

Statisztika

Politológus képzés

Daróczi Gergely

Politológia Tanszék

2011. május 4.

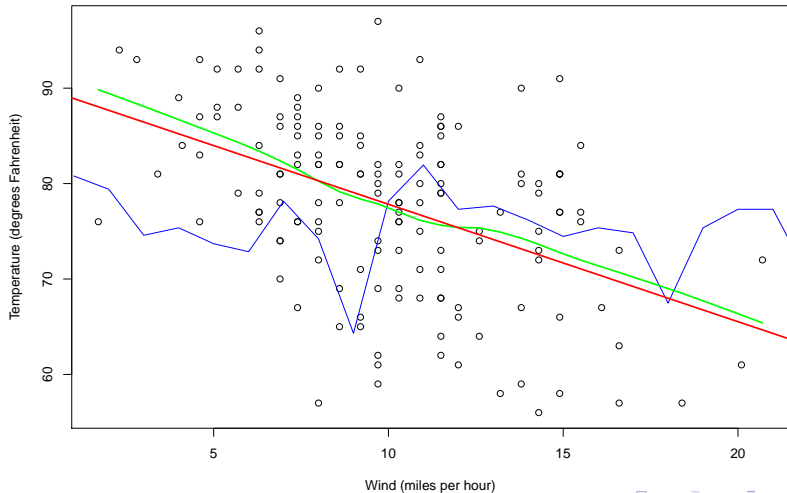


PÁZMÁNY PÉTER
KATOLIKUS EGYETEM
Bölcsészettudományi Kar

- 1 Korreláció
- 2 Standardizálás és dekompozíció
- 3 Grafikonok, ábrák

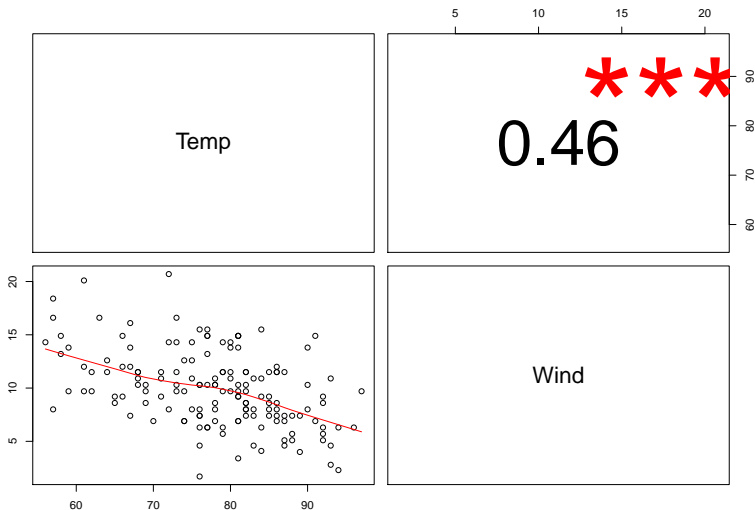
Korreláció

Valós kapcsolat?



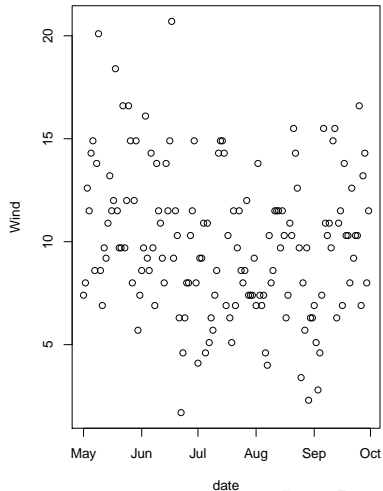
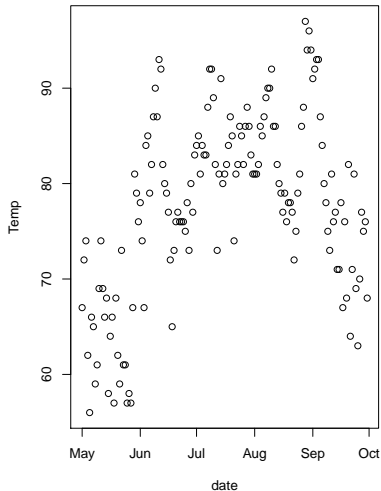
Korreláció

Valós kapcsolat?



