

# Automated Table Generation



CS448b

Pontus Orraryd  
Oskar Ankarberg

# Project Description

An automatic table generator for the web.

The generator generates tables according to Edward Tufte and Stephen Few's visualization principles for tables.

The automatic table generator will also be user interactive with the ability to make custom changes to the initial table to produce a personalized table.

Input a data file and automatically get a well designed table back.

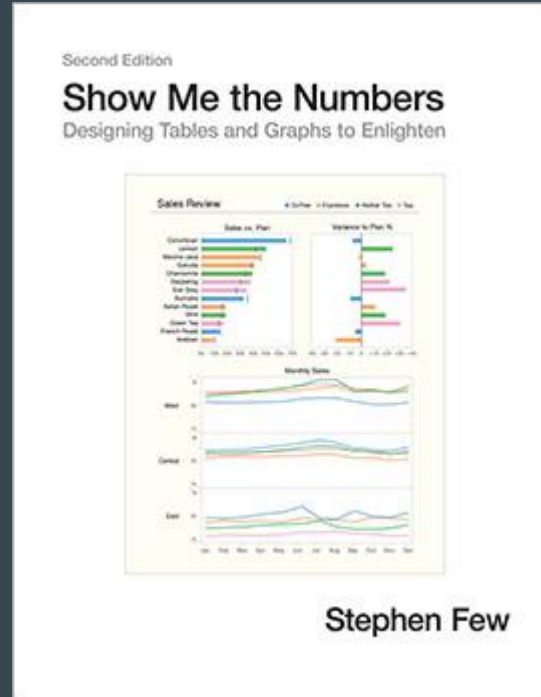
# Related Work

## Show me the Numbers

by Stephen Few

Have a great chapter on how to design tables.

Our main reference on table design.

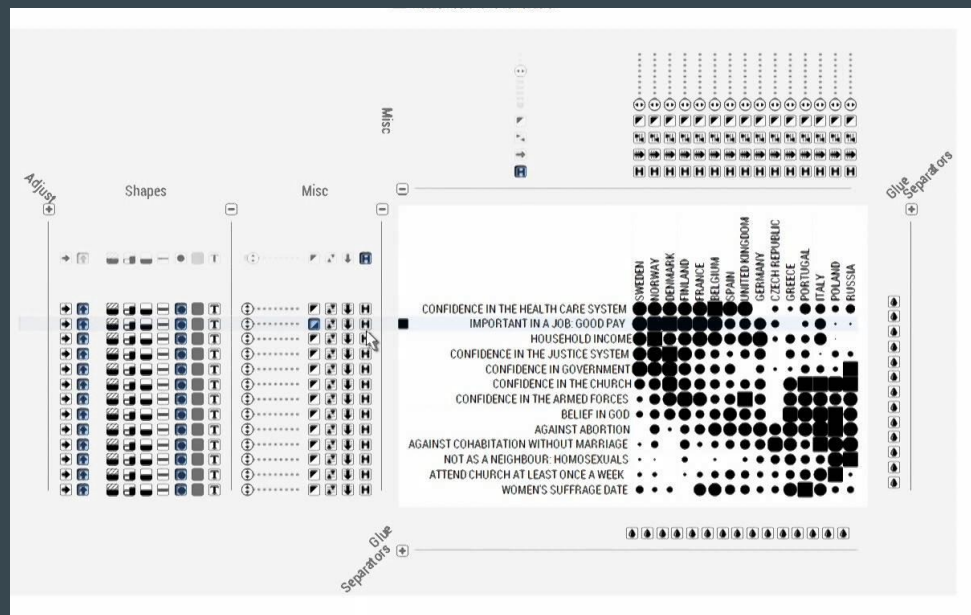


# Related Work

Bertifier

<http://www.bertifier.com/>

- More focused on data exploration.
- Many great interactive features.



# Table Design Principles

- Minimize graphics that are not directly related to data. Gridlines only when necessary.
- Emphasis on white space. A lot of white space makes it much easier to quickly scan the table.
- Use well-established fonts.
- Keep it simple, don't try to show off.

# Example

Cancer survival rate by type.

