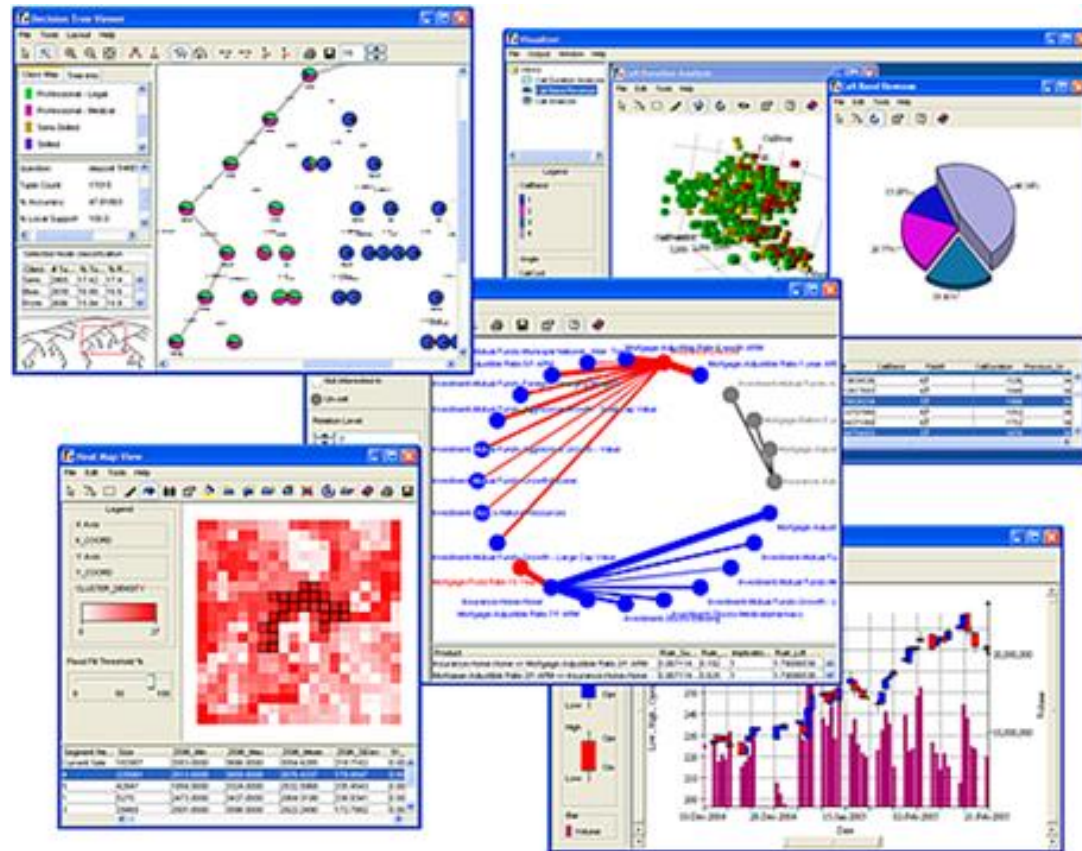


# Vizualizace dat



# Vizualizace - úvod



- ❑ Obrázek je cennější než tisíce slov!
- ❑ Vizualizace umožňuje snadnější pochopení dat!