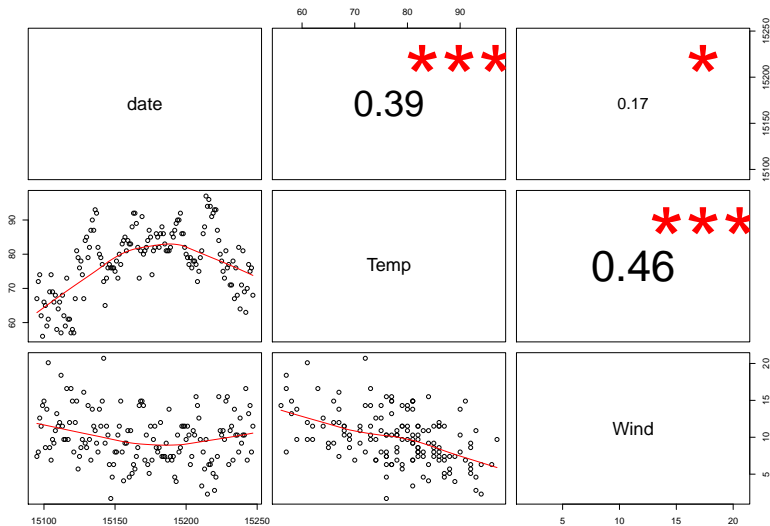
Korreláció

Valós kapcsolat?



Korreláció

Valós kapcsolat?



Kettős vak, randomizált kísérletet végeztünk végzős orvostanhallgatók körében.

A kísérleti gyógyszert, ill. placebo-t fogyasztó alanyok egy éven keresztül minden nap benették a kiosztott pirulákat, majd a kísérlet végeztével az infarktusra való hajlamot vizsgálták:

	szívroham	nincs szívroham
Aspirin	104	10933
placebo	189	10845

Keresztábra

Egy klinikai példa

	szívroham	nincs szívroham	Σ	ráta (1000 fő)
Aspirin	104	10 933	11 937	9.4
Placebo	189	10 845	11 034	17.1
Σ	293	21 778	22 071	13.3

	szívroham	nincs szívroham	Σ
Aspirin	0,9 %	99,1 %	100 %
Placebo	1,7 %	98,3 %	100 %
Σ	1,3 %	98,7 %	100 %

Keresztábra

Egy klinikai példa

Heart Attack?

	Yes	No	Total
Aspirin	104	10933	11037
	146.52	10890.48	
Placebo	189	10845	11034
	146.48	10887.52	
Total	293	21778	22071

$$\text{Chi-Sq} = 12.339 + 0.166 + 12.343 + 0.166 = 25.014$$

$$\text{DF} = 1, \text{ P-Value} = 0.000$$

Keresztábla

Egy klinikai példa tizede

Heart Attack?

	Yes	No	Total
Aspirin	10	1093	1103
	14.49	1088.51	
Placebo	19	1085	1104
	14.51	1089.49	
Total	29	2178	2207

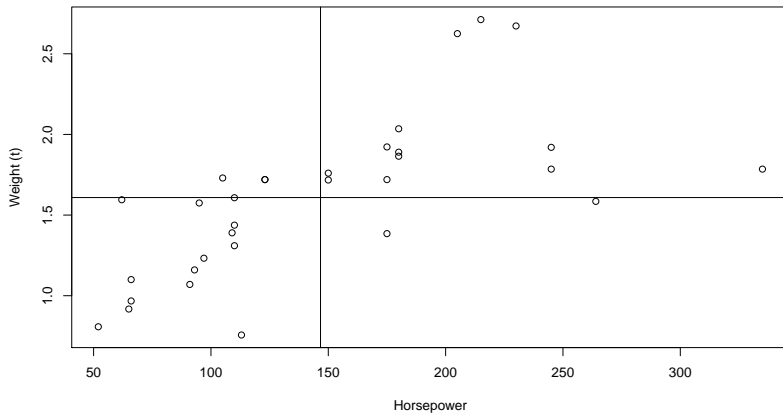
$$\text{Chi-Sq} = 1.393 + 0.019 + 1.392 + 0.019 = 2.822$$

