

# Survey Public Visualization Services

Stefan Kölbl  
Markus Unterleitner  
Benedict Wright

Graz University of Technology  
A-8010 Graz, Austria

4 May 2012

## **Abstract**

This survey is about public visualization services. It gives insight to five popular visualization services, points out their strengths and weaknesses and compares them. The survey includes both online and offline services. Two of the key points are the handling of user data, for instance if user data is kept private, and the different visualizations types offered by the service.



# Contents

- 1 Introduction** **1**
  - 1.1 Outline of the survey . . . . . 1
  
- 2 Many Eyes** **3**
  - 2.1 Uploading Data . . . . . 3
  - 2.2 Visualizations . . . . . 3
  - 2.3 Strength and Weaknesses . . . . . 4
  
- 3 Spotfire** **7**
  - 3.1 Description . . . . . 7
  - 3.2 Data Import . . . . . 7
  - 3.3 Visualizations . . . . . 7
  - 3.4 Strengths and Weaknesses . . . . . 9
  
- 4 Tableau** **10**
  - 4.1 Description . . . . . 10
  - 4.2 Data Import . . . . . 10
  - 4.3 Visualization . . . . . 10
  - 4.4 Strengths and Weaknesses . . . . . 11
  
- 5 VIDI** **15**
  - 5.1 Uploading and exploring a dataset . . . . . 16
  - 5.2 Visualizations . . . . . 16
  
- 6 DataMarket** **18**
  - 6.1 Upload and publish data . . . . . 18
  - 6.2 Exploring datasets . . . . . 19
  
- 7 Comparison and Conclusion** **21**
  
- References** **22**



# Chapter 1

## Introduction

Public visualization services allow users to create visualizations from data sets without the need of much background knowledge. The users can provide their own data sets or use existing ones and choose from a number of visualizations. Some services only need a modern browser and only require to register at the website.

In this survey Many Eyes, Spotfire, Tableau, Datamarket and Vidi are compared. This service have been chosen to have a wide range of visualization services.

To give a fair comparison of the services two common data sets have been used and visualized

- **Vienna Expenses:** data of the expenses of the city Vienna provided by [Open Data Vienna, 2011]
- **Worldwide Alcohol Consumption:** data of each countries alcohol consumption per capita provided by [World Health Organization, 2005]

### 1.1 Outline of the survey

First the five visualization services are presented to point out their characteristics. Examples are given for the two data sets and other noteworthy visualizations are shown. In the next section a overview of the features is given to compare the different services. The last section gives a conclusion over the current state of public visualization services.



## Chapter 2

# Many Eyes

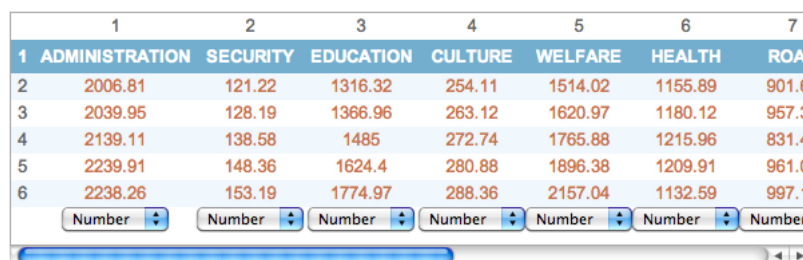
Many Eyes is a web site by IBM that provides a collaborative information Visualization service [Viégas et al., 2007]. The users can upload data sets, visualize them and share them with other users. It also allow the users to rate and discuss Visualizations from other users. The only requirement is a modern browser with Java and Flash and registration on the website <http://www.many-eyes.com>. The service is free of charge without restrictions. All the data sets and Visualizations created by the users are public.

### 2.1 Uploading Data

Many Eyes provides a HTML form to upload data. Uploading your own data set is done in a few steps

1. Prepare the data and paste the data set
2. Check if Many Eyes recognized the data set as desired.
3. Add meta-information to your data set

Many Eyes accepts only two types of data to upload onto the website. The first ones are tables in the CSV format with tabs as separators. The other format is free text. Some Visualizations have different interpretations on the data set and have to be adjusted accordingly to get the desired result. The maximum size is limited to 5 megabytes. The HTML form allows to simply paste the data into a text field. Many Eyes gives a preview to show you how it interprets your data (see figure 2.2).



