

# PA055 Vizualizace komplexních dat

## Týden 1

Katedra informačních technologií  
Masarykova Univerzita Brno

Podzim 2013

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



# Outline

# Kontaktní údaje

- ▶ Dr. Matej Lexa, C506 (lexa@fi.muni.cz)
- ▶ Kurz: Út 10:00-12:00 (B103)
- ▶ Konzultace: Čt 13:00-15:00 (C506)

# Literatura zakoupena do knihovny

## OBECNĚ

- ▶ Chen et al. (2008). Handbook of data visualization Springer ISBN:978-3540330363
- ▶ Telea (2007). Data visualization AK Peters ISBN:978-1568813066

## PRO R

- ▶ Sarkar (2008). Lattice: Multivariate data visualization with R Springer ISBN:978-0387759685
- ▶ Maindonald and Braun (2006). Data analysis and graphics using R Cambridge Univ Press ISBN:978-0521861168

## PRO PROCESSING

- ▶ Fry (2008). Visualizing data: Exploring data with the processing environment O'Reilly ISBN:978-0596514556
- ▶ Greenberg (2007). Processing: Creative coding and computational art. Friends of ED ISBN:978-1590596173

# Studijní URL

- ▶ <https://is.muni.cz/auth/el/1433/podzim2012/PA055/um/McGarthwaite.pdf>  
(str.1-68)
- ▶ <http://www.wikiviz.org>
- ▶ <http://cran.r-project.org/>
- ▶ <http://www.ggobi.org/>
- ▶ <http://www.processing.org/>
- ▶ <http://had.co.nz/>

# Klasifikace

- ▶ Hodnotí se
  - ▶ Ukázka vizualizace 5 bodů
  - ▶ Cvičení 15 bodů
  - ▶ Semestrální skupinový projekt 30 bodů
  - ▶ Zkouška 50 bodů
- ▶ Klasifikační stupnice
  - ▶ A 91 - 100
  - ▶ B 81 - 90
  - ▶ C 71 - 80
  - ▶ D 61 - 70
  - ▶ E 51 - 60
  - ▶ F méně než 51

# Outline

# Definice

## Vizualizace

- ▶ Úzce chápáno:  
tvorba grafické reprezentace dat pro jejich pochopení
- ▶ Šířeji:  
tvorba mentální reprezentace dat pomocí grafiky

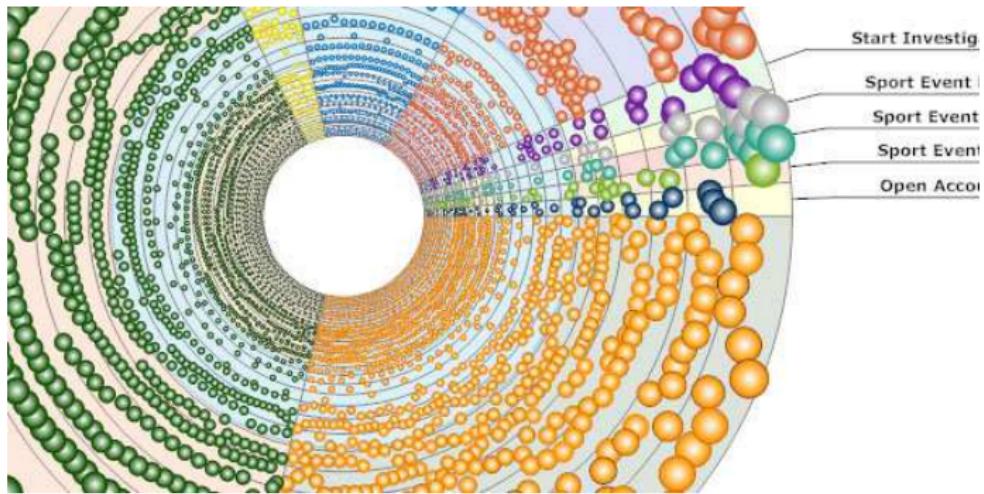
# Typy vizualizace

- ▶ Information visualization
- ▶ Knowledge visualization
- ▶ Scientific visualization
- ▶ Visual data-mining
- ▶ Object visualization
- ▶ Visual communication

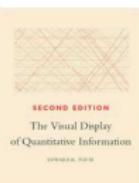
- ▶ Akvizice dat
- ▶ Analýza dat (parsing)
- ▶ Filtrace
- ▶ Dolování
- ▶ Reprezentace
- ▶ Interakce