$$\text{DF} = 1, \text{ P-Value} = 0.093$$

Standardizálás és dekompozíció

Egy egyszerű példán keresztül bemutatva

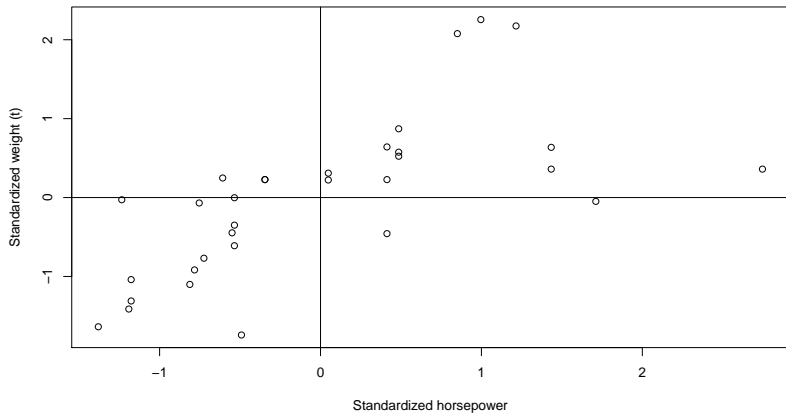
**Henderson & Velleman (1981):
Building multiple regression models interactively**



Standardizálás és dekompozíció

Egy egyszerű példán keresztül bemutatva

Henderson & Velleman (1981):
Building multiple regression models interactively

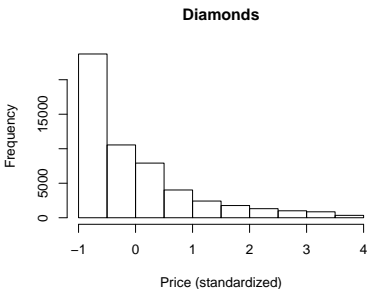
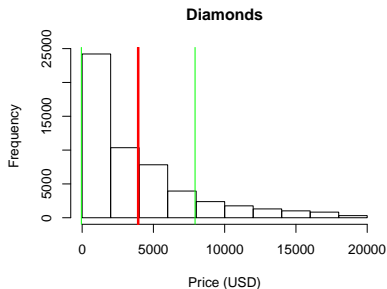


Standardizálás és dekompozíció

A standardizálás egyszerű esete

Normált érték (z-values, z-scores, normal scores, standardizált változó) alatt azt értjük, hogy a megfigyelések az átlagtól hány szórásnyira esnek:

$$z = \frac{x - \mu}{\sigma}$$



Standardizálás és dekompozíció

Dekompozíció

**Population and Deaths by Age in 1970 for White Females in
Miami, Alaska, and the U.S.**

Age	Miami			Alaska			U.S.		
	Pop.	Deaths	Rate*	Pop.	Deaths	Rate*	Pop.+	Deaths+	Rate*
< 15	114,350	136	1.19	37,164	59	1.59	23,961	32	1.34
15-24	80,259	57	0.71	20,036	18	0.90	15,420	9	0.58
25-44	133,440	208	1.56	32,693	37	1.13	21,353	30	1.40
45-64	142,670	1,016	7.12	14,947	90	6.02	19,609	140	7.14
65+	92,168	3,605	39.11	2,077	81	39.00	10,685	529	49.51
	562,887	5,022		106,917	285		91,028	740	
Crude death rate*			8.92			2.67			8.13

* Deaths per 1,000 population

+ in thousands

Meghatározás

Közvetlen transzformáció (direct standardization) során a rétegspecifikus arányokat, rátákat értelmezzük a teljes populációra jellemző rétegnagyságokon.

$$r\acute{a}ta = \frac{\sum \text{rétegspecifikus arány} \times \text{standard súlyok}}{\sum \text{standard súlyok}}$$

$$Miami = \frac{(1.19 \times 23,961) + \dots + (39.11 \times 10,685)}{91,208} = 6.92$$

$$Alaska = \frac{(1.59 \times 23,961) + \dots + (39 \times 10,685)}{91,208} = 6.71$$

Meghatározás

A közvetett transzformáció (indirect standardization) esetében pontosan fordítva, a teljes populációból származtatjuk az arányokat/rátákat, a súlyokat pedig a rétegek alapján számoljuk.

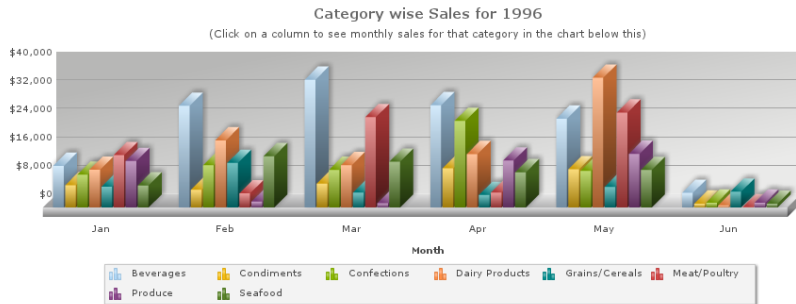
$$ráta = \frac{\sum \text{tapasztalt értékek}}{\sum \text{várt értékek}}$$