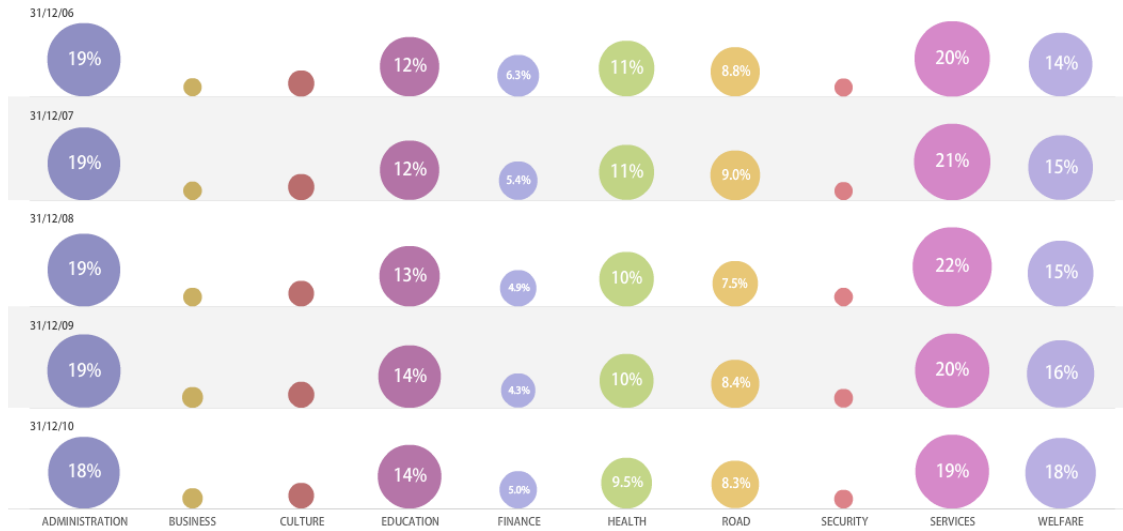
	1	2	3	4	5	6	7
1	ADMINISTRATION	SECURITY	EDUCATION	CULTURE	WELFARE	HEALTH	ROAD
2	2006.81	121.22	1316.32	254.11	1514.02	1155.89	901.6
3	2039.95	128.19	1366.96	263.12	1620.97	1180.12	957.3
4	2139.11	138.58	1485	272.74	1765.88	1215.96	831.4
5	2239.91	148.36	1624.4	280.88	1896.38	1209.91	961.0
6	2238.26	153.19	1774.97	288.36	2157.04	1132.59	997.1

Figure 2.1: Many Eyes preview of data

Adding additional information to your data set is important to make it easier for other people to both understand and find your data set.

### 2.2 Visualizations

The process of creating a Visualization consists of three steps



**Figure 2.2:** Visualization of the expenses of Vienna with a bubble matrix

1. Choose a data set
2. Choose a Visualization
3. Customize and publish

Many Eyes provides almost 300.000 data sets uploaded by the users and the possibility to upload your own data set (see 2.1). After selecting the data the user can choose from various types of Visualizations which are combined into the following groups

- Analyze a text
- Compare a set of values
- See relationships among data points
- See the parts of a whole
- See the world

For every Visualization exists a detailed guide which explains what it is useful for, how it works and what the data requirements are. The Visualizations are interactive and allow the user to change parameters and explore the data set in different ways depending on the type of Visualization. Export is only possible in the PNG format. The data set "Vienna Expenses" is visualized as a bubble matrix in figure ???. Many Eyes provides map data for a few countries and allows to visualize appropriate data. The data set provided by the WHO is visualized in figure 2.3.

Many Eyes provides different text based visualizations like Word Trees, Phrase Nets and Word Clouds (see figure 2.4).

## 2.3 Strength and Weaknesses

Many Eyes is designed to share everything with everybody. There is no possibility to keep your data or visualizations private, which is a drawback. On the other side this offers great value with hundreds of thousands visualizations and data sets due to the open structure of this service. An issue is that there are not many options to upload the data and there is almost none help to make the data clean. Many Eyes is very easy to use and it is possible to create good visualizations in minutes. The visualizations are interactive and there are lots of different types to choose from.







# Chapter 3

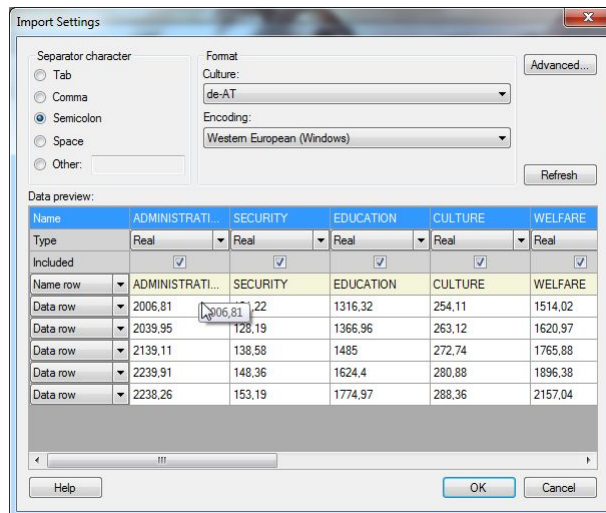
# Spotfire

## 3.1 Description

Spotfire is a free Software for creating information visualizations [Tibco, 2012] It requires an Account with TIBCO, which is the publisher of this software, to log in and use the application. Although an online authentication is required, the Data and the visualizations are Stored locally.