# Vizualizace - úvod

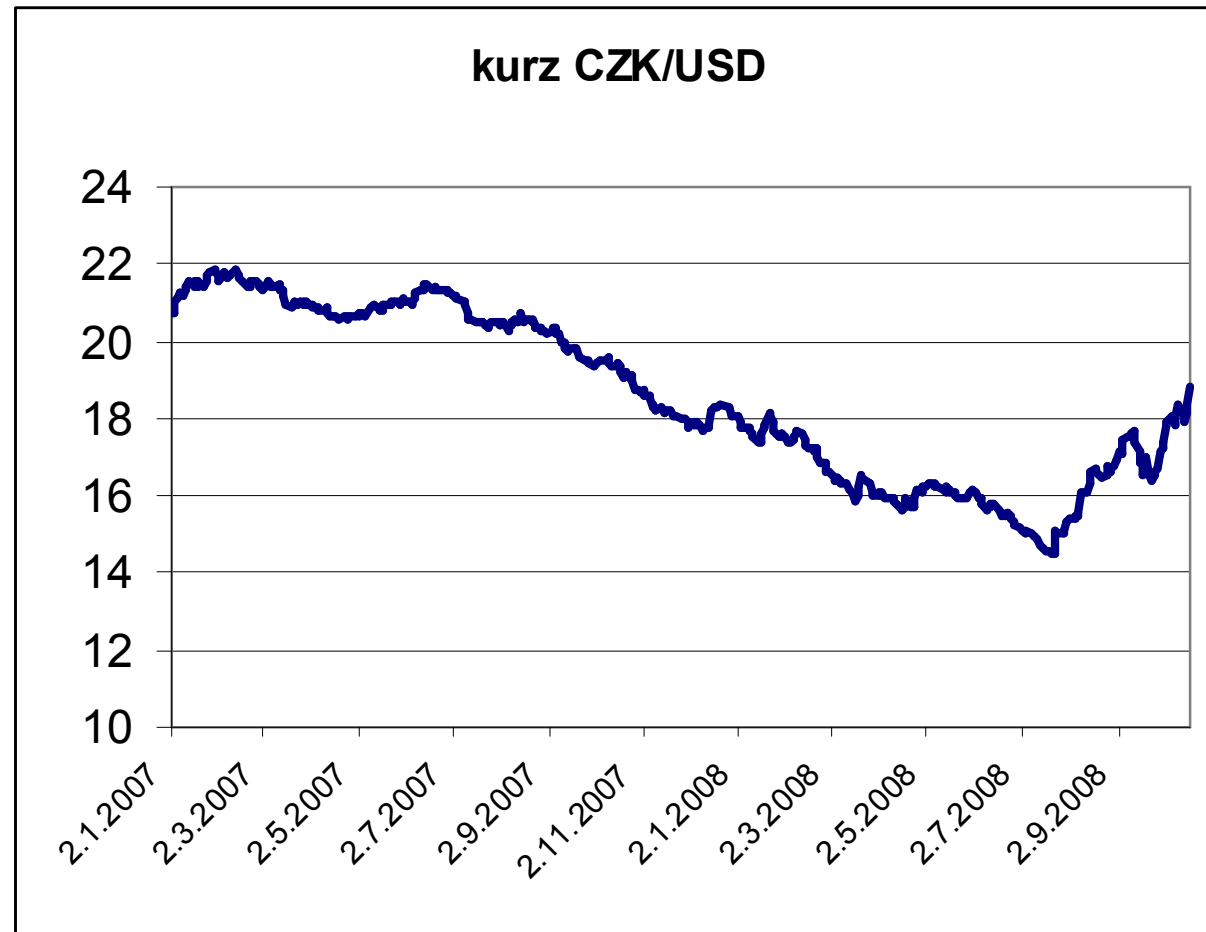
Kurz CZK/USD:

- 2.1.2007: 20,745
- 17.10.2008: 18,819

	1 USD
2.1.2007	20,745
3.1.2007	20,733
4.1.2007	21,064
5.1.2007	21,12
8.1.2007	21,256
9.1.2007	21,202
10.1.2007	21,35
11.1.2007	21,392
12.1.2007	21,53
15.1.2007	21,457
16.1.2007	21,434
17.1.2007	21,587
18.1.2007	21,554
19.1.2007	21,424
22.1.2007	21,495
23.1.2007	21,402
24.1.2007	21,579
25.1.2007	21,687
26.1.2007	21,811
29.1.2007	21,836
30.1.2007	21,814
31.1.2007	21,737