## Cílem vizualizace je

- ▶ usnadnit porovnávání dat, rozeznávání vzorů a detekci změn v datech
  - ▶ spřístupnit komplikované sady dat lidskému vnímání



- ▶ J.H. Lambert - nahrazení tabulek grafikou
- ▶ William Playfair - souřadnicové grafy
- ▶ Jacques Bertin - klasifikace grafických elementů
- ▶ Edward Tufte - maximalizace hustoty užitečných informací



# Zdroje inspirace

- ▶ <http://www.visualcomplexity.com/vc>
- ▶ <http://personal.csics.ch/mvalle/visualization/tools.html>

Data set: Analysis U.S. Presidential Inaugural Addresses

Your visualization will look like this:

Select a phrase

Show 46 of 2125 terms

word1 and word2  
word1 's word2  
word1 of the word2  
word1 the word2  
word1 a word2  
word1 at word2  
word1 is word2  
word1 [space] word2

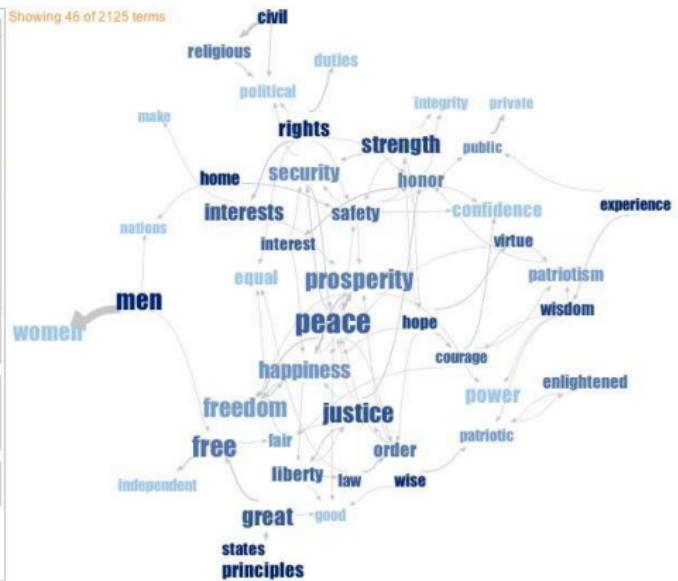
or enter your own  
" and " Submit

Filters

Show top: 50 Hide common words

Zoom

In Out Reset



<http://davidwarlick.com/2cents/>

# Typické prostředky informační grafiky (vizualizace dat)

- ▶ grafy
- ▶ mapy
- ▶ diagramy
- ▶ tabulky
- ▶ nákresy

# Proč vizualizace umožňuje lepší pochopení dat

- ▶ lokalita
- ▶ minimalizace textu
- ▶ vyvolání vjemů

- ▶ zvýrazňování
- ▶ ukrývání
- ▶ škálování
- ▶ společných/rozdílných
  - ▶ tvarech
  - ▶ barvě
  - ▶ orientaci
- ▶ projekci
- ▶ vrstvení
- ▶ změně pozice
- ▶ mikro/makro kombinace
- ▶ opakování
- ▶ časoprostorové ose