## 3.2 Data Import

Spotfire can read a wide variety of data formats, including CSV, XLS (Microsoft Excel Spreadsheet), TXT and MDB (Microsoft Access Database). During the Data-Import action further settings can be chosen, e. G. encoding, language and other format specific settings like the separator character in CSV files. After import a sample visualization is generated. This visualization can be tweaked or removed. The application lets you define the Dimensions of the chart and its colour coding.



**Figure 3.1:** This image demonstrates the data-input dialogue, here the data-types for each column can be checked and altered.

## 3.3 Visualizations

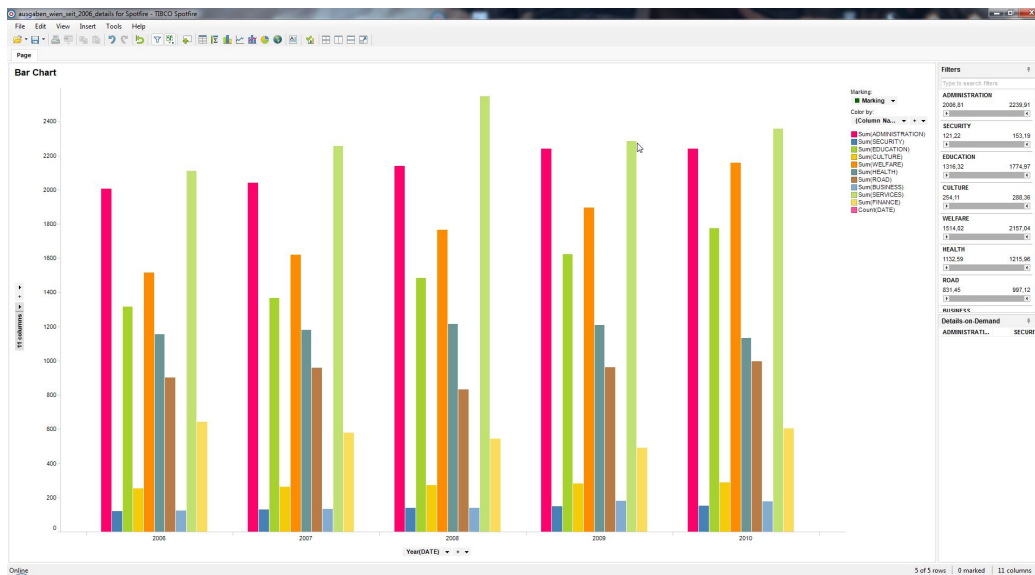
Supported visualizations:

- bar-charts

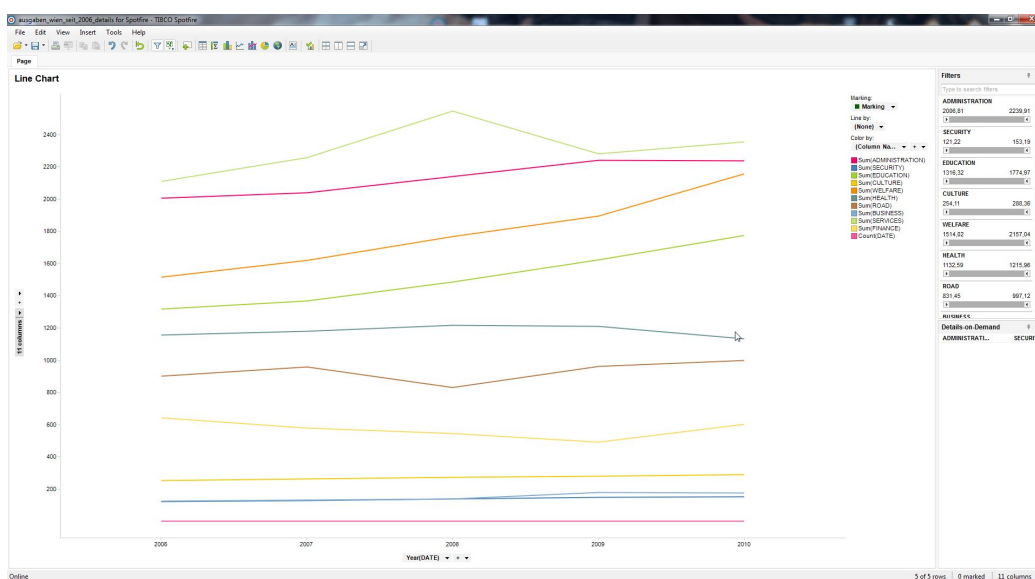
- line-charts
- mixed (bar and line) charts
- map-chart.

Spotfire however does not support the actual map data. This has to be supplied by the dataset. Other Software provides this data, and can map country names to a map representation. The creating of the visualization is completely interactive and intuitive. Data can be filtered and ranges defined, dimensions changed. Clicking on an entry will reveal its details. Spotfire supports two export features: as HTML and as Image. Although the HTML version only contains an image of the visualization but no interactivity.