# Vizualizace - úvod

□ Obrázek poskytuje mnohem více informací



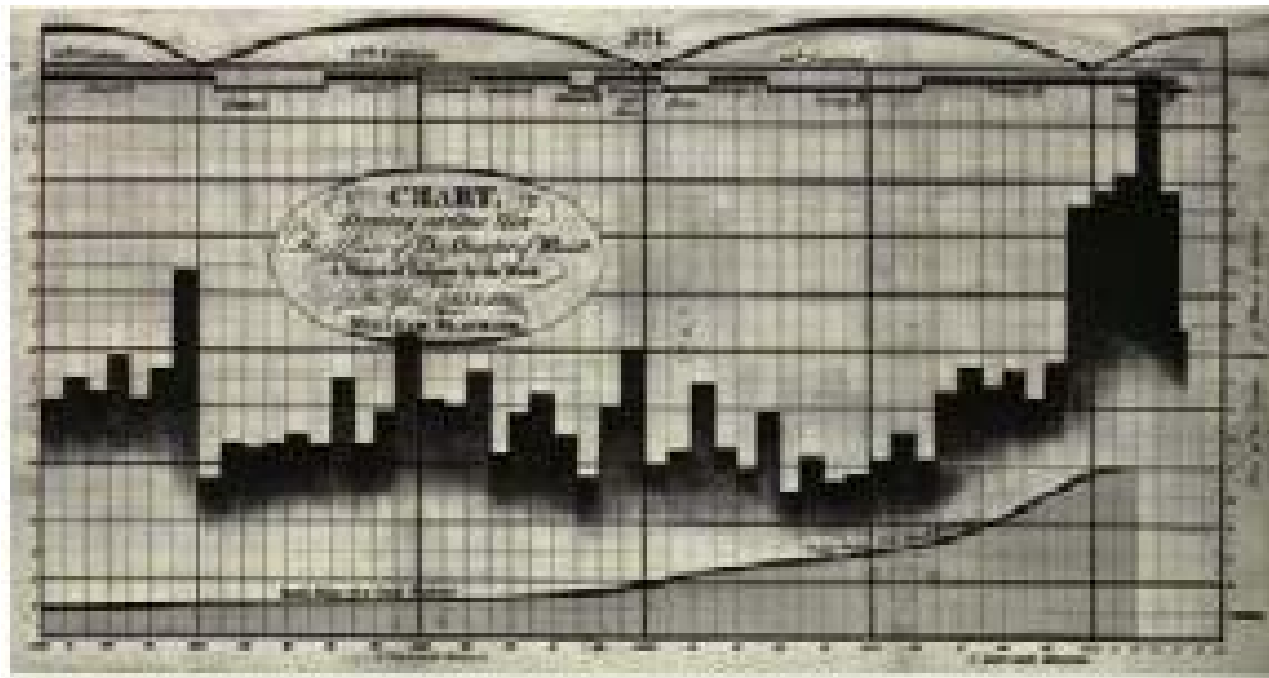
# Vizualizace – zdroje



- ⌘ Na prvním místě se obvykle citují knihy prof. Tufteho, např. Tufte E.R. (1983) The Visual Display of Quantitative Information, Graphic Press, Cheshire, Conn.
- ⌘ Weby o vizualizaci, např.
  - ☞ <http://www.math.yorku.ca/SCS/Gallery/noframes.html> - galerie s poučným výkladem a příklady i nezdařených či lživých grafů
  - ☞ <http://www.agocg.ac.uk/> - John Lansdown (1992) Aspects of Design in Computer Graphics: Some Notes –  
<http://www.agocg.ac.uk/train/hitch/hitch.htm>
- ⌘ Jiné weby, např. stránky různých vizualizačních programů a organizací
  - ☞ <http://www.cybergeography.org/atlas/atlas.html> nebo <http://miner3d.com/products/gallery.html>