Várt értékek = Rétegspecifikus súlyok \times populációbeli arányok/ráta

	Minta	Alapsokaság
Directly-standardized rate	ráta	súly
Indirectly-standardized rate	súly	ráta

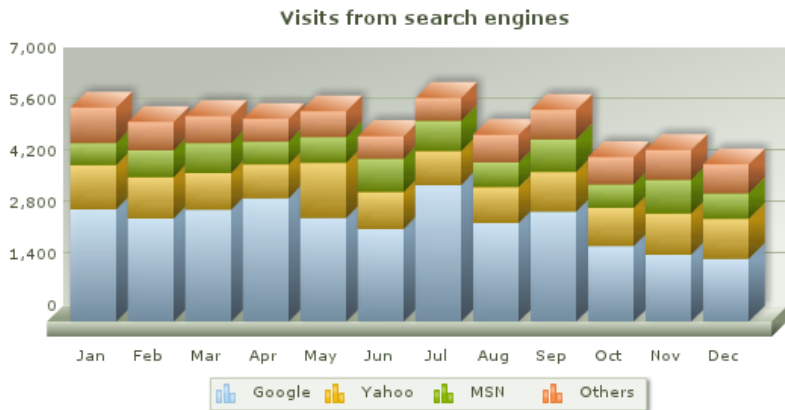
Grafikonok, ábrák

Csoportosított oszlopdiaagram



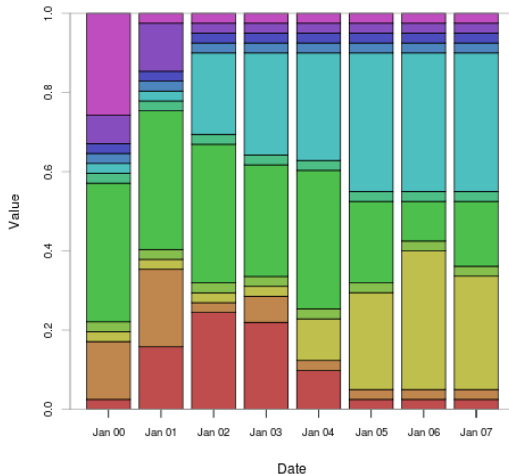
Grafikonok, ábrák

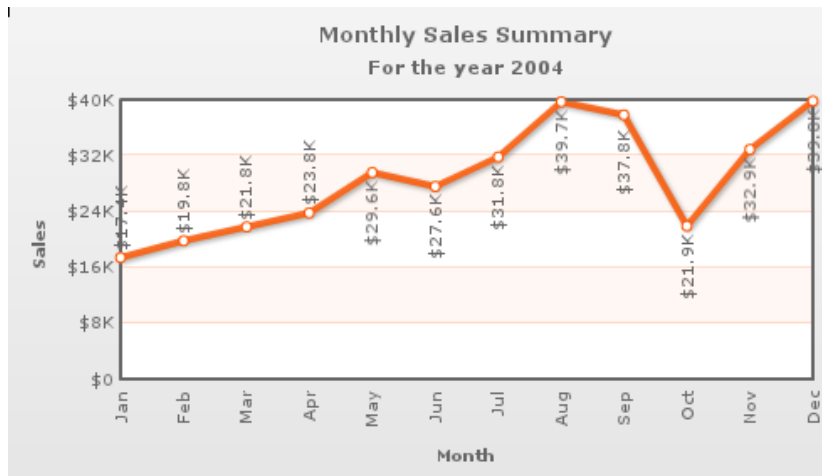
Rétegzett oszlopdiagram



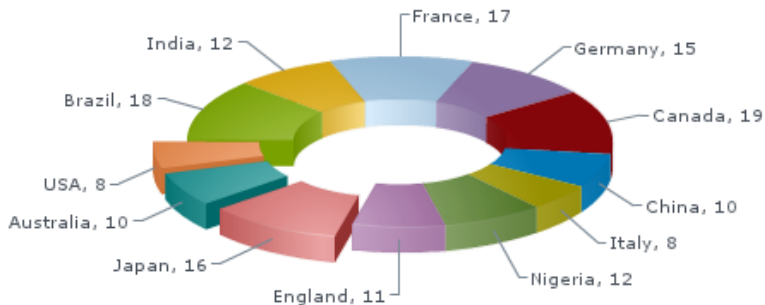