Following images show the application and a two visualizations.



**Figure 3.2:** Barchart showing the expenses of Vienna from 2006 to 2010.



**Figure 3.3:** This shows a more clear line-chart of the Vienna expenses.

### **3.4 Strengths and Weaknesses**

The main strength of Spotfire compared to other free visualization tools, is the fact that the data and the visualization stay private. Spotfire offers a HTML export function, to easily publish the visualization on the internet. Sadly this export is not interactive, which would make Spotfire a stronger competitor to the other tested applications.

# Chapter 4

## Tableau

### 4.1 Description

Tableau, similar to Spotfire, is a free Software for creating information visualizations [Tableau Software, 2012]. To be able to use the software, an account has to be created. Every visualization created with Tableau is stored online and open to the public, with the option to buy your own server. So when using Tableau one has to keep in mind, that the underlying data is also available to the public. Tableau offers a wide range of examples and training material for the beginner and advanced user. When launching the application the user the option to either view visualizations and comment on them, viewing training material or creating their own graphic.

### 4.2 Data Import

Tableau support a wide range of data sources, from local files like CSV and Excel to data stored on in a database on a server. The most data sources however are restricted to professional users with a paid licence. The data is automatically ordered in to Dimensions and Measures, but can be changed manually. After the data is imported the user can interactively create the visualization by dragging the dimensions and measures on to the workspace.

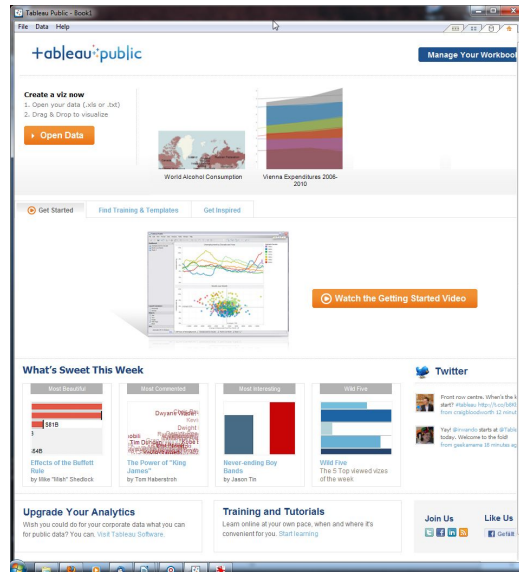
### 4.3 Visualization

Tableau supports many different types of visualizations:

- heat-maps
- highlight tables
- different bar-charts
- scatter-plots
- area-charts
- pie-charts
- line-charts
- map-charts

Tableau even provides a map data, and can interpret country names to match the data to the country, province. After the creation process, the visualization together with the data is stored online. This stored version is the access able via a modern browser. The resulting visualization is very interactive so the viewer can apply filters or get detailed information about a specific entry.

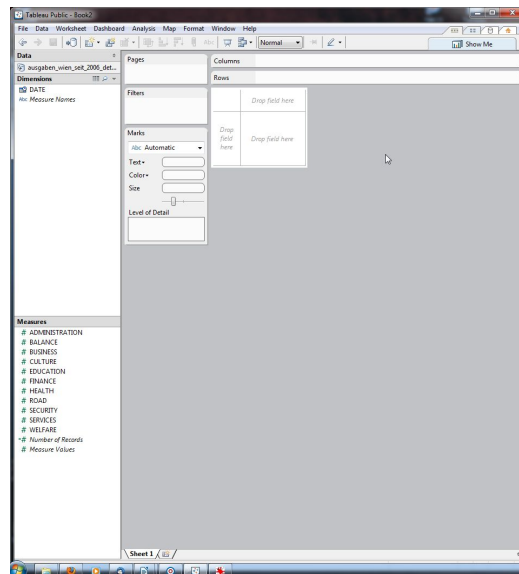
Following images show the application and two visualizations.



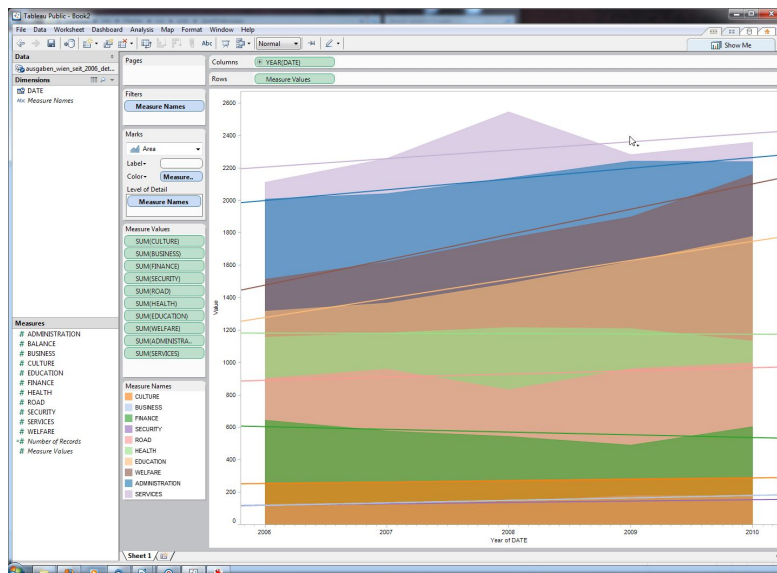
**Figure 4.1:** This shows the Tableau start screen. From here one can load existing projects or start a new one. One can also have a look at training material and visualizations by other users.