Example by Tufte

Cancer site	Relative survival rate, % (SE)			
	5 years	10 years	15 years	20 years
Oral cavity and pharynx	56.7 (1.3)	44.2 (1.4)	37.5 (1.6)	33.0 (1.8)
Oesophagus	14.2 (1.4)	7.9 (1.3)	7.7 (1.6)	5.4 (2.0)
Stomach	23.8 (1.3)	19.4 (1.4)	19.0 (1.7)	14.9 (1.9)
Colon	61.7 (0.8)	55.4 (1.0)	53.9 (1.2)	52.3 (1.6)
Rectum	62.6 (1.2)	55.2 (1.4)	51.8 (1.8)	49.2 (2.3)
Liver and intrahepatic bile duct	7.5 (1.1)	5.8 (1.2)	6.3 (1.5)	7.6 (2.0)
Pancreas	4.0 (0.5)	3.0 (0.5)	2.7 (0.6)	2.7 (0.8)
Larynx	68.8 (2.1)	56.7 (2.5)	45.8 (2.8)	37.8 (3.1)
Lung and bronchus	15.0 (0.4)	10.6 (0.4)	8.1 (0.4)	6.5 (0.4)
Melanomas	89.0 (0.8)	86.7 (1.1)	83.5 (1.5)	82.8 (1.9)
Breast	86.4 (0.4)	78.3 (0.6)	71.3 (0.7)	65.0 (1.0)
Cervix uteri	70.5 (1.6)	64.1 (1.8)	62.8 (2.1)	60.0 (2.4)
Corpus uteri and uterus, NOS	84.3 (1.0)	83.2 (1.3)	80.8 (1.7)	79.2 (2.0)
Ovary	55.0 (1.3)	49.3 (1.6)	49.9 (1.9)	49.6 (2.4)
Prostate	98.8 (0.4)	95.2 (0.9)	87.1 (1.7)	81.1 (3.0)
Testis	94.7 (1.1)	94.0 (1.3)	91.1 (1.8)	88.2 (2.3)
Urinary bladder	82.1 (1.0)	76.2 (1.4)	70.3 (1.9)	67.9 (2.4)
Kidney and renal pelvis	61.8 (1.3)	54.4 (1.6)	49.8 (2.0)	47.3 (2.6)
Brain and other nervous system	32.0 (1.4)	29.2 (1.5)	27.6 (1.6)	26.1 (1.9)
Thyroid	96.0 (0.8)	95.8 (1.2)	94.0 (1.6)	95.4 (2.1)
Hodgkin's disease	85.1 (1.7)	79.8 (2.0)	73.8 (2.4)	67.1 (2.8)
Non-Hodgkin lymphomas	57.8 (1.0)	46.3 (1.2)	38.3 (1.4)	34.3 (1.7)
Multiple myeloma	29.5 (1.6)	12.7 (1.5)	7.0 (1.3)	4.8 (1.5)
Leukaemias	42.5 (1.2)	32.4 (1.3)	29.7 (1.5)	26.2 (1.7)

Rates derived from SEER 1973-98 database (both sexes, all ethnic groups).<sup>12</sup>  
NOS=not otherwise specified.

Table 4: **Most recent period estimates of relative survival rates, by cancer site**

# Example

Changes:

- Fewer and more discrete grids
- Ordered by survival rate
- Introduced a hierarchy between % and std error.

## Estimates of relative survival rates, by cancer site

	% survival rates and standard errors							
	5 year		10 year		15 year		20 year	
Prostate	98.8	0.4	95.2	0.9	87.1	1.7	81.1	3.0
Thyroid	96.0	0.8	95.8	1.2	94.0	1.6	95.4	2.1
Testis	94.7	1.1	94.0	1.3	91.1	1.8	88.2	2.3
Melanomas	89.0	0.8	86.7	1.1	83.5	1.5	82.8	1.9
Breast	86.4	0.4	78.3	0.6	71.3	0.7	65.0	1.0
Hodgkin's disease	85.1	1.7	79.8	2.0	73.8	2.4	67.1	2.8
Corpus uteri, uterus	84.3	1.0	83.2	1.3	80.8	1.7	79.2	2.0
Urinary, bladder	82.1	1.0	76.2	1.4	70.3	1.9	67.9	2.4
Cervix, uteri	70.5	1.6	64.1	1.8	62.8	2.1	60.0	2.4
Larynx	68.8	2.1	56.7	2.5	45.8	2.8	37.8	3.1
Rectum	62.6	1.2	55.2	1.4	51.8	1.8	49.2	2.3
Kidney, renal pelvis	61.8	1.3	54.4	1.6	49.8	2.0	47.3	2.6
Colon	61.7	0.8	55.4	1.0	53.9	1.2	52.3	1.6
Non-Hodgkin's	57.8	1.0	46.3	1.2	38.3	1.4	34.3	1.7
Oral cavity, pharynx	56.7	1.3	44.2	1.4	37.5	1.6	33.0	1.8
Ovary	55.0	1.3	49.3	1.6	49.9	1.9	49.6	2.4
Leukemia	42.5	1.2	32.4	1.3	29.7	1.5	26.2	1.7
Brain, nervous system	32.0	1.4	29.2	1.5	27.6	1.6	26.1	1.9
Multiple myeloma	29.5	1.6	12.7	1.5	7.0	1.3	4.8	1.5
Stomach	23.8	1.3	19.4	1.4	19.0	1.7	14.9	1.9
Lung and bronchus	15.0	0.4	10.6	0.4	8.1	0.4	6.5	0.4
Esophagus	14.2	1.4	7.9	1.3	7.7	1.6	5.4	2.0
Liver, bile duct	7.5	1.1	5.8	1.2	6.3	1.5	7.6	2.0
Pancreas	4.0	0.5	3.0	1.5	2.7	0.6	2.7	0.8