Grafikonok, ábrák

Rétegzett oszlopdiagram



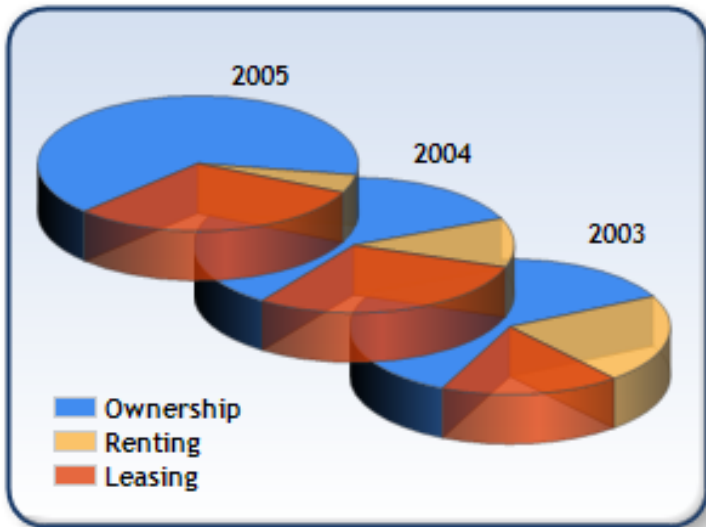


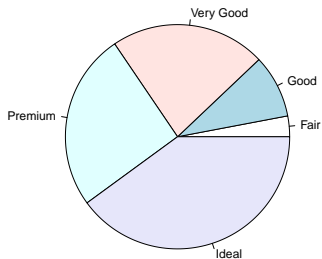
**Industrial Growth Rate
(Country)**



Grafikonok, ábrák

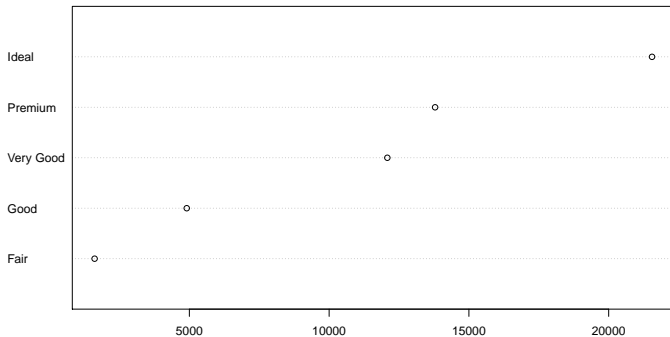
Kördiagram





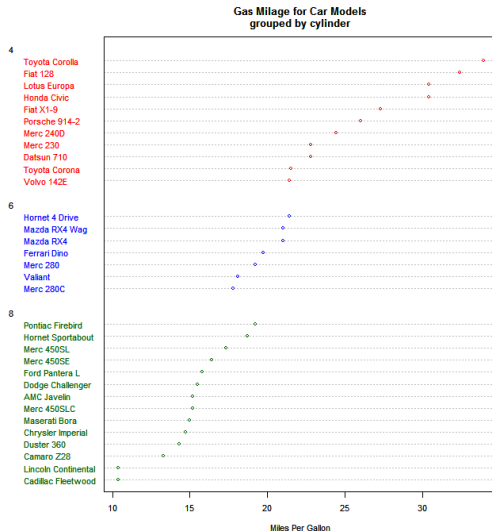
Grafikonok, ábrák

Kördiagram



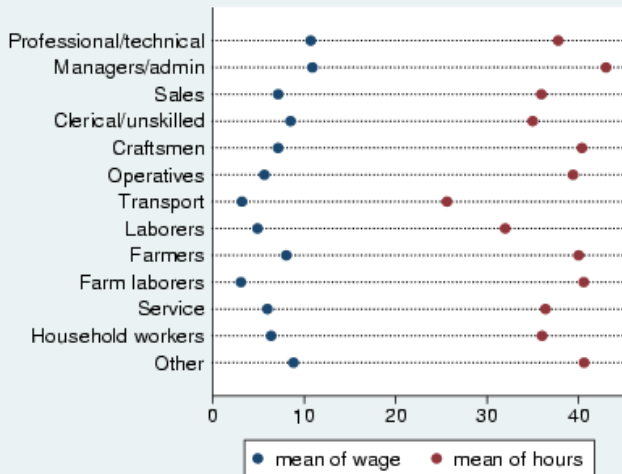
Grafikonok, ábrák

Dot plot



