

Alexander Kowarik¹,
Bernhard Meindl¹ and
Matthias Templ^{1,2}
1. Statistics Austria
2. Vienna University of
Technology
useR! 2011
Coventry, August 16-18,
2011

sparkTable: Generating Graphical Tables for Websites and Documents with R

Motivation

Introductory examples

The R-package sparkTable

Example

Outline

- ▶ Providing quick access to additional insights by the use of *graphical tables*
- ▶ Presentation of numerous data in a well-arranged way
- ▶ Improving data density by using spark-graphs
- ▶ Results should be easy to modify
- ▶ Development of a tool to create graphical tables easily

How can this be achieved? → R-package `sparkTable`

Example (production indices)
















Seasonally adjusted production indices by branch, monthly.

Produktionsindex ÖNACE 2008 (Ø 2005=100), EU harmonisiert saisonal bereinigt

Berichts- periode	Insgesamt (B bis F)	Industrie (B bis E)	EU insgesamt (B, C, D, F ohne 353)	EU Industrie (B, C, D ohne 353)	Verwendungskategorien (MIG)						ÖNACE 2008 - Abschnitte					
					Vorleistungen	Energie	Investitionsgüter	Konsumgüter			Bergbau und Gewinnung von Steinen und Erden (B)	Herstellung von Waren (C)	Energieversorgung (D)	Wasserversorgung und -ent-sorgung; Abfallentsorgung (E)	Bau (F)	
								Insgesamt	Langlebige Konsumgüter	Kurzlebige Konsumgüter						
Q 2006	107,2	107,5	107,1	107,3	106,7	108,7	110,4	105,1	108,4	104,5	110,8	107,4	106,8	108,6	106,2	
Q 2007	113,1	113,7	113,2	113,8	113,8	109,1	119,0	110,1	112,1	109,7	110,7	114,4	107,8	110,0	110,9	
Q 2008	114,7	116,0	114,2	115,3	113,8	124,1	122,0	108,0	113,7	106,8	117,5	115,6	118,8	119,0	110,2	
Q 2009	103,1	102,1	102,8	101,5	96,7	118,7	99,6	104,6	110,4	103,4	108,0	100,1	114,3	128,5	107,8	
2005	I.	95,2	95,7	95,6	96,2	97,5	90,5	94,6	98,8	93,7	99,9	109,2	96,2	87,9	92,7	93,6
	II.	96,6	97,0	96,9	97,2	97,9	95,4	96,7	98,0	97,7	98,3	101,0	97,2	95,2	90,6	95,6
	III.	96,2	97,2	96,3	97,3	98,8	95,7	96,4	98,4	94,9	99,3	98,1	97,4	95,1	95,7	92,6
	IV.	99,3	99,4	99,5	99,5	99,0	97,6	98,5	99,7	101,1	99,4	98,1	99,7	96,4	96,1	99,1
	V.	98,6	98,3	98,8	98,5	98,3	96,9	99,9	100,5	102,1	100,3	95,6	98,4	98,0	95,2	99,9
	VI.	100,9	100,5	100,7	100,1	100,2	101,7	100,8	99,5	99,8	100,1	95,8	100,2	102,6	108,6	102,6
	VII.	100,9	101,3	101,1	101,1	100,5	99,3	100,6	101,9	98,9	102,2	109,8	101,3	99,6	99,8	99,4
	VIII.	99,7	99,6	99,9	99,7	100,3	96,8	99,0	101,5	103,5	101,3	97,8	99,8	98,9	97,5	100,4
	IX.	101,9	102,0	101,5	101,5	100,7	104,0	101,4	100,4	100,7	100,1	96,0	101,7	105,0	106,8	101,4
	X.	102,3	101,7	102,2	101,4	102,5	103,9	102,9	99,2	102,3	98,6	97,4	101,6	104,2	100,5	104,9
	XI.	104,1	103,9	103,8	103,5	104,0	107,1	106,9	100,6	103,7	100,0	97,8	103,8	107,1	100,7	104,9
	XII.	104,2	103,9	103,6	103,0	99,9	110,7	102,9	100,9	101,5	100,4	103,8	103,0	109,9	116,1	105,8



Example (production indices enhanced)