## 4.4 Strengths and Weaknesses

The most appealing feature of Tableau is the way visualizations are created. Also its versatility puts it ahead of other Tools like Spotfire. Another strength of this application is the interactive result. Users can filter the data and manipulate the visualization after it has been published. On setback is that all the data has to be public, except if you buy a licence for your own server.



**Figure 4.2:** After importing data to Tableau the application splits it up in to Dimensions and Measures. Measures being numeric real values and Dimensions being strings.



**Figure 4.3:** This area chart shows the expenses of Vienne between 2006 and 2010.



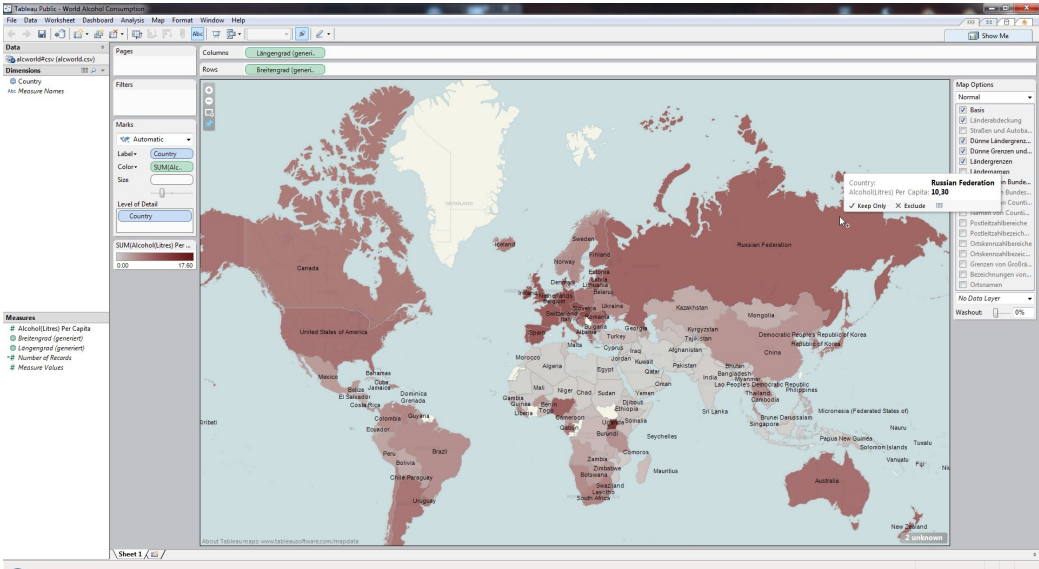


Figure 4.4: This map visualizes the per capita consume of alcohol.



# Chapter 5

## VIDI

VIDI is a module for Drupal [Drupal, 2012], a public and open-source content management system (CMS). VIDI is free and can be downloaded at [Jefferson Institute, 2012].

As it is a module for Drupal it can be installed on any webserver (like Apache or IIS). This allows to implement various visualizations on a website. With the importer module data can be imported to a MySQL database as a new data table. A user with the required access rights can select which MySQL data table is available for the visualizations. Then the data can be selected within the visualization modules to be displayed. Visualizations can be added in the CMS to a selected webpage.

The CMS webpages can be configured such that some pages are only visible to users that are logged in. Therefore visualizations can be selected to be private or publicly available. Of course it is always possible to configure the webserver to be accessible only for the internal network of a company. So there are many options to keep data and visualizations private.