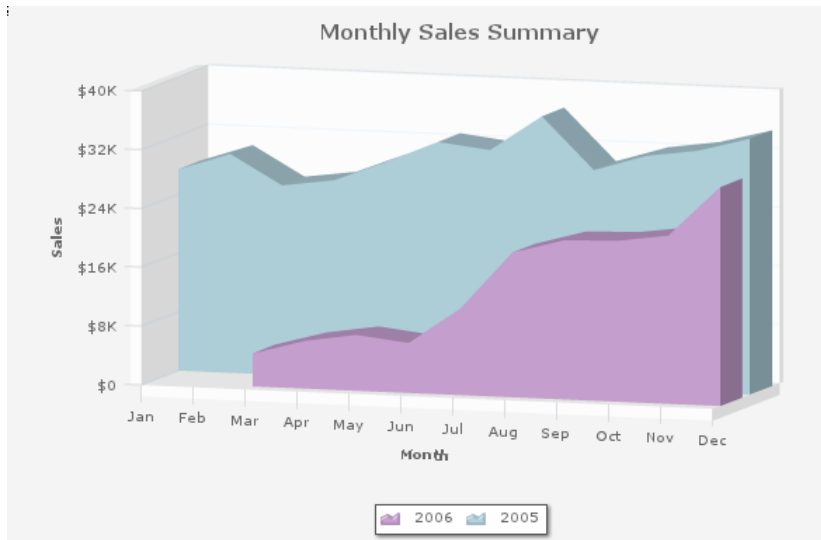
Grafikonok, ábrák

Dot plot

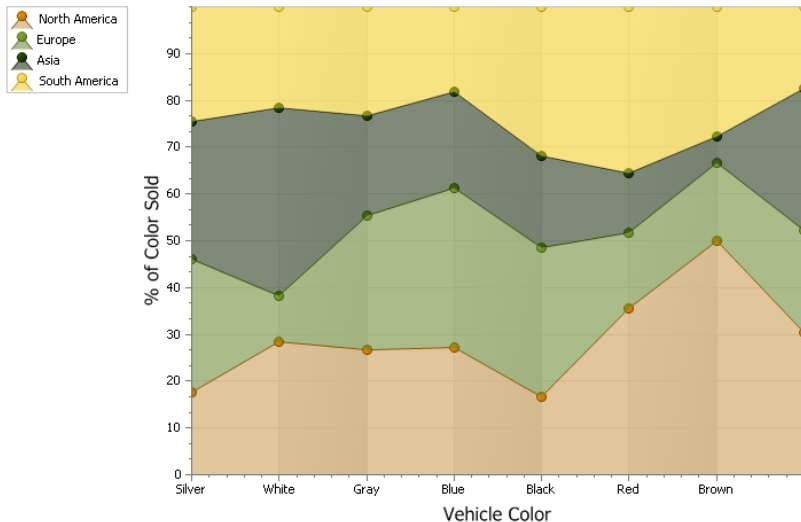


Grafikonok, ábrák

Terület



Worlds Most Popular Vehicle Colors



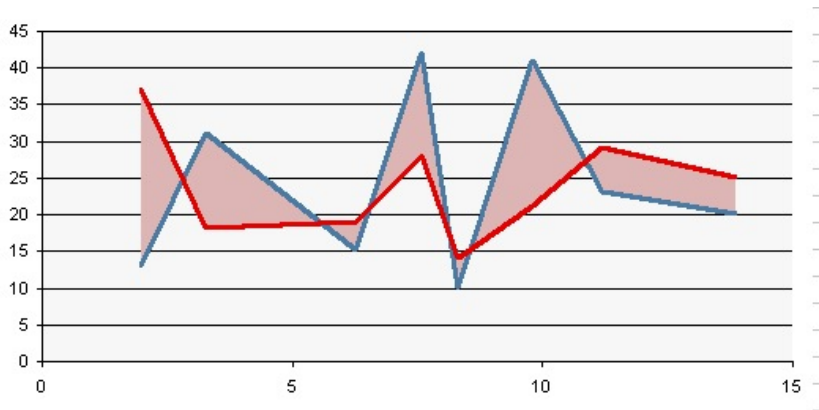
Grafikonok, ábrák

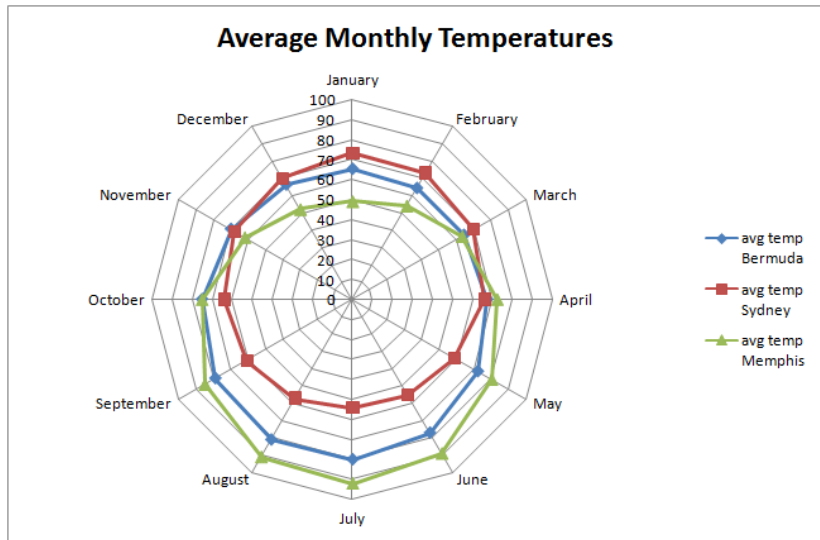
Összetett, kombinált grafikonok



Grafikonok, ábrák

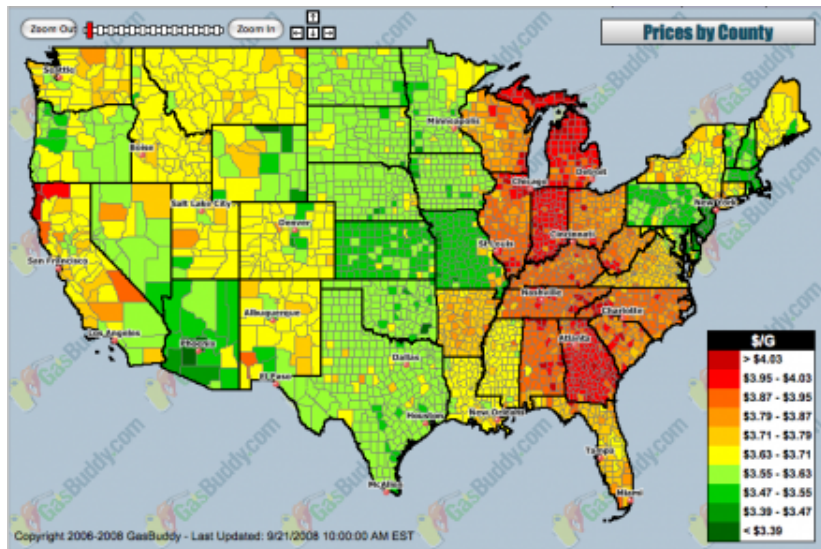
Összetett, kombinált grafikonok





