. The figure below is an example tree map showing gross domestic product distribution of the top 15 countries in the world.

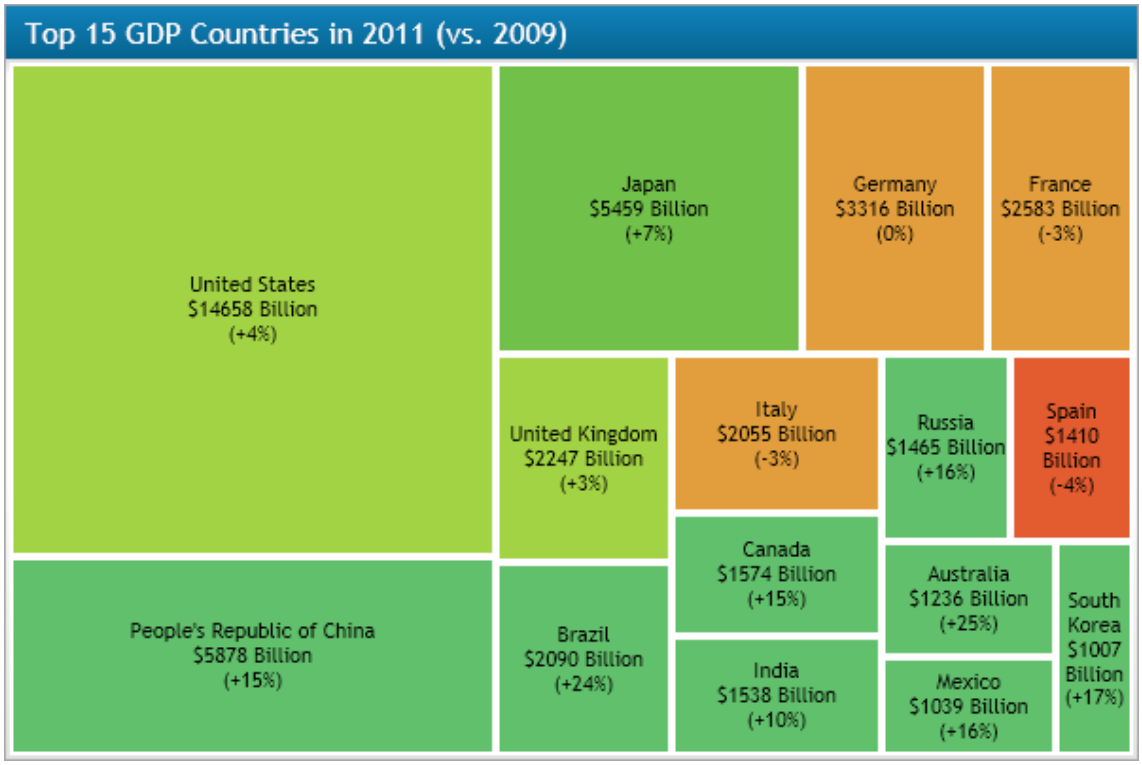

Figure 10: GDP (gross domestic product) distribution of the top 15 countries in 2011. Copied from Dundas Data Visualization, Inc. (2015) [14].

Each rectangle represents a category that can also be divided into sub-categories. Colour can be used to emphasize intensity of the segment. Similarity in colour and size of the rectangles can guide users into identifying a pattern in the data represented. [10.]

5 Visualization of Financial Budget Data

Financial data is defined as a set of price and volume values for a certain quantity of income [15]. There are my charting tools used in visualization of financial data. These tools can be used to generate various types of charts in order to find trend and relation in the data. Charting tools like pie charts and bar graphs use features such as size and color to show trends and distinguish sets of data visually. [16, 8.]

Visualizing financial data puts the story in front of everyone's eyes making it easy to digest and understand. Development of new visualization tools has enabled dynamic visualization to be achieved through dashboards and interactive charts. [17, 35.]

The financial budget for many countries is presented in printed format as a booklet or PDF document in electronic format, but because of the complexity and huge financial numbers involved it takes time for one to read and understand. Visualizing the same data in the form of charts and perhaps with some interactivity makes it easy for an average person to understand how the money has been budgeted. [17, page 35]

The web provides a platform for sharing data easily to many people. The purpose of visualizing budget information is to generate interest of the reader making it easy to know how the tax payer money has been spent, explain the work that the government performs and enable the citizens to know functions that the government has prioritized in terms of spending.

6 Illustrations of Budget Visualization

6.1 OpenSpending

Open Spending is an open platform where people can upload financial data on spending and visualize using built-in tools. The platform is a JavaScript framework and it is a community run project with open-source code and open data. Visualization done on this application can be embedded to other websites using an available widget [18].

There are more than 28 million transactions from 1071 datasets over 75 nations on the website. One application in Open Spending that shows visualization of financial data is named **where does my money go?** It is a visualization of how the money of taxpayers from the United Kingdom has been budgeted by the government. The visualization has been done in two ways, namely **country and regional analysis** and **the daily bread** [18].