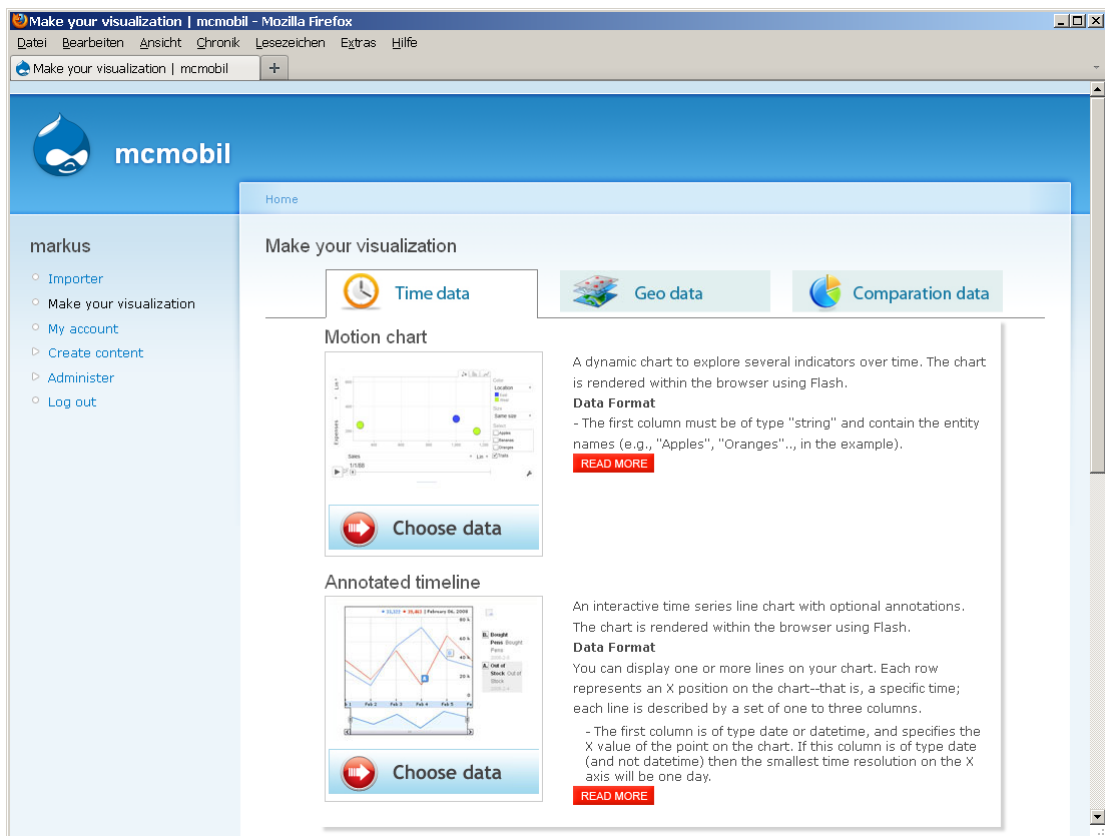
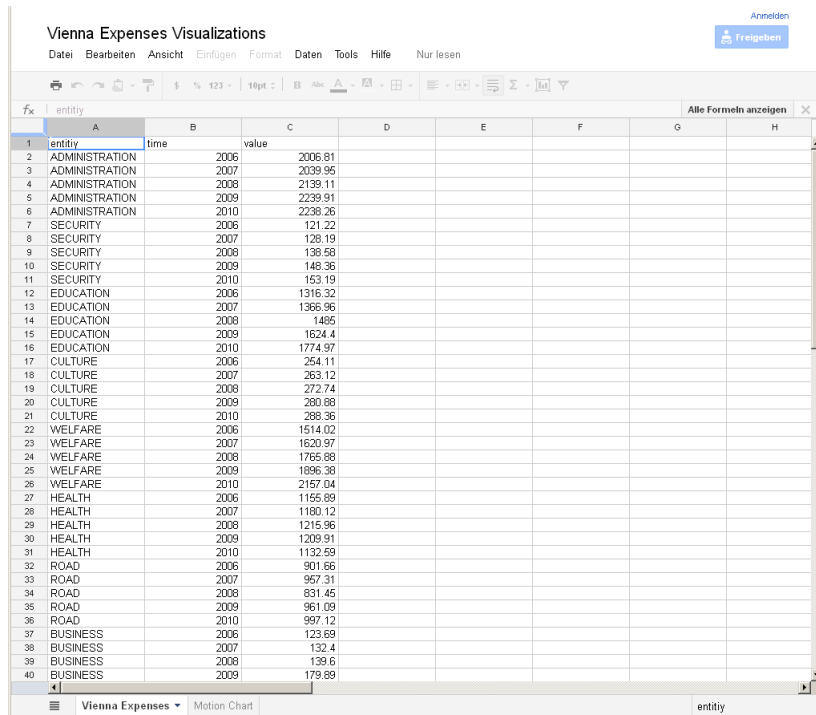


Figure 5.1: Screenshot of the selected VIDI module on the Drupal CMS.

## 5.1 Uploading and exploring a dataset

The setup of the Drupal CMS and the configuration of the VIDİ module worked fine on two different web-servers. But when uploading the dataset Vienna expenses 2006-2010 the system always generated an unspecific error or crashed. So the visualizations of the VIDİ module could not be tested. As the VIDİ module uses the Google Visualization API the dataset has been visualized by uploading the data Google Docs and a flash-based motion-chart was generated. It shows, how the VIDİ module would visualize the data and what can be done with it.