Grafikonok, ábrák

Heatmap



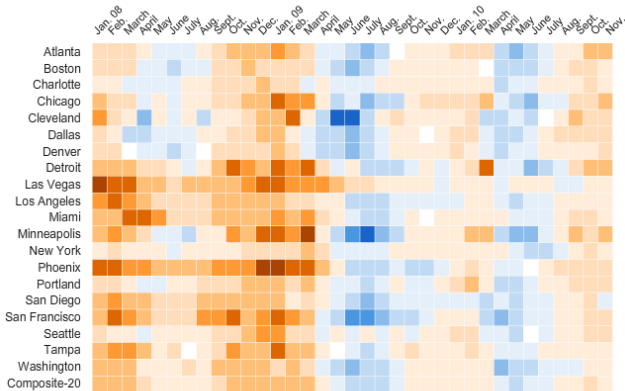
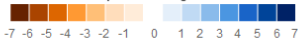
Home statistics: A national gauge of the housing market

HOME PRICE INDEX

HOME SALES

Track the month-to-month percent change in the Standard & Poor's/Case-Shiller 20-city home price index and the cities that comprise it.

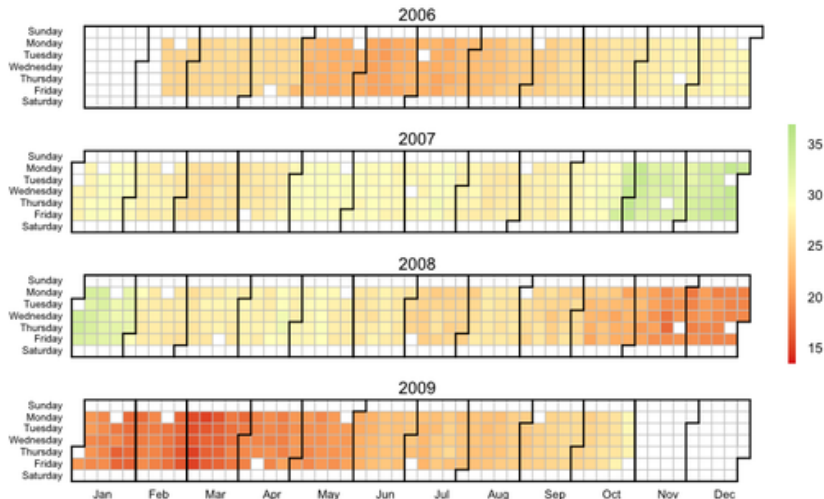
Month-to-month percent change:



Updated: Feb. 23, 2011

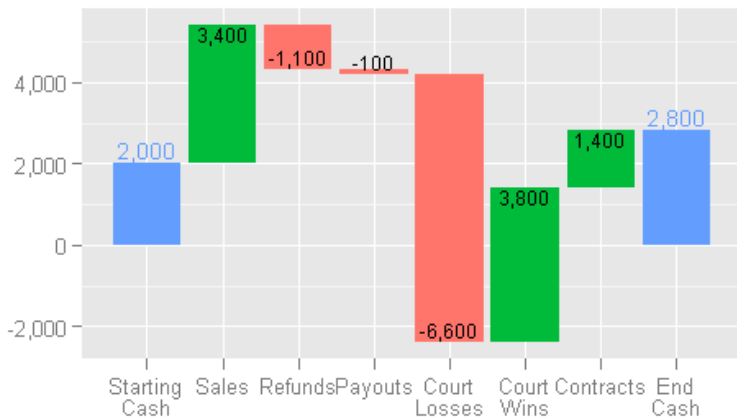
SOURCES

Calendar Heat Map of MSFT Adjusted Close



Grafikonok, ábrák

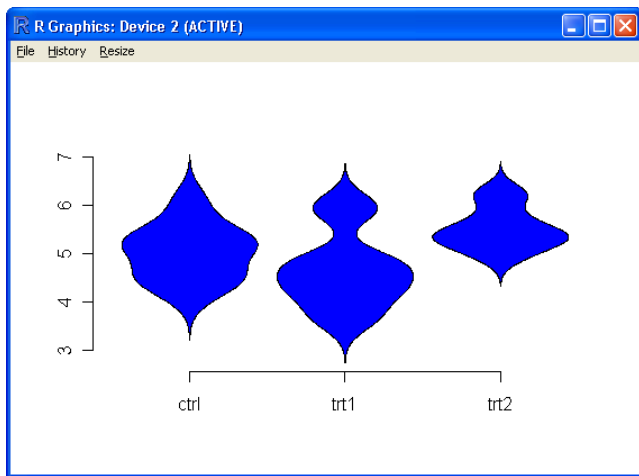
Waterfall





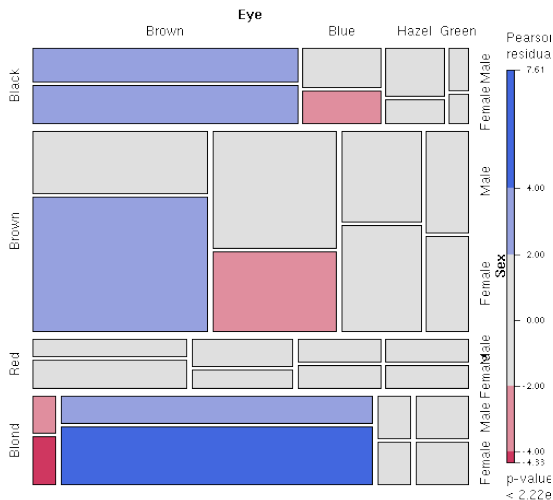
Grafikonok, ábrák

Violin plot



Grafikonok, ábrák

Mosaic chart



Szófelhő



„Crayola Color Chart, 1903-2010”



Grafikonok, ábrák

Érdekes, felkeresésre javasolt honlapok:

- http://www.visual-literacy.org/periodic_table/periodic_table.html
- <http://www.edwardtufte.com/tufte/>
- <http://www.perceptualedge.com/>
- <http://www.visualcomplexity.com/vc/>
- <http://flowingdata.com/>
- <http://infosthetics.com/>
- <http://chartsgraphs.wordpress.com/>
- <http://www.informationisbeautiful.net/>
- <http://chartporn.org/>

To be continued...

Daróczi Gergely
daroczi.gergely@btk.ppke.hu