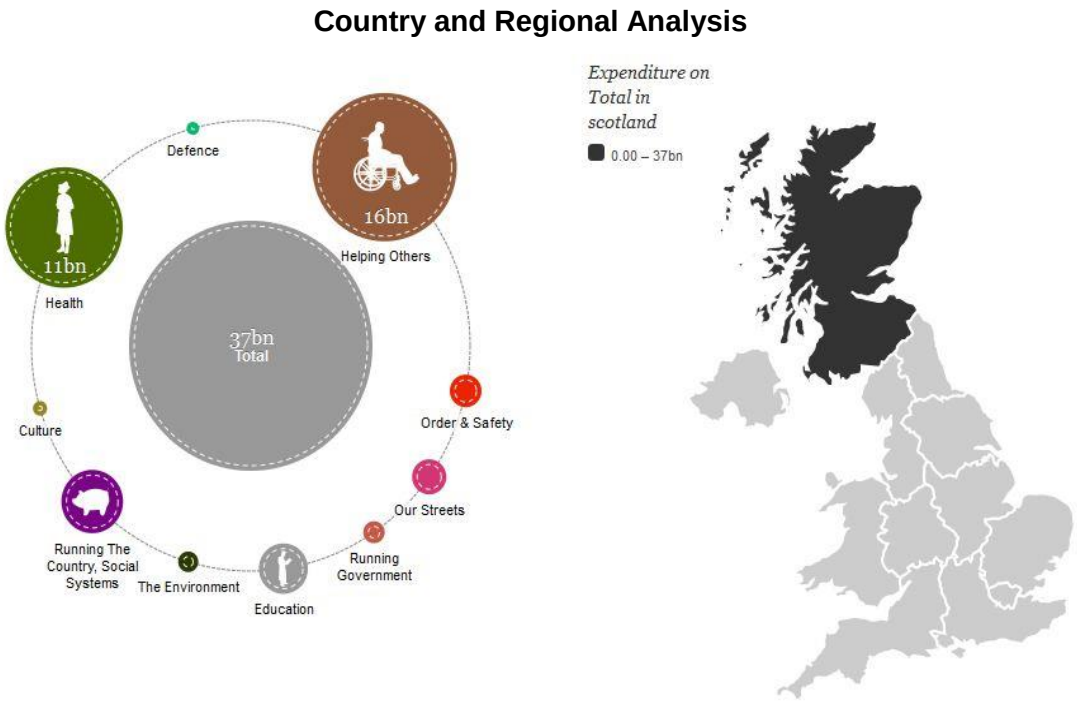

Figure 11: Bubble Tree Map representation of the United Kingdom government spending for the years 2004-2010. Copied from Where Does My Money Go (UK) (2015) [18].

One can compare how various regions in the United Kingdom have done financial budgeting by clicking on various regions on the map. Its interactivity generates user interest to understand how budgeting has been done by the regional government. [18.]

Figure 12: The Daily Bread Visualization of the United Kingdom government spending of tax to various ministries. Copied from Where Does My Money Go (UK) (2015) [18].

In **the daily bread** application, dragging the shown slider in figure nine adjusts the salary scale which then displays how much tax is levied on that salary level annually. It also show a breakdown of the amount of tax paid on a single day to various government functions on the basis of that salary scale. Its simplicity and interactivity makes it easy for a person to understand tax levied and expenditure by the British government functions. [18.]

6.2 Meie Raha

The term "**Meie Raha"** is an Estonian word for "**Our Money**". Meie raha application visualizes the Estonian financial budget spending. It was created by a team of volunteers at a local hackathon known as Garage48. The application has two sides of the equation, the income group that mainly consists of taxes and the expenditure group. The financial size of the budget item is directly proportional to the size of the bubble. The picture below is an illustration of what the visualization looks like [19.]

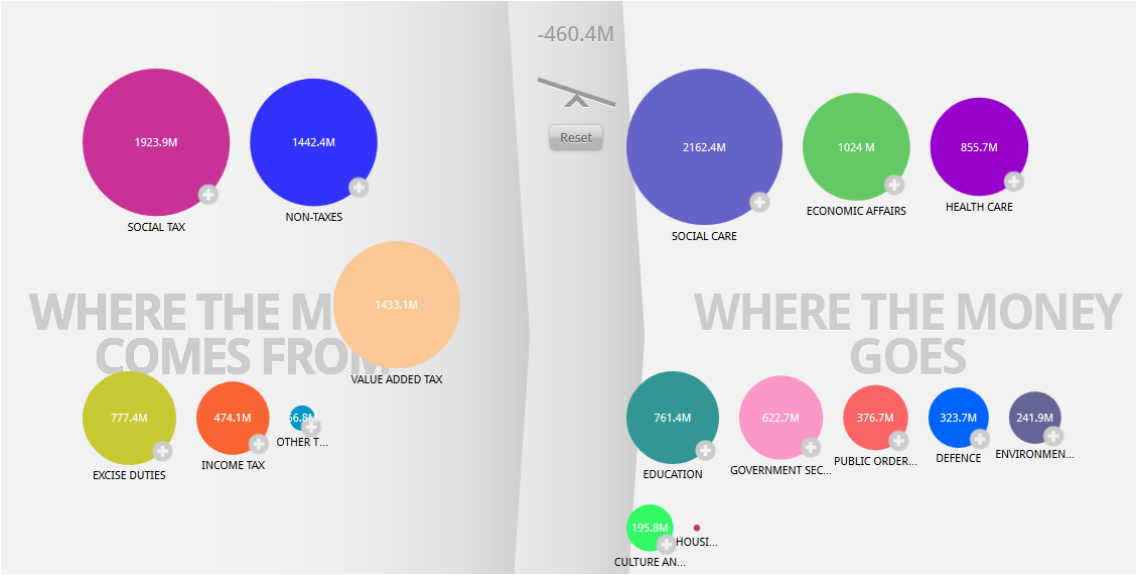

Figure 13: Screen capture of Meie raha visualization of the Estonian government financial budget (2015) [19].

The application can be manipulated by dragging the bubble from one position to another resulting in different balance of the budget. However, it is important to remember that a budget has to be balanced and so a change in the expenditure side has to be balanced on the income side. It is also possible to expand and see the content of a bubble by clicking on it. [19.]

6.3 Think Tank Fund - Open Society Foundations

Think tanks are organizations that are focused on data analysis in Central and Eastern Europe using numbers and context to make a difference. Their aim is to make it easy for users to understand valuable data by creating interactive visualization. This helps both citizens and policy makers to better understand data and guide them in future decision-making. [20.]

Think Tank Fund has supported projects since 2010 that are aimed at providing new perspective of how facts are viewed in a certain policy field. The picture below is a screen capture of projects that were supported by open society foundation between the years 2010 and 2013. [20.]