	2005-2010	2009 - 3	2009 - 4	2009 - 5	2009 - 6	2009 - 7	2009 - 8	2009 - 9	2009 - 10	2009 - 11	2009 - 12	2010 - 1	2010 - 2
BbisF		103	101.7	100.5	101.6	102.6	102.2	103.5	105.2	104.8	104.3	104.2	101.6
BbisE		102.1	99.8	99.2	100.2	101.8	100.9	102.6	104.8	103.8	103.6	103.6	101.4
EU_BCDF		102.7	101.1	99.8	100.9	102	101.9	102.4	104.5	104	103.9	105.1	101.4
EU_Industrie		101.2	98.8	98.4	99.1	100.9	100.7	100.8	104.1	102.6	103.4	104	100.6
Vorleistungen		98.2	92.8	93.4	95.2	96.6	96	98.7	98.2	100.2	99.6	99.6	98.6
Energie		119.9	121.3	118.7	118.9	118.4	115.6	117.5	119.8	118.8	118	114.9	115.4
Investitionsgueter		98.6	97.7	96.9	98.1	97.8	98.5	99.2	99.7	104.5	103	97.3	99.3
Konsumgueter		105.7	104.9	104	103.3	103.4	103.4	104.4	105.1	104.1	105.2	104.4	105.2
Lang_Konsumg		122.9	117.7	121.7	107.7	109.4	104.1	103.1	105.3	108.6	110.7	109.9	107.4
Kurz_Konsumg		103.1	103.1	101.5	102.6	102.1	103.3	104.3	104.5	103	103.7	103.2	103.8
Bergbau		103.3	105	108.3	114.5	115	111	110.1	108.7	111.9	102.8	98.6	105.7
HerstellungC		99.9	97.8	96.7	98	99.8	98.9	100.3	102.8	102.4	101.8	102	99.1
D		118.1	112.1	114.9	112	111.5	111.7	116.5	118.8	111.3	116.9	114.9	115.3
Wasser		122.4	132.4	130.6	132.9	135	134.6	133	128.1	125.8	124.3	128.8	134.2
Bau		107.3	109	105.9	107.6	106.7	107.6	107.7	107.4	109.3	107.3	106.2	102.7

Example (developments of the Austrian population)









Bevölkerung zu Jahresbeginn seit 1981 nach Geschlecht bzw. breiten Altersgruppen (Absolutwerte)

Jahr	Insgesamt	Nach Geschlecht		Männer auf 1.000 Frauen	Nach Altersgruppen			
		Männer	Frauen		0 bis 19 Jahre	20 bis 64 Jahre	65 Jahre und älter	dar.: 75 Jahre und älter
1981	7.553.326	3.570.172	3.983.154	896	2.184.224	4.212.971	1.156.131	454.278
1982	7.584.094	3.590.286	3.993.808	899	2.159.778	4.292.823	1.131.493	465.300
1983	7.564.185	3.582.589	3.981.596	900	2.115.305	4.348.057	1.100.823	473.838
1984	7.559.635	3.583.422	3.976.213	901	2.070.767	4.415.758	1.073.110	480.749
1985	7.563.233	3.588.116	3.975.117	903	2.028.352	4.465.937	1.068.944	491.279
1986	7.566.736	3.594.380	3.972.356	905	1.988.702	4.499.348	1.078.686	500.239
1987	7.572.852	3.602.199	3.970.653	907	1.950.892	4.528.383	1.093.577	508.013
1988	7.576.319	3.608.710	3.967.609	910	1.911.761	4.553.802	1.110.756	519.409
1989	7.594.315	3.623.136	3.971.179	912	1.879.112	4.589.333	1.125.870	527.740
1990	7.644.818	3.654.915	3.989.903	916	1.862.258	4.642.719	1.139.841	534.306
1991	7.710.882	3.696.200	4.014.682	921	1.856.653	4.700.847	1.153.382	526.559
1992	7.798.899	3.746.551	4.052.348	925	1.864.333	4.770.187	1.164.379	511.086
1993	7.882.519	3.793.245	4.089.274	928	1.876.578	4.831.640	1.174.301	494.349
1994	7.928.746	3.820.889	4.107.857	930	1.880.290	4.862.793	1.185.663	479.964
1995	7.943.489	3.831.200	4.112.289	932	1.875.112	4.871.503	1.196.874	481.743
1996	7.953.067	3.836.950	4.116.117	932	1.871.831	4.873.219	1.208.017	494.972
1997	7.964.966	3.844.019	4.120.947	933	1.870.818	4.877.700	1.216.448	511.436
1998	7.971.116	3.848.305	4.122.811	933	1.866.873	4.880.028	1.224.215	528.564
1999	7.982.461	3.856.029	4.126.432	934	1.862.619	4.890.127	1.229.715	545.049
2000	8.002.186	3.868.331	4.133.855	936	1.857.356	4.911.163	1.233.667	559.914
2001	8.020.946	3.881.104	4.139.842	938	1.844.074	4.938.856	1.238.016	575.493
2002	8.063.640	3.906.734	4.156.906	940	1.827.823	4.986.599	1.249.218	593.437
2003	8.100.273	3.929.599	4.170.674	942	1.819.450	5.030.344	1.250.479	601.901
2004	8.142.573	3.952.600	4.189.973	943	1.813.186	5.068.488	1.260.899	612.140
2005	8.201.359	3.984.866	4.216.493	945	1.809.717	5.083.697	1.307.945	625.028
2006	8.254.298	4.014.344	4.239.954	947	1.803.687	5.093.024	1.357.587	638.263
2007	8.282.984	4.030.062	4.252.922	948	1.790.880	5.093.505	1.398.599	648.843
2008	8.318.592	4.048.633	4.269.959	948	1.777.869	5.115.684	1.425.039	658.531
2009	8.355.260	4.068.047	4.287.213	949	1.763.948	5.140.425	1.450.887	665.415