# Progress

Implemented so far:

- Read csv file and create table.
- Sorting of columns.
- Optional circles for numerical data.
- Tried to keep the design simplistic and uphold design principles.

So far we have been focusing on implementing basic ways to control the look of the table.

## Showing data by Transaction Date

	Transaction Date	Product	Price	Payment Type	Name	City	State	Country	Account Created	Last Login	Latitude	Longitude
Transaction Date	1/2/09 6:17	Product1	●	Mastercard	carolina	Basildon	England	United Kingdom	1/2/09 6:00	1/2/09 6:08	●	●
Product	1/2/09 4:53	Product1	●	Visa	Betina	Parkville	MO	United States	1/2/09 4:42	1/2/09 7:49	●	●
Price	1/2/09 13:08	Product1	●	Mastercard	Federica e Andrea	Astoria	OR	United States	1/1/09 16:21	1/3/09 12:32	●	●
Payment Type	1/3/09 14:44	Product1	●	Visa	Gouya	Echuca	Victoria	Australia	9/25/05 21:13	1/3/09 14:22	●	●
Name	1/4/09 12:56	Product2	●	Visa	Gerd W	Cahaba Heights	AL	United States	11/15/08 15:47	1/4/09 12:45	●	●
City	1/4/09 13:19	Product1	●	Visa	LAURENCE	Mickleton	NJ	United States	9/24/08 15:19	1/4/09 13:04	●	●
State	1/4/09 20:11	Product1	●	Mastercard	Fleur	Peoria	IL	United States	1/3/09 9:38	1/4/09 19:45	●	●
Country	1/2/09 20:09	Product1	●	Mastercard	adam	Martin	TN	United States	1/2/09 17:43	1/4/09 20:01	●	●
Account Created	1/4/09 13:17	Product1	●	Mastercard	Renee Elisabeth	Tel Aviv	Tel Aviv	Israel	1/4/09 13:03	1/4/09 22:10	●	●
Last Login	1/4/09 14:11	Product1	●	Visa	Aidan	Chatou	Ile-de-France	France	6/3/08 4:22	1/5/09 1:17	●	●
Latitude	1/5/09 2:42	Product1	●	Diners	Stacy	New York	NY	United States	1/5/09 2:23	1/5/09 4:59	●	●
Longitude	1/5/09 5:39	Product1	●	Amex	Heidi	Eindhoven	Noord-Brabant	Netherlands	1/5/09 4:55	1/5/09 8:15	●	●
	1/2/09 9:16	Product1	●	Mastercard	Sean	Shavano Park	TX	United States	1/2/09 8:32	1/5/09 9:05	●	●
	1/5/09 10:08	Product1	●	Visa	Georgia	Eagle	ID	United States	11/11/08 15:53	1/5/09 10:05	●	●
	1/2/09 14:18	Product1	●	Visa	Richard	Riverside	NJ	United States	12/9/08 12:07	1/5/09 11:01	●	●
	1/4/09 1:05	Product1	●	Diners	Leanne	Julianstown	Meath	Ireland	1/4/09 0:00	1/5/09 13:36	●	●
	1/5/09 11:37	Product1	●	Visa	Janet	Ottawa	Ontario	Canada	1/5/09 9:35	1/5/09 19:24	●	●
	1/6/09 5:02	Product1	●	Diners	barbara	Hyderabad	Andhra Pradesh	India	1/6/09 2:41	1/6/09 7:52	●	●

# Todo

Some things we will implement next:

- Review visualization principles
- Detect “header columns” automatically.
- Extend to support more advanced data types (Excel etc).
- Reordering of columns.
- Add some design options (color, margins etc).
- Edit specific fields.

# Feedback Questions

- How much freedom should the user have on the design of the table?
- What feature would you want in a system like this?