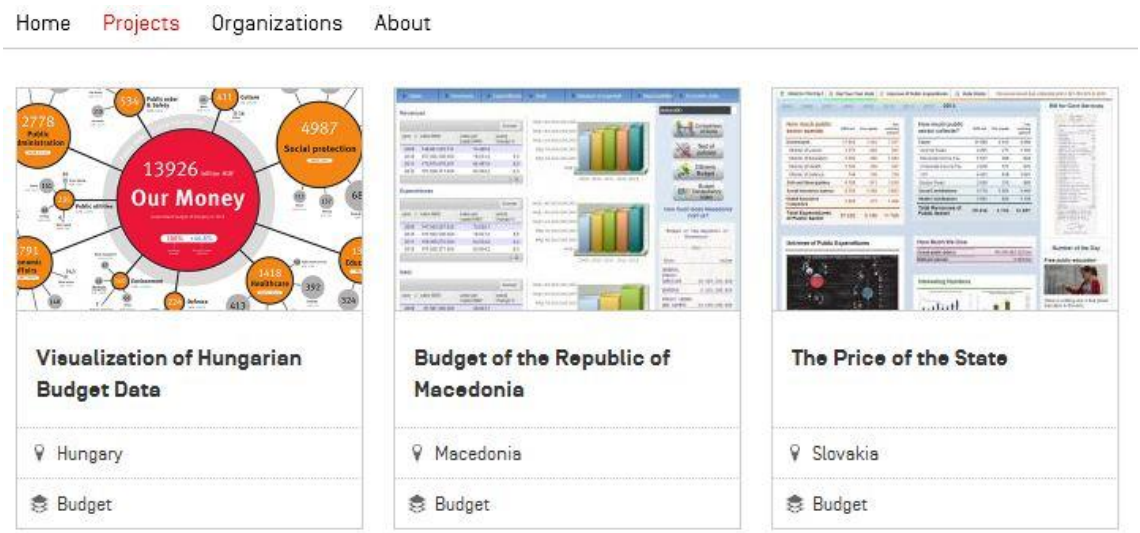

Figure 14: Screen capture of financial visualizations supported by the Open Society Foundation. Copied from Think Thank projects (2015) [20].

Think Tank Fund has 16 supported projects that are divided into three named categories: accountability, crime and budget. The visualization makes the understanding of complex data simple. [20.]

6.4 Four Ways to Slice Obama's 2013 Budget Proposal

Four Ways to Slice Obama's 2013 Budget is a budget data visualization application created by Shan Cater with the help of others for the New York Times. Its aim is to provide an interactive way to explore the Obama budget proposal for the year 2013 and also to compare how he changed it with respect to the previous years. It uses bubbles to represent data, and it has smooth transitions between each of the four distinct views making it easy for users to explore it. The picture below shows how the visualization has been made. [21.]

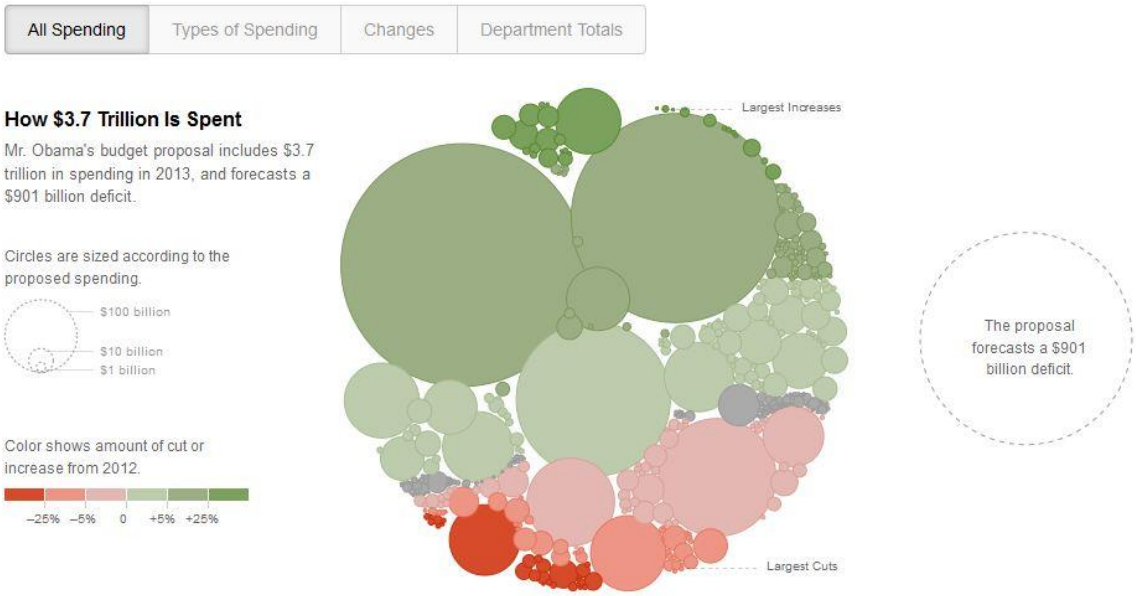

Figure 15: Screen capture of New York Times visualization of Obama 2013 budget. Copied from the New York Times (2015) [21].

The size of the bubble is directly proportional to the amount of money allocation in US dollars. The colour green means more money allocation and red means less money allocation to the various governments departments. [21.]

6.5 Infographics

Information graphics are a visual representation of data designed to convey information clearly and quickly. They are used for many different kinds of information visualization. However, their use for representation of financial budget data will be discussed in this study. [22].

Infographics have been used to visualize financial and budget data with very good results. The demonocracy.info website offers a number of detailed economic infographics for a wide area of financial topics. The Canadian financial budget for the years 2012-2013 has also been visualized using infographics in a very interesting manner as shown in the figure below. [22].