Q: STATISTIK AUSTRIA, Statistik des Bevölkerungsstandes.- Revidierte Ergebnisse für 2002 bis 2008. Erstellt am: 27.05.2009.

Example (developments of the Austrian population, enhanced)

Bevölkerung zu Jahresbeginn seit 1981 nach Geschlecht bzw. breiten Altersgruppen (Überblick)





















	1981-2009	1981	2009	Minimum	Maximum
Insgesamt		7.5533.26	8.355.260	7.553.326	8.355.260
Männer		3.570.172	4.068.047	3.570.172	4.068.047
Frauen		3.983.154	4.287.213	3.967.609	4.287.213
Männer auf 1.000 Frauen		896	949	896	949
0-19 Jahre		2.184.224	1.763.948	1.763.948	2.184.224
20-64 Jahre		4.212.971	5.140.425	4.212.971	5.140.425
65+ Jahre		1.156.131	1.450.887	1.068.944	1.450.887
75+ Jahre		454.278	665.415	454.278	665.415

Bevölkerung zu Jahresbeginn seit 1981 nach Geschlecht bzw. breiten Altersgruppen (Absolutwerte)

Jahr	Geschlecht			Männer auf 1.000 Frauen	Altersgruppen			
	Insgesamt	Männer	Frauen		0-19 Jahre	20-64 Jahre	65+ Jahre	75+ Jahre
2009	8.355.260	4.068.047	4.287.213	949	1.763.948	5.140.425	1.450.887	665.415
2008	8.318.592	4.048.633	4.269.959	948	1.777.869	5.115.684	1.425.039	658.531
2007	8.282.984	4.030.062	4.252.922	948	1.790.880	5.093.505	1.398.599	648.843
2006	8.254.298	4.014.344	4.239.954	947	1.803.687	5.093.024	1.357.587	638.263
2005	8.201.359	3.984.866	4.216.493	945	1.809.717	5.083.697	1.307.945	625.028
2004	8.142.573	3.952.600	4.189.973	943	1.813.186	5.068.488	1.260.899	612.140
...

Example (team performance)

League Table - Premier League 2011/12

	Games	Performance	Wins	Draws	Losses	Points
Man Utd	38		23	11	4	80
Chelsea	38		21	8	9	71
Man City	38		21	8	9	71
Arsenal	38		19	11	8	68
Tottenham	38		16	14	8	62
Liverpool	38		17	7	14	58
Everton	38		13	15	10	54
Fulham	38		11	16	11	49
A. Villa	38		12	12	14	48
Sunderland	38		12	11	15	47
WBA	38		12	11	15	47
Newcastle	38		11	13	13	46
Stoke	38		13	7	18	46
Bolton	38		12	10	16	46
Blackburn	38		11	10	17	43
Wigan	38		9	15	14	42
Wolves	38		11	7	20	40
Birmingham	38		8	15	15	39
Blackpool	38		10	9	19	39
West Ham	38		7	12	19	33



Optimal Allocation with Sparklines

IS Population : Debt :		NO Population : Debt :	SE Population : Debt :	FI Population : Debt :	EE Population : Debt :
	NL Population : Debt :	DE Population : Debt :	DK Population : Debt :	LT Population : Debt :	LV Population : Debt :
UK Population : Debt :	IE Population : Debt :	BE Population : Debt :	CZ Population : Debt :	PL Population : Debt :	SK Population : Debt :
LU Population : Debt :	CH Population : Debt :	LI Population : Debt :	AT Population : Debt :	HU Population : Debt :	RO Population : Debt :
PT Population : Debt :	FR Population : Debt :	SI Population : Debt :	HR Population : Debt :	BG Population : Debt :	TR Population : Debt :
	ES Population : Debt :	IT Population : Debt :	MT Population : Debt :	GR Population : Debt :	CY Population : Debt :