# Vizualizace – historie

- William Playfair, 1786: první publikovaná prezentační grafika



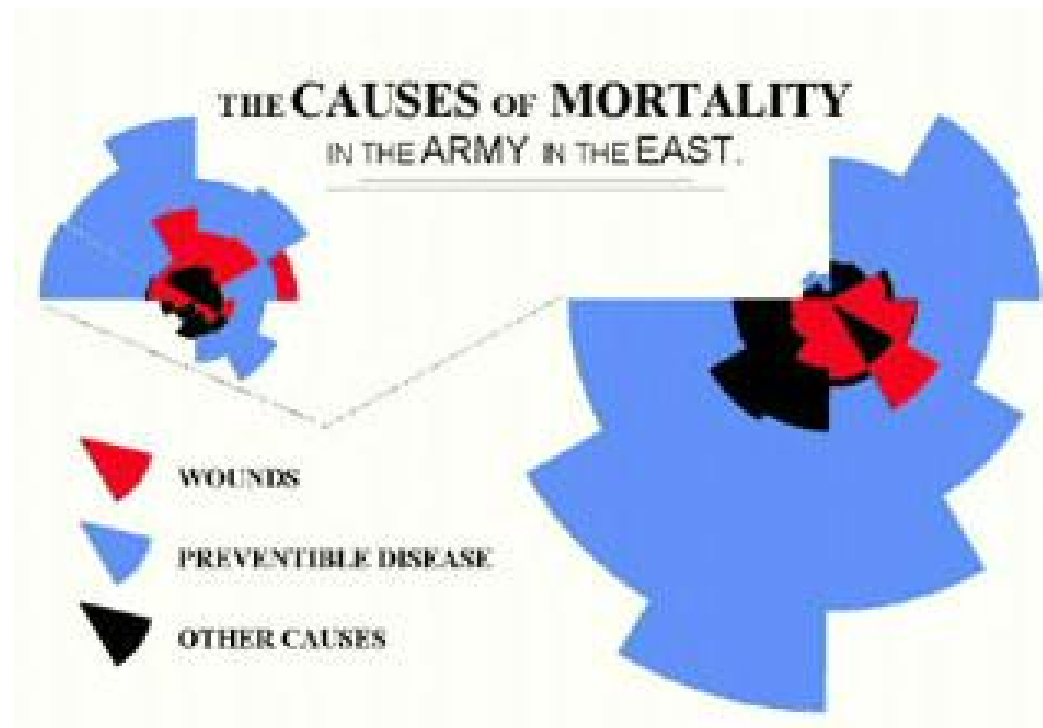
# Vizualizace – historie

- Dr. John Snow, 1845: epidemie cholery v Londýně



# Vizualizace – historie

- Florence Nightingale, 1858: důvody úmrtí v průběhu Krymské války (1853-1856)





# Vizualizace – historie

- Harry Beck, 1931: schéma Londýnského metra



# Vizualizace – investigativní analýza



<http://www.i2inc.com/>



## Law Enforcement

- » Counterterrorism
- » Narcotics investigations
- » Organized crime
- » Intelligence analysis
- » Fraud
- » Missing persons
- » Major investigations
- » Counterfeiting
- » Immigration control
- » Major event security
- » Money laundering
- » Gang investigations

## Government

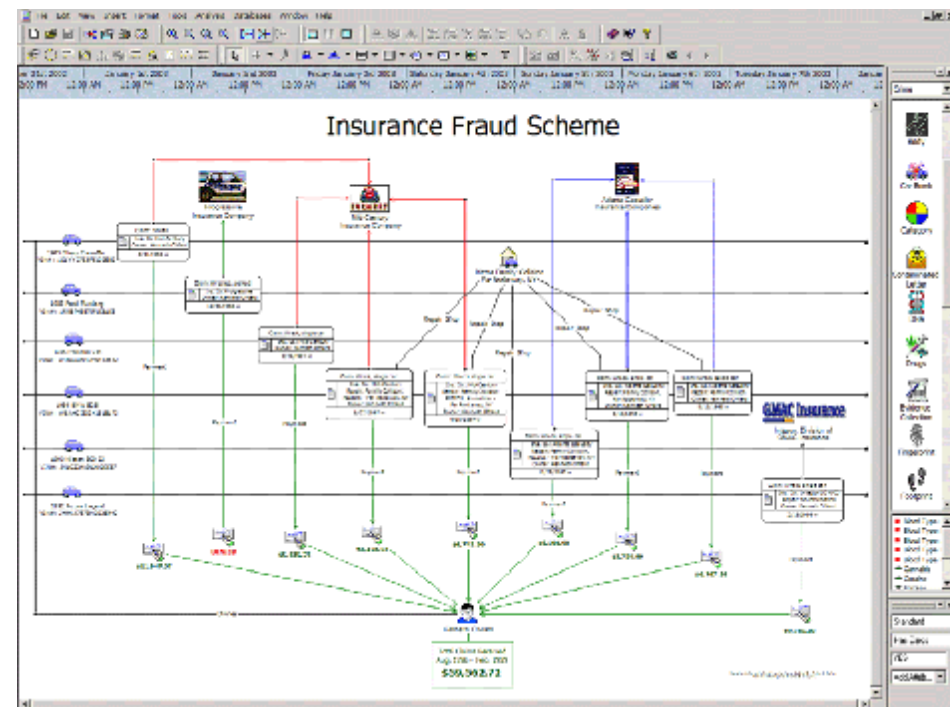
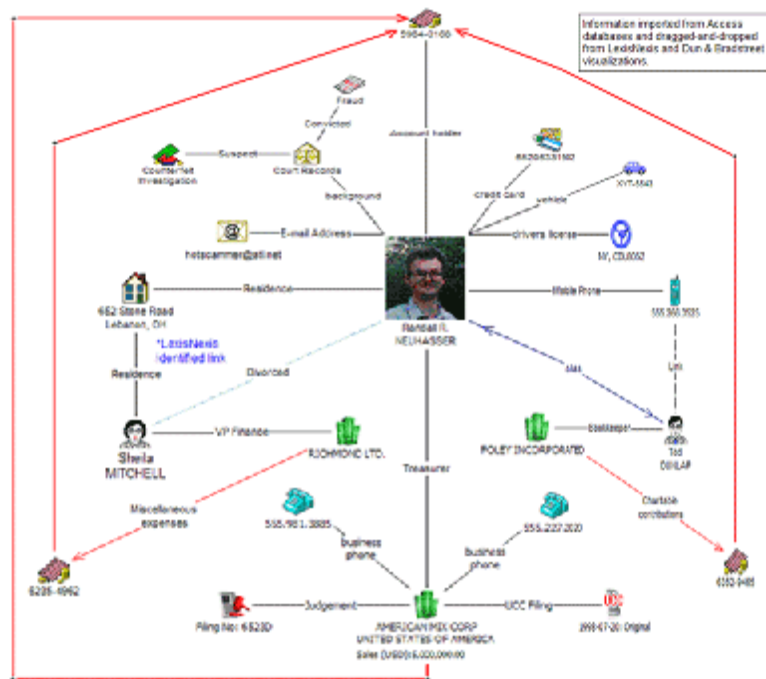
- » Criminal prosecutions
- » National security
- » Military intelligence
- » Embassy security
- » Postal inspection and fraud
- » Prison investigations
- » Park and wildlife services
- » Antitrust investigations
- » Tax fraud investigations
- » Customs investigations