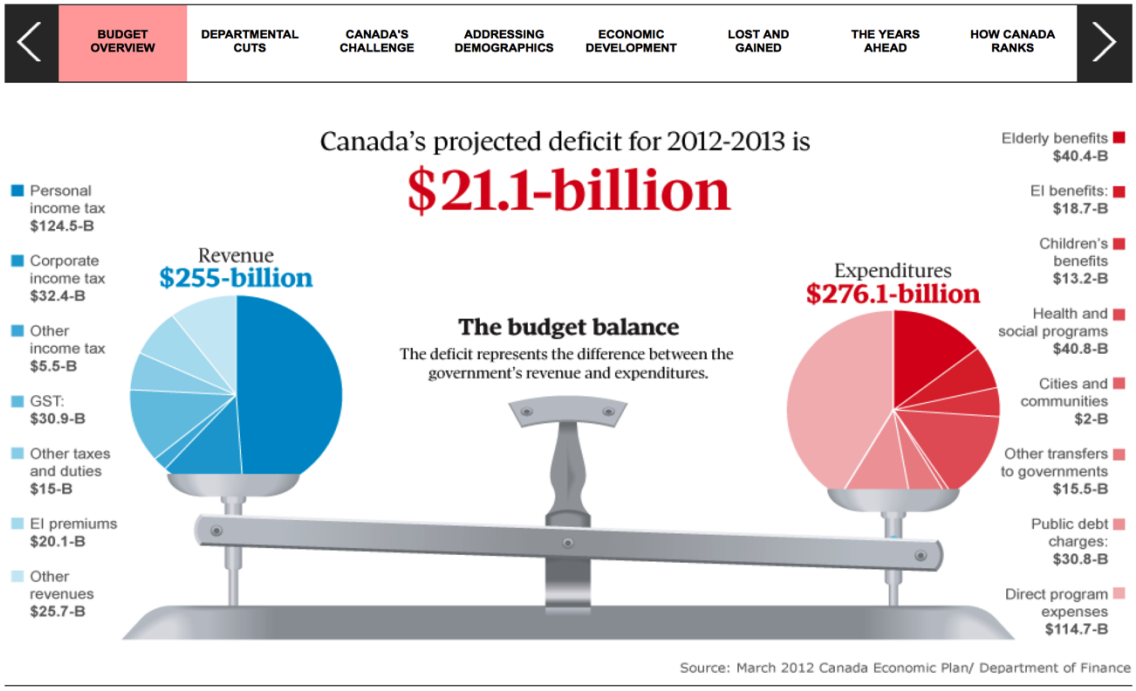

Figure 16: Screen capture of Canada 2012 budget visualization. Copied from the globe and mail (2015) [22].

The slider at the top of the infographics provides a way to view the budget from different perspectives. The blue and red colour indicates revenue and expenditure respectively [22].

6.6 Kenyan Budget Explorer

Kenyan budget explorer is a budget visualization developed by the Twaweza organization. It has been designed in the form of a dashboard and it visualizes nine Kenyan budget proposals between the years 2002-2003 and 2011-2013. Data can be visualized in the form of bubbles and bar charts. It has filters to focus the view and also the option of downloading the data. The figure below is how the visualization is a picture of the application user interface [23].

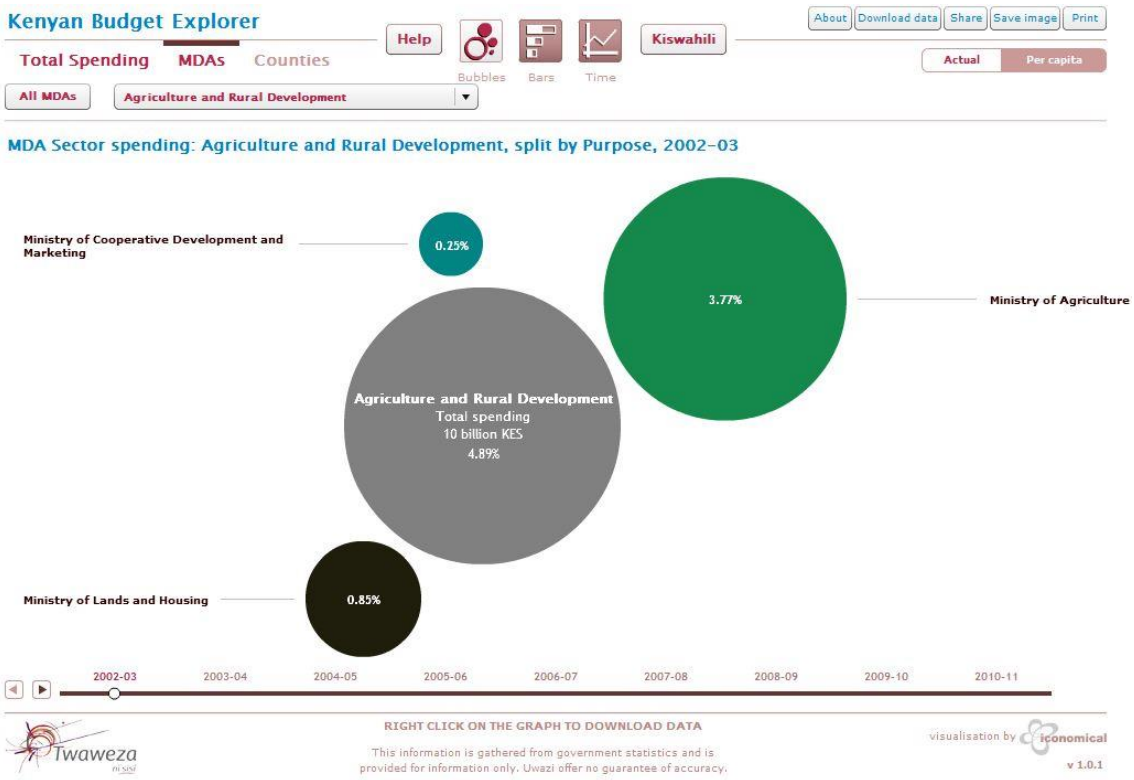

Figure 17: Screen capture of the Kenya Budget Explorer application 2015. Copied from the Twaweza organization (2015) [23].

The application visualizes budget estimates for nine consecutive years enabling one to see how the budget allocation has been changing over the years and possibly how to find trends in money allocation to various government ministries [23].

7 Visualizing Kenyan Financial Budget Data from 2010/2011 to 2015/2016

The process of data visualization used in the study of Kenya financial budget has been modelled as the below flow chat diagram shows.

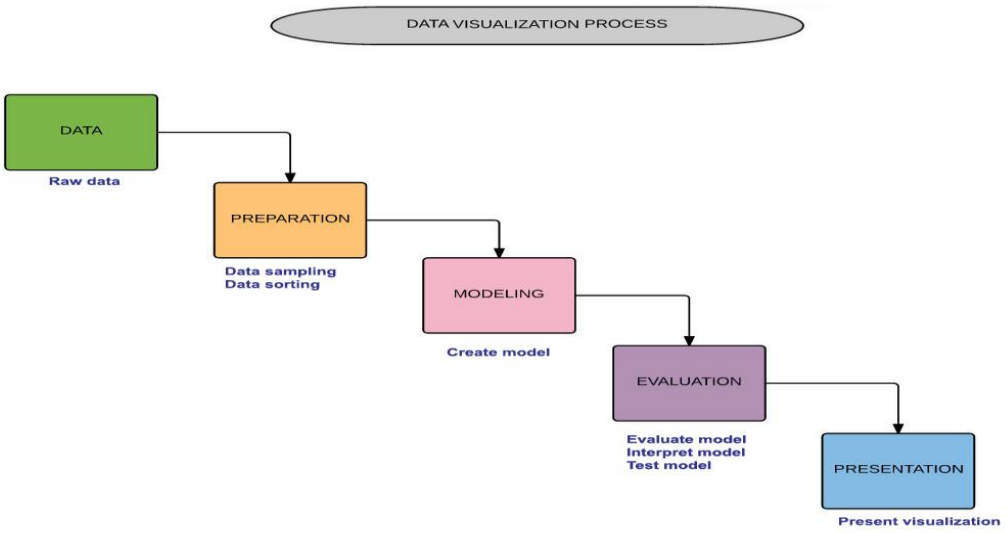

Figure 18: Data visualization model process of the Kenyan budget data.