# Vizuální kódování dat

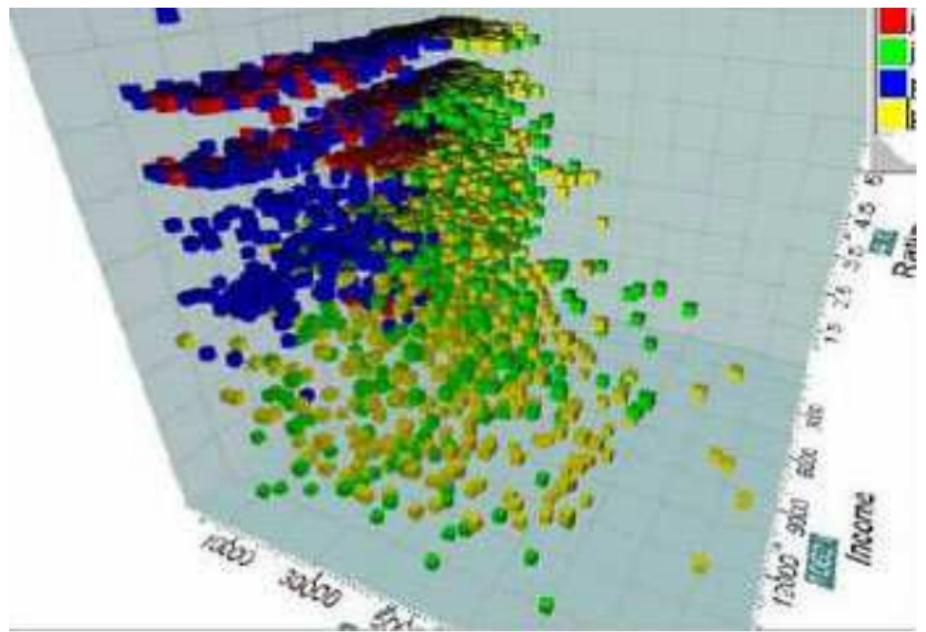
- ▶ základní prostor
  - ▶ nedělený
  - ▶ jmenný
  - ▶ pořadový
  - ▶ kvantitativní)
- ▶ grafické symboly
  - ▶ bod
  - ▶ čára
  - ▶ plocha (případně objem)
- ▶ vlastnosti symbolů
  - ▶ barva
  - ▶ velikost
  - ▶ orientace
  - ▶ intenzita
  - ▶ textura
  - ▶ tvar
  - ▶ pohyb (blikání, směr, rychlosť)
  - ▶ propojení
  - ▶ uzavření

# Kolik proměnných lze zobrazit na jednom diagramu?

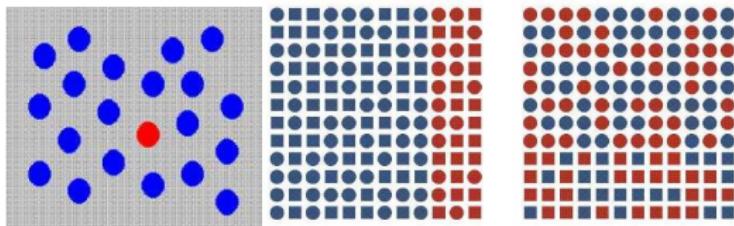
	Quantitative	Ordinal	Nominal
More Accurate			
Position	• :	Position	• :
Length	==	Density	● ● ○
Angle	<	Saturation	● ● ○
Slope	<	Hue	● ○ ○
Area	• ●	Length	==
Density	● ○ ○	Angle	<
Saturation	● ● ○	Slope	<
Hue	● ○ ○	Area	• ●
Shape	● ■ ■	Shape	● ■ ■
Less Accurate			

Figure 16. Visual Encoding Accuracy by Task type.

Kolik proměnných lze zobrazit na jednom diagramu?

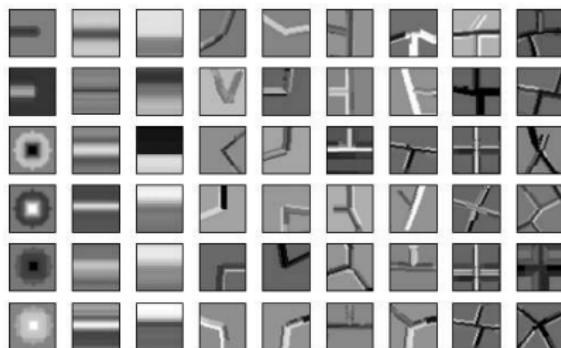


- ▶ rychlé zpracování nevyžadující soustředění
  - ▶ detekce objektu
  - ▶ detekce hranic
  - ▶ sledování pohybu
  - ▶ odhad počtu
- ▶ soustředěné zpracování



## textony

- ▶ tvar
- ▶ křížení
- ▶ ukončení



# Data

## 1. Nature of the data

- ▶ Numerical data
- ▶ Ordinal data
- ▶ Categorical data

## 2. Number of data dimensions

- ▶ Univariate data
- ▶ Bivariate data
- ▶ Trivariate data
- ▶ Multivariate data

## 3. Structure of the data

- ▶ Linear data
- ▶ Temporal data
- ▶ Spatial data (or geographic data)
- ▶ Hierarchical data
- ▶ Network data
- ▶ Textual data

# Outline

- ▶ grafické knihovny pro 2D a 3D grafiku
- ▶ graficky orientované jazyky a prostředí
- ▶ programy pro různé typy vizualizace
- ▶ univerzální vizualizační programy

# Outline

## For Further Reading

X