## Commercial

- » Forensic accounting
- » Money laundering
- » Insider trading violations
- » Corporate security
- » Anti-pirating investigations
- » Entertainment copyright violations
- » Competitive intelligence
- » Civil lawsuits
- » Fraud:
  - » Credit card
  - » Insurance
  - » Retail
  - » Health care
  - » Commercial
  - » Telephone

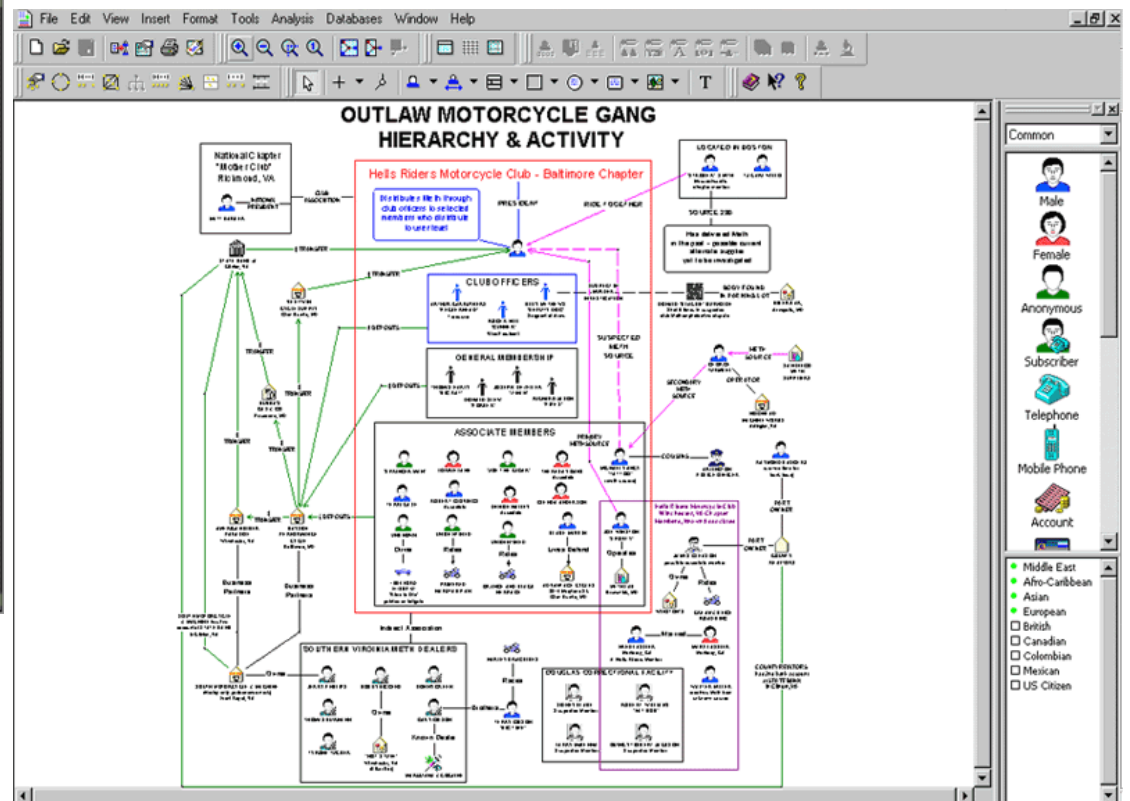