The process of data visualization used in this study was structured in a number of steps as indicated in the above flow chart. First, it was important to find a good source of data to use in the study and then sorted it and finally prepares it for use in the visualization.

A model was then created based on the collected data. The type of data determined the type of model used. The visualized model was evaluated and tested to see if the goal of the study had been achieved and also to make any correction if necessary. When all had been done then design and presentation of the final model was then done.

7.1 Data Source

The motivation of any data visualization is to communicate the meaning of volumes of data and make people understand the story behind the data in a simple way. Choosing a good source of data is an important prerequisite of any visualization. The data source of financial transactions and budget expenditures of many African governments is not easy to be accessed. However, the Kenyan government has made good effort of trying to make data of its budget expenditures more open to the general public [24].

The government of Kenya established a Kenyan open data portal in an effort to make public datasets more easily accessible to the public and also to try to be accountable to its citizens and to make them more proactively involved and informed. The portal holds more than 680 datasets from various sectors such as education, energy, environment and agriculture. Also many kinds of visualizations have been done on these datasets. These datasets can also be downloaded in various standard file formats. [24.]

The Kenyan government budget estimates for the dataset of the years 2015-2016 was downloaded from a website owned by the Institute of Economic Affairs (IEA - Kenya). It is a think tank organization that provides a platform for informed discussions with the intention of influencing public policy in Kenya [11]. I chose to use Kenyan government data because of the following reasons:

- To show comparison of financial budget allocation to the five government ministries over the chosen period of time.

- To show trends overtime if allocation to a particular ministry is progressively going up or down and to show what fluctuation has been in the ministries over the year's chosen.

- To demonstrate that a budget presentation can be made easy for people to understand by visualizing it in an interesting way. Normally Kenya budget presentations are in paper print or PDF files.

- To empower citizens by presenting information about the government budget expenditure.

7.2 Data Preparation

It is a good practice to prepare data before the visualization process starts. This will help the people doing the visualization to come up with effective results. Data preparation is a time consuming process because mostly it has to be done manually but it is important if a good visualization has to be achieved [26, 23-24].

The budget data used in creating the application for this study as shown in figure nineteen originally was in the PDF file format. Only what was required for the visualization was identified and selected. A Microsoft Excel file was used to sort through and prepare the data. Inconsistent entries, errors and any discrepancies were checked and corrected at this point. The currency in the budget estimates is in Kenyan shillings and the entire file document is in English.

7.3 Bubble Tree

The bubble tree chart type has been used in visualizing the Kenyan government budget estimates. Gregor Aisch and David McCandless originally developed it for visualizing spending data in the OpenSpending platform but now it is possible to create a visualization using the application independently. Bubble tree has been developed on top of JQuery and Raphael JS libraries. [27.]

Using this interactive platform creates interest in users to try to understand the budget and also help them comprehend the important aspects of the budget spending estimates. Reading and understanding the financial budget in the original PDF document takes more time and does not attract the interest of its readers.

Different colours have been used to bring out the contrast in the bubble tree. The icons in the bubbles help the user to easily identify the meaning of what the bubble represents. The little dotted white line on the inside of the bubbles indicates that it has child bubble and expands to other children by clicking on the icon. The picture below shows the bubble tree structure general look. [27.]

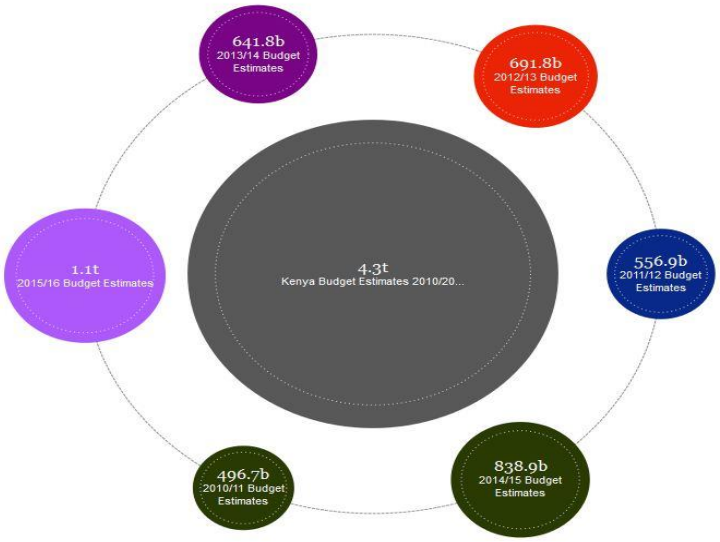

Figure 19: Kenyan financial budgets estimate for the years 2010-2016. Data from the Institute of Economic Affairs in Kenya [11].

The bubble tree structure makes it easy for users to understand the general overview of the budget proposal without the need of reading a lot of pages of financial statements in print copies that is a custom of the traditional budget presentations.

Setup

Radial bubble tree chart visualization can be independently set up from the OpenSpending platform. Data can be sourced from a local file or it can be fetched from the OpenSpending platform using the platform API. In this study, each individual data entry was inserted within the specific tag in the code file. The bubble chart constructor have one argument, an object which defines the settings for the chart as shown by the below example. [27]

```
new BubbleTree(
            {
            data: data,
            container: '.bubbletree',
            bubbleType: 'donut',
            });
```

Listing 12: Bubble chart contractor [27].