R-package `sparkTable` (available on CRAN) provides the following features¹

- ▶ Different graphical representations
- ▶ Combining numerical with graphical information into a table
- ▶ Creation of tables for online- or print publication

¹for details, see: *The Visual Display of Quantitative Information*, E. R. Tufte

- ▶ **newSparkBar(), newSparkBox(), newSparkLine():** Functions to create new Spark object
- ▶ **newSparkTable():** Function to create new SparkTable object
- ▶ **newGeoTable():** Functions to create a new object of class 'geoTable'
- ▶ **plotSparks(), plotSparkTable(), plotGeoTable():** Generate Output from Spark, SparkTable and GeoTable objects
- ▶ **getParameter(), setParameter():** Functions to view or change the parameters of the object

- ▶ We show a very simple example to create a single spark-graph
- ▶ Generating random input-data:

```
dat <- rnorm(25, 100, 25)
```

- ▶ Creating a suitable input-object and setting several graphic parameters:



```
a <- newSparkLine(values=dat, pointWidth=8)
```


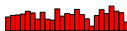

- ▶ Drawing the *png*-image

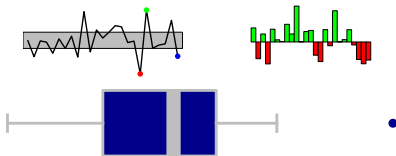
```
plotSparks(a, outputType='png',  
  filename='testLine1')
```

- ▶ Changing parameters and drawing the *pdf*-image

```
a <- setParameter(a, c("darkred", "darkgreen",  
  "darkblue", "white", "black", "red"),  
  type='allColors')  
getParameter(a, type='allColors')  
a <- setParameter(a, 3, type='pointWidth')  
a <- setParameter(a, 1, type='lineWidth')  
plotSparks(a, outputType="pdf",  
  filename='testLine2')
```

- ▶ Displaying the image as pdf  (scalable!)
- ▶ Displaying the image as png  (not scalable!)

























- ▶ Creation of customized time-series, box- and barplots
 - ▶ Time series plots 
 - ▶ Bar plots 
 - ▶ Box plots 
- ▶ Graphical options include (among others)
 - ▶ Highlighting of specific values in time series plots
 - ▶ Choice of colors
 - ▶ Choice of boundaries
 - ▶ Inclusion of interquartile-range in time-series plots



The user can decide to generate *pdf*, *eps* or *png* output

- ▶ *Pdf*, *Eps* graphics can be included in \LaTeX for scientific use
- ▶ *Png* can be used to show both graphics on websites

Example (population) I

	Mean value	Boxplot	LinePlot	BarPlot	Current value
Total	7884750.66				8355260
Men	3791573.55				4068047
Women	4093177.1				4287213
Men per 1.000 Women	925.97				949
Age 0-19	1899456.83				1763948
Age 20-64	4778240				5140425
Age 65+	1207053.83				1450887
Age 75+	541649.59				665415

► The data:

```
data(pop)
head(pop,5)
  time variable  value
1 1981 Ingesamt 7553326
2 1982 Ingesamt 7584094
3 1983 Ingesamt 7564185
4 1984 Ingesamt 7559635
5 1985 Ingesamt 7563233
```

- ▶ Create a content object:

```
content <- list()
content[['Mean value']] <- function(x) { round(mean(x),2) }
content[['Boxplot']] <- newSparkBox()
content[['LinePlot']] <- newSparkLine()
content[['BarPlot']] <- newSparkBar()
content[['Current value']] <- function(x) { round(tail(x,1),2) }
```

- ▶ Creating a sparkTable object:

```
varType <- rep("value",5)
pop <- pop[,c("variable","value","time")]
pop$time <- as.numeric(as.character(pop$time))
dat <- reshapeExt(pop, idvar="variable", varying=list(2))
sparkTab <- newSparkTable(dat, content, varType)
```

- ▶ Create the HTML output:





















```
plotSparkTable(sparkTab, outputType="html", filename="t1")
```

- ▶ Create the TEX output:

```
plotSparkTable(x1, outputType="tex", filename="t2")
```