	A	B	C	D	E	F	G	H
1	entity	time	value					
2	ADMINISTRATION	2006	2006.81					
3	ADMINISTRATION	2007	2039.96					
4	ADMINISTRATION	2008	2139.11					
5	ADMINISTRATION	2009	2239.91					
6	ADMINISTRATION	2010	2238.26					
7	SECURITY	2006	121.22					
8	SECURITY	2007	128.19					
9	SECURITY	2008	138.58					
10	SECURITY	2009	148.36					
11	SECURITY	2010	153.19					
12	EDUCATION	2006	1316.32					
13	EDUCATION	2007	1366.96					
14	EDUCATION	2008	1485					
15	EDUCATION	2009	1624.4					
16	EDUCATION	2010	1774.97					
17	CULTURE	2006	254.11					
18	CULTURE	2007	263.12					
19	CULTURE	2008	272.74					
20	CULTURE	2009	280.88					
21	CULTURE	2010	288.36					
22	WELFARE	2006	1514.02					
23	WELFARE	2007	1620.97					
24	WELFARE	2008	1765.88					
25	WELFARE	2009	1896.38					
26	WELFARE	2010	2157.04					
27	HEALTH	2006	1155.89					
28	HEALTH	2007	1180.12					
29	HEALTH	2008	1215.96					
30	HEALTH	2009	1209.91					
31	HEALTH	2010	1132.59					
32	ROAD	2006	901.66					
33	ROAD	2007	957.31					
34	ROAD	2008	831.45					
35	ROAD	2009	961.09					
36	ROAD	2010	997.12					
37	BUSINESS	2006	123.69					
38	BUSINESS	2007	132.4					
39	BUSINESS	2008	139.6					
40	BUSINESS	2009	179.89					

Figure 5.2: Screenshot of the uploaded data, Vienna expenses 2006-2010

## 5.2 Visualizations

Three modules are used to visualize user defined datasets. There is a module to make visualizations with time-data, with geo-data and to compare datasets with each other. All modules use the Google Visualization API.

The visualization of time-data is based on a timeline map plugin for Drupal, which was published in a Google Code Project as JavaScript Code. This module allows to display several datasets in a Motion Chart or in an annotated timeline. Both modules are using flash. It uses a window for displaying a specified range in the dataset, where only the data within the defined range is visible.

The visualization of Geo-data uses the Tagmap plugin for Drupal. It allows to present location based data on a geographical map. There are three types of visualizations: the Tag Map, the Intensity Map and the Geomap. The Tag Map displays labels on a geographical map weighted by the provided numerical data. As geographical map the Google map is used. The Intensity Map displays selected countries in a specific colour intensity according to the given numeric value. The Geo Map uses flash for displaying data on a specific map. The numeric values are displayed as color scales for a defined location. Therefore the geo-coordinates or the address of the location must be provided.

To compare datasets also the Google Visualization API is used. The user can choose from seven different visualization types. Pie chart, Gauge chart, Bar chart, Area chart, Column chart, Line chart and Scatter chart.



**Figure 5.3:** Screenshot of the motion chart generated by the Google Visualization API

Each chart type needs the data to be prepared in a defined way to make the chart work. The charts are saved using vector graphics as SVG (Scalable Vector Graphics) or VML (Vector Markup Language).

Examples for the VID-Module:

- The Jefferson Institute: Blogs on the <http://www.jeffersoninst.org/> Website.
- Patchwork Nation communities by the Jefferson Institute: <http://www.patchworknation.org/>
- The Knight Foundation (Knight Soul of the Community Website). Example of an emotion chart in Akron, Ohio <http://www.soulofthecommunity.org/akron>

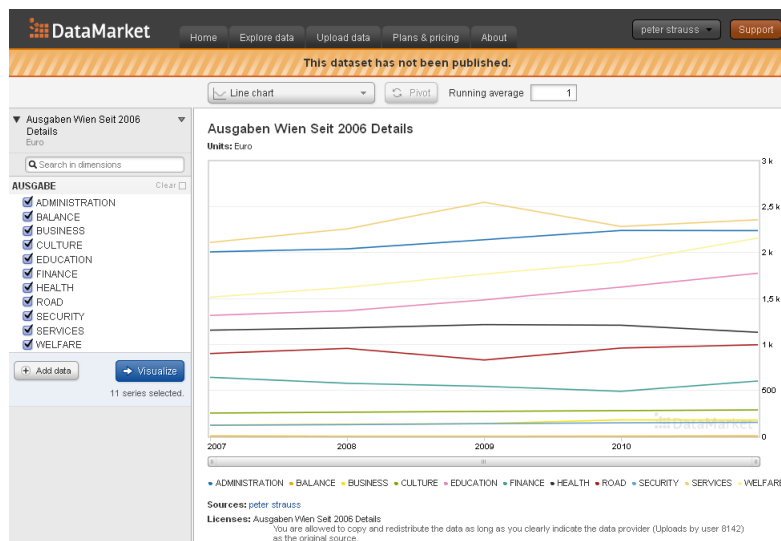
# Chapter 6

## DataMarket

DataMarket is a website for sharing and visualizing data for free [DataMarket, 2012]. DataMarket is online and available for every one. There are many visualization tools to display the data in the desired way. For using some features the user has to register and create an account, this is also for free.