The position where the bubble tree is to be inserted in the HTML DOM structure is defined by setting the container property as shown above. The data format is nested in a tree structure format. Two elements in the nodes *label* and *amount* must be defined. The format is shown by the following code example. [27].

```
var data={
           label: 'Kenya Budget Estimates',
           amount: 430448165055,
           }
```

Listing 13: Bubble chart important variables [27].

The child nodes are nested within the children array as shown by the example below

```
var data={
            label: 'Kenya Budget Estimates',
            amount: 430448165055,
            children: [{
                    label:"Agriculture",
                    amount:200000
                     },
                     {
                     label:"Education",
                     amount:200000
                        }]
                    }
```

Listing 14: Bubble chart nesting [27].

Data on the child nodes can be read by clicking on the parent node in the tree structure that then opens up as shown in the following figure.

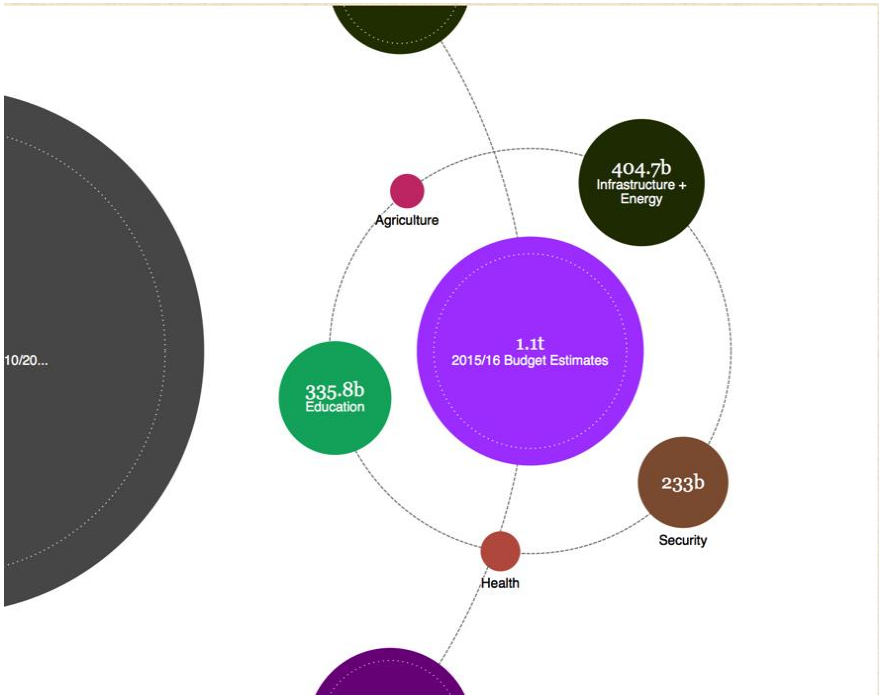

Figure 20: Visualization of the Kenyan budgets estimate from 2010/2011 to 2015/2016. Data from the Institute of Economic Affairs in Kenya [11].

A unique identifier can be assigned to the node by setting the *id* property as demonstrated in the example below.

```
var data={

          id:"Ken-Bud",
          label:'Kenya Budget Estimates',
          amount:430448165055

          }
```

Listing 15: Bubble tree unique identifier definition [27].

An alternative way of identifying the nodes is by assigning a taxonomy and name property to the node as shown below.

```
var data={

          taxonomy: "KB",
          name:'2015/2016 Budget',
          label: 'Kenya Budget Estimates',
          amount: 430448165055

          }
```

Listing 16: Taxonomy and name identifiers of a bubble tree [27].

Defining the *bubbletype* property sets the display type. This determines the kind of class used to display the bubble icons, three different types of values can be used as shown by the below examples [27]

```
bubbleType: 'donut',
bubbleType: 'plain',
bubbleType: 'icon',
```

Listing 17: Display type definition of a bubble tree [27].

Defining the bubble style helps to customize the look of the icons. Styling can be done for each taxonomy or for individual identifiers. This can be done by defining the colour and icon property. By default bubble sorting is done by amount in alternating manner. This can be changed by defining the sortby property to “label” as illustrated below. [27.]

```
sortBy = 'label';
```

Listing 18: Bubble sorting definition [27].

Tooltip is not yet integrated as part of the bubble chart implementation so far but it can be added to the tree chart by the use of a simple API. Event handlers for the tooltip show and hide event can be set with the ***tooltipCallback*** property in the configuration. The code below is an illustration of how it can be done.

```
var getTooltip = function() {
                return this.getAttribute('tooltip');};

var initTooltip = function(node, domnode) {

domnode.setAttribute('tooltip',node.label+' <b><br
/>KSH '+node.famount+'</b><br /><b>');

vis4.log(domnode.getAttribute('tooltip')); $(domnode).tooltip (
{delay: 20, bodyHandler: getTooltip});};
```

Listing 19: Tooltip and event handlers of bubble tree [27].

The tooltip is initiated by calling it as like below

```
initTooltip: initTooltip,
```

Listing 20: Initiating a tooltip [27].

7.4 Tree Map

The entire budget can be looked at as a whole by visualizing it using a TreeMap. The visualization was generated using the Open Spending budget visualization tool. Data is prepared in the comma separated value file format and uploaded to the Open Spending platform following a step process provided on the website. The tree map chart type is easily generated using the site API. Users can explore the budget and compare allocations for various government ministries. The picture below is a TreeMap visualization of the Kenyan Budget 2015-2016 estimates. [29]

Figure 21: Screen capture of the Kenyan budget 2015-2016 estimates. Copied from open spending (2015) [29].

From the visualization above, one can observe that apart from consolidated fund services, agriculture rural and urban development was allocated the highest spending and East African affairs was allocated the least. Users can have a fast and easy view of the entire budget estimate by hovering the mouse pointer on the various segmented rectangular boxes that separate each spending category.

Bar Charts

The Kenyan budget allocation to various functions of the government has varied in the financial years from 2010/2011 to 2015/2016. For instance allocation to education and infrastructure sectors has had an increasing trend in the five years' budget allocations as shown in the below bar chart.