Example (team performance) I

League Table - Premier League 2011/12

	Games	Performance	Wins	Draws	Losses	Points
Man Utd	38		23	11	4	80
Chelsea	38		21	8	9	71
Man City	38		21	8	9	71
Arsenal	38		19	11	8	68
Tottenham	38		16	14	8	62
Liverpool	38		17	7	14	58
Everton	38		13	15	10	54
Fulham	38		11	16	11	49
A. Villa	38		12	12	14	48
Sunderland	38		12	11	15	47
WBA	38		12	11	15	47
Newcastle	38		11	13	13	46
Stoke	38		13	7	18	46
Bolton	38		12	10	16	46
Blackburn	38		11	10	17	43
Wigan	38		9	15	14	42
Wolves	38		11	7	20	40
Birmingham	38		8	15	15	39
Blackpool	38		10	9	19	39
West Ham	38		7	12	19	33

- ▶ The data:

```
head(soccer[,1:19]) # first half of the season
  teams 1  2  3 4  5  6  7  8  9 10 11 12 13 14 15 16 17 18
  Man Utd 1  0  1 0  1  0  0  0  1  1  1  0  0  1  1  1  1 -1
  Chelsea 1  1  1 1  1 -1  1  0  1  1 -1  1 -1 -1  0  0  0  1
  Man City 0  1 -1 0  1  1  1  1 -1 -1  1  0  0  1  0  1  1 -1
  Arsenal 0  1  1 1  0 -1 -1  1  1  1 -1  1  1 -1  1  1 -1  1
  Tottenham 0  1 -1 0  1 -1  1  1  0 -1 -1  0  1  1  1  0  0 -1
  Liverpool 0 -1  1 0 -1  0 -1 -1  1  1  1  0 -1  1 -1  1 -1  1
```

- ▶ reshaping the data:

```
nrGamedays <- ncol(soccer)-1
dat <- reshapeExt(soccer,
  idvar='team', v.names=c('result'), varying=list(2:(nrGamedays+1)),
  timeValues=1:nrGamedays
)
```

- ▶ Create a content object:

```
b <- newSparkBar()
b <- setParameter(b, c("red","green","black"), type="barCol")

content <- list()
content[['Games']] <- function(x) { length(x) }
content[['Performance']] <- b
content[['Wins']] <- function(x) { length(which(x==1)) }
content[['Draws']] <- function(x) { length(which(x==0)) }
content[['Losses']] <- function(x) { length(which(x== -1)) }
content[['Points']] <- function(x) {
  length(which(x==1))*3 + length(which(x==0)) }
}
```

- ▶ Creating a sparkTable object:

```
obj <- newSparkTable(dat, content, varType=rep("result",6))
```

- ▶ Creating the TEX output:

```
plotSparkTable(obj, outputType="tex", filename="tabSoccer")
```

Example - Optimal Allocation with Sparklines

IS Population : Debt :		NO Population : Debt :	SE Population : Debt :	FI Population : Debt :	EE Population : Debt :
	NL Population : Debt :	DE Population : Debt :	DK Population : Debt :	LT Population : Debt :	LV Population : Debt :
UK Population : Debt :	IE Population : Debt :	BE Population : Debt :	CZ Population : Debt :	PL Population : Debt :	SK Population : Debt :
LU Population : Debt :	CH Population : Debt :	LI Population : Debt :	AT Population : Debt :	HU Population : Debt :	RO Population : Debt :
PT Population : Debt :	FR Population : Debt :	SI Population : Debt :	HR Population : Debt :	BG Population : Debt :	TR Population : Debt :
	ES Population : Debt :	IT Population : Debt :	MT Population : Debt :	GR Population : Debt :	CY Population : Debt :

► Data Preparation:

```
data(popEU, package="sparkTable")
data(debtEU, package="sparkTable")
data(coordsEU, package="sparkTable")
popEU <- popEU[popEU$country%in%coordsEU$country,]
debtEU <- debtEU[debtEU$country%in%coordsEU$country,]
EU <- cbind(popEU,debtEU[,-1])
EUlong <- reshapeExt(EU, idvar="country",
  v.names=c("pop", "debt"), varying=list(2:13,14:25),
  geographicVar="country", timeValues=1999:2010)
```


- ▶ Defining the content and generating the geoTable object:

```
l <- newSparkLine()  
l <- setParameter(l, 'lineWidth', 2.5)  
content <- list(function(x){"Population:"},1,  
  function(x){"Debt:"},1)  
varType <- c(rep("pop",2),rep("debt",2))  
xGeoEU <- newGeoTable(EUlong, content, varType,  
  geographicVar="country",geographicInfo=coordsEU)
```

- ▶ Writing the geoTable in a tex-File:

```
plotGeoTable(xGeoEU, outputType="tex",  
graphNames="outEU",filename="testEUT",  
transpose=TRUE)
```

- ▶ Package `sparkTable` is at an early stage of development
- ▶ some improvements (e.g. additional customization options, additional graphs) are already planned to be implemented
- ▶ User-feedback is (of course) very welcome