There is also a premium service where the user can access premium datasets for which he has to pay. Premium datasets can be for example research data or financial data. The prices for the datasets vary. The premium service provides three types of subscriptions: a Pro, Corporate and an Enterprise Subscription. With the premium account users for example can sell their datasets and create and publish their own interactive charts.

DataMarket provides also a programming API. The price depends on the number of requests. Up to 50 requests per month the service is free. The API can be integrated in various existing systems and returns the requested result in CSV, XML or JSON data format. The data tables and charts can be configured to be embedded in an external website or articles.



**Figure 6.1:** Screenshot of the user-defined dataset Vienna expenses 2006-2010, generated by datamarket.com

### 6.1 Upload and publish data

Right after creating a free account the user can log in and select the button upload in the main navigation menu of the website. The Excel file of the Vienna expenses 2006-2010 can be selected and uploaded. It is important that the Excel file has the right structure. There are three sample Excel files where the user can see the required

file structures for different data structures (like multidimensional or hierarchical datasets). After the upload a preview is generated and if this looks correct the dataset can be described with some additional information (Title, Source Url, Author, Research method, Accessibility, Price,...). A different accessibility than private and a price, as explained before, can only be selected if the subscription package is Pro or higher.

Data can be uploaded as Excel-, CSV- or SPSS-file. The data can be selected as private, as public or as shared within a user group. The ability to select one of these options and also the size of the uploaded data depends on the type of the subscription package. The free account only supports Excel files with a maximum size of 15MB. The selectable accessibility is only private therefore a price cannot be chosen. For selecting an other accessibility minimum the Pro subscription package is required.

## 6.2 Exploring datasets

On the website many datasets can be explored. Each displayed dataset on the website can be explored. A click on "explore" below the displayed chart opens a new browser-window. On the left side of the browser-window the displayed data can be chosen. The possibilities for selecting sub datasets vary and depend on the uploaded data. For all datasets the same chart types can be selected. In some datasets it is not possible to use all chart types, because the data doesn't provide the right format. The following charts can be selected:

Line chart, relative line chart, stacked area chart, 100% stacked area chart, bar chart, stacked bar chart, column chart, stacked column chart, 100% stacked column chart, table, current values and pie chart. Sometimes a pivot function for a calculating running average can be selected.

Many further examples can be selected and explored on <http://www.datamarket.com>.





## Chapter 7

# Comparison and Conclusion

There are a lot of different visualization services to choose from and each of them has his strengths and weaknesses. The right choice depends largely on what the users requirements are and whether the data and visualizations need to stay private.

If you just want to quickly visualize something, Many-Eyes and DataMarket are a good choice, because you only need to register at the according homepage and can start straight on uploading your data and visualize it. Tableau is a more sophisticated solution and offers the most options to make a individual visualization. Spotfire seems to be a bit behind on terms of usability compared to Tableau. Vidi looks like a nice solution if you have your own website using Drupal, but unfortunately we could not look deeper into it due to technical problems.

Feature	Many-Eyes	Spotfire	Tableau	Vidi	DataMarket
free	X	X	X	X	X
data public	X		X	1	1
vis. public	X		X	1	1
offline		X	X		
SVG export		X		X	X <sup>2</sup>
payed privacy			X		
payed features		X	X		X
map chart	X		X	X	
tag/word cloud	X				
interactive gen.	X	X	X		X
interactive result	X		X	X	X

**Figure 7.1:** Comparison of the public visualization services

<sup>1</sup>selectable by user    <sup>2</sup>only for PRO users

# References

DataMarket [2012]. *DataMarket*. <http://datamarket.com>.

Drupal [2012]. *Drupal - Open Source Content Management System*. <http://drupal.org/>.

Jefferson Institute [2012]. *Vidi - Drupal Visualization module*. <http://www.dataviz.org/>.

Open Data Vienna [2011]. *Expenses of the city Vienna*. <http://data.wien.gv.at/katalog/finanzgebarung-ausgaben.html>.

Tableau Software [2012]. *Tableau Software*. <http://www.tableausoftware.com/>.

Tibco [2012]. *TIBCO Spotfire*. <http://spotfire.tibco.com/>.

Viégas, Fernanda B., Martin Wattenberg, Frank van Ham, Jesse Kriss, and Matt McKeon [2007]. *Many Eyes: a Site for Visualization at Internet Scale*. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 13(6), pages 1121–1128. ISSN 1077-2626. doi:10.1109/TVCG.2007.70577. <http://www.research.ibm.com/visual/papers/viegasinfovis07.pdf>. Proc. IEEE Symposium on Information Visualization (InfoVis 2007), Phoenix, Arizona, USA.

World Health Organization [2005]. *Alcohol Consumption*. <http://apps.who.int/ghodata/>.