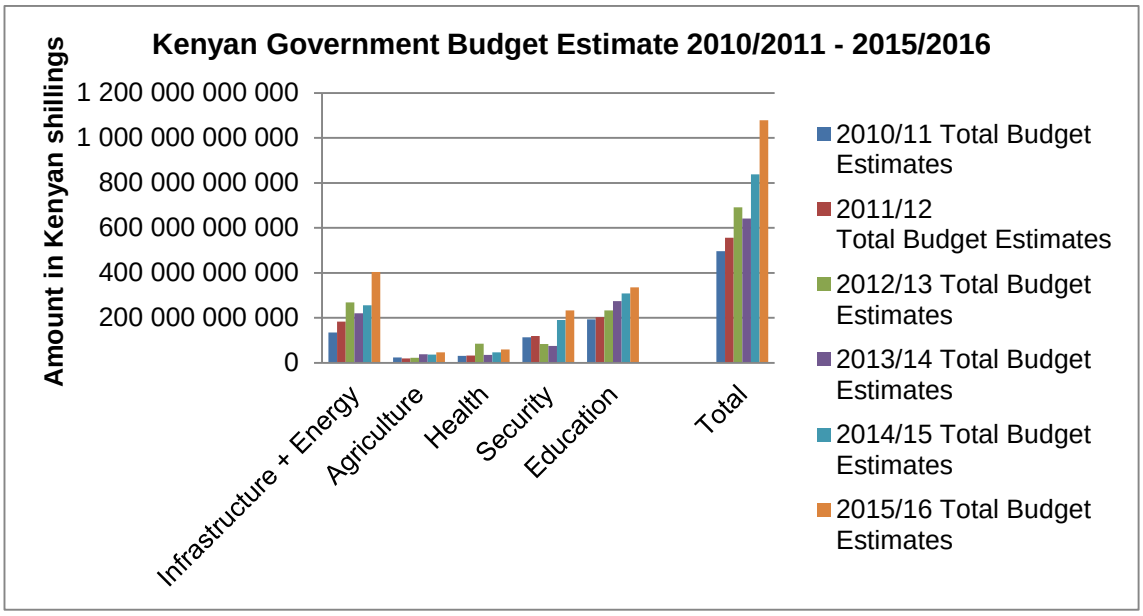

Figure 22: Kenya Budget Estimates for a 6 Year period focusing on 5 ministries. Data from the Institute of Economic Affairs in Kenya [11].

It is easy to notice a trend in financial allocation when data is viewed as the above bar chart. For instance it is noticeable that agriculture and health sectors have had the lowest budget allocation in the financial years visualized while infrastructure and education have had an increasingly higher allocation in the same period of time.

The interactive bar chart increases user interest in exploring the visualization. D3 plus library which is an extension of d3.js provides a very easy way of creating an interactive bar graph. The Kenyan budget estimates for the year 2010/11 to 2015/16 has also been visualized using the library as shown in the following figure.

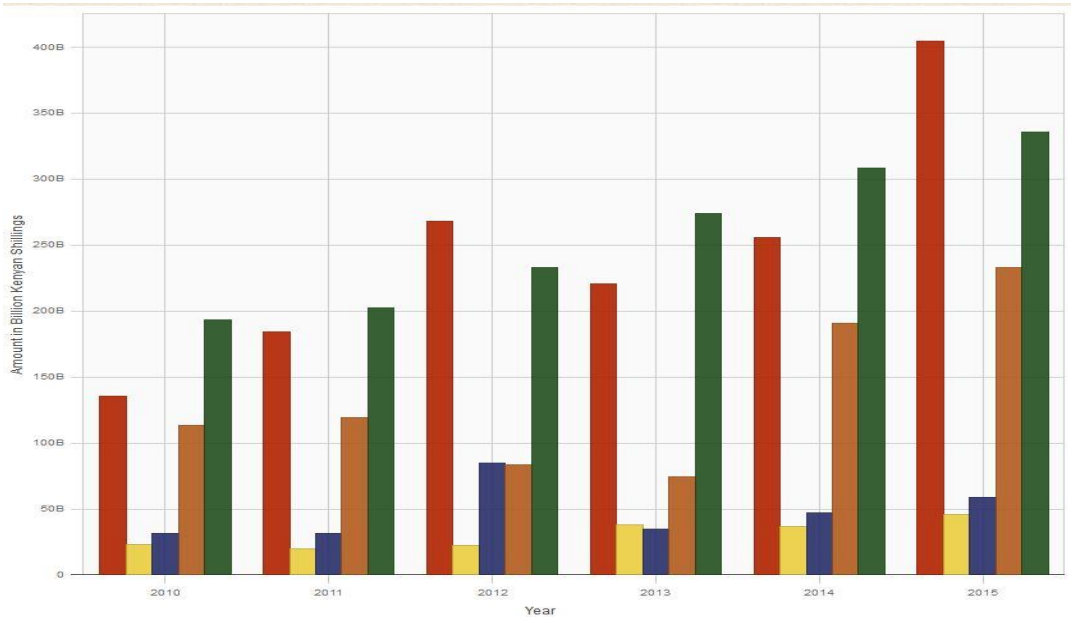

Figure 23: D3 plus bar chart visualization. Data from the Institute of Economic Affairs in Kenya [11].

D3 plus has a similar way of coding as D3.js. The chart is easily created by the chaining method as illustrated below.

```
var visualization = d3plus.viz()
    .container("#viz")
    .data(data)
    .type("bar")
    .id("name")
    .x("year")
    .y("value")
    .draw()
```

Listing 21: Chaining method of D3 plus [30].

The method ***.data()*** appends data values for the graph and ***.draw()*** creates the visualization. Reference to D3 and D3plus library has to be made in the head of the document file.

8 Analysis

The aim of this thesis has been to provide insight into the importance of visualization of financial data and the benefits it has for ordinary citizens and also for organizations. This study has explained the techniques used in visualization of data and various tools available for easy data visualization. Also different approaches and examples for visualizing various kinds of data have been explained.

Visualizing financial data is a process with many steps in which every step plays an important role towards the realization of a successful data representation. A reliable data source is an important prerequisite in the visualization process, because an error in the data source may mislead intended users to make critical decisions who may suffer huge financial loss. Data preparation consumes much of the visualization time but it is worth in the end because it makes the other following steps easy.

It is a challenge to find a perfect visualization tool. Usability of the technique is vital and it should offer enough information to the users for it to be considered effective. The scalability of the techniques used is a challenge since most visualizations are designed to be viewed on standard screen devices like computers and not on mobile phone platforms. Also a lack of standards to measure effectiveness of methods used makes it hard to evaluate if the technique used is effective for a particular type of data representation.

The choice of tools and methods to use in representing data is also an important factor because it is what determines how the target audience will perceive the data represented. Understanding of budget and spending information helps citizens to know how their taxes are being spent by the government and guides them in future decision making.

9 Discussion and conclusion

Data visualization provides users an easy way to check through data and easily understand information or a story that is behind the numbers. This is because it is not an easy task to go through / understand data which is complex in nature or which is big in quantity. Understanding of data is important to users because it helps them to have different perspectives on issues related to them and also guides them in decision making.

A lot of institutions publish monetary data on their budgets, spending data or procurements using conventional methods such as PDF files. Users often find it hard to read through and understand financial data that is presented using these methods. Visualizing these data makes it interesting and easy for users to read through and understand reports that otherwise might not have been easy for them.

There are many techniques that can be used to visualize data. The kind of method to be used in visualization is determined by factors such as type of data, story that is supposed to be communicated and time available to do the visualization. Also there are many online applications that can be used to visualize data in an easy way. Some of these tools might be free for everyone to use but others might involve a cost for using them.

The study gives examples of successful budget and financial data visualization that has been done by various different people and organizations. Pictures of their visualization have been shown in the discussion. It is important to know that different kinds of monetary data might need different tools and approaches in visualization. In order to have a successful visualization, the choice of tools and knowing the kind of story to tell is essential.

Information empowers people by increasing their knowledge and understanding of issues. Financial data visualization simplifies and increases people's understanding of how budgeting and spending has been done by governments and other institutions.

References

1. Ward M, Grinstein G, Keim D. Interactive Data Visualization Foundations, Techniques, and Application. India: A K Peters Ltd; 2010.

2. Data Visualization [online]. Seattle, WA; 26th November 2014. URL:http://www.tableausoftware.com/solutions/data-visualization. Accessed 16 November 2014.

3. Bostock M. Data-Driven Documents [online]. Sans Francisco, CA: The New York Times; March 2014. URL: http://d3js.org/. Accessed 17 November 2015.

4. Gutierrez S. Learn how to make data visualization with D3.js [online]. Rye, NY: August 2013. URL: https://www.dashingd3js.com. Accessed 16 November 2014.

5. Meeks E. D3.js in Action. Shelter Island, NY: Manning Publications; 2014.

6. Belmonte N. JavaScript InfoVis Toolkit; 2013. [online] URL: http://philogb.github.io/jit/index.html. Accessed 27 October 2015

7. Watts G. jQuery Sparklines [online]. URL: http://omnipotent.net/jquery.sparkline/#quickstart. Accessed 29 October 2015.

8. Google Developers. Google Charts [online]. 7 August 2014. URL: https://developers.google.com/chart/interactive/docs/index. Accessed 3 January 2015.

9. Highcharts Developers. Highcharts [online]. 16 September 2014. URL:http://www.highcharts.com/docs/getting-started/system-requirements. Accessed January 2015.

10. Muzammil K, Sarwar Shah K. Data and Information Visualization Methods, and Interactive Mechanisms [online]. Pakistan; November 2011. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.4640&rep=rep1&type=pdf. Accessed 10 September 2015.

11. Institute of Economic Affairs [online]. URL: http://www.ieakenya.or.ke/. Accessed 2 September 2015

12. Combining Chart Types in Excel [online]. URL: https://blogs.office.com/2012/06/21/combining-chart-types-adding-a-second-axis/. Accessed 11September 2015.

13. InterMath Scatter Plot Visualization [online]. URL: http://intermath.coe.uga.edu/dictnary/descript.asp?termid=320

14. Sample Data Visualization [online]. Accessed 15 September 2015 URL: http://www.dundas.com/gallery/data-visualization-gallery/treemap/

15. Bamberger M, Rugh J, Mabry L. Real World Evaluation, Working Under Budget, Time, Data, and Political Constraints. India: SAGE Publication Limited; 2012.

16. Diamond S, Yuk M. Data Visualization for Dummies. United States: John Wiley & Sons, Incorporated; January 2014.

17. Andy K. Data Visualization: a successful design process. United Kingdom: Birmingham Packt Publishing; January 2014.

18. Where Does My Money Go? [online].
URL: http://wheredoesmymoneygo.org/. Accessed 12 August 2015.

19. Meie Raha [online].
URL: http://meieraha.eu/?lang=en. Accessed 15 August 2015

20. Open Society Foundations projects [online].
URL: http://ttf.visualizingdata.opensocietyfoundations.org/projects. Accessed 21 August 2015.

21. Four Ways to Slice Obama's 2013 Budget Proposal [online].
URL: http://www.nytimes.com/interactive/2012/02/13/us/politics/2013-budget-proposal-graphic.html. Accessed August 2015

22. Your 2012 federal budget explained [online].
URL: http://www.theglobeandmail.com/news/politics/budget/infographic-your-2012-federal-budget-explained/article2384109/. Accessed 29 August 2015

23. Kenya Budget Explorer [online].
URL: twaweza.org/uploads/flash/budget-visualization-kenya-000/Kenya.html. Accessed 1 September 2015

24. Kenya Open Data [online].
URL: https://opendata.go.ke/. Accessed 2 August 2015

25. Pyle D. Data preparation for data mining. Sans Francisco, CA: Morgan Kaufmann Publishers; 1997.

26. McCandless D. Radial Bubble Tree [online].
URL: http://okfnlabs.org/bubbletree/. Accessed 4 July 2015.

27. McCandless D. Radial Bubble Documentation [online].
URL: https://github.com/okfn/bubbletree/wiki/Bubble-Tree-Documentation . Accessed 04 July 2015.

28. Shneiderman B. Treemaps for space-constrained visualization of hierarchies [online]. 25 June 2009.
URL: www.ifs.tuwien.ac.at/~silvia/wien/vu-infovis/articles/shneiderman_treemap-history_1998-2009.pdf. Accessed 03 July 2015.

29. Kenya Budget Visualization [online].
URL: https://openspending.org/kenyabudget-2015-2016/views/kenyan-budget-estimates-for-the-year-2015-2016. Accessed 10 September 2015.

30. D3 plus documentation [online].
URL: https://github.com/alexandersimoes/d3plus/wiki. Accessed 22 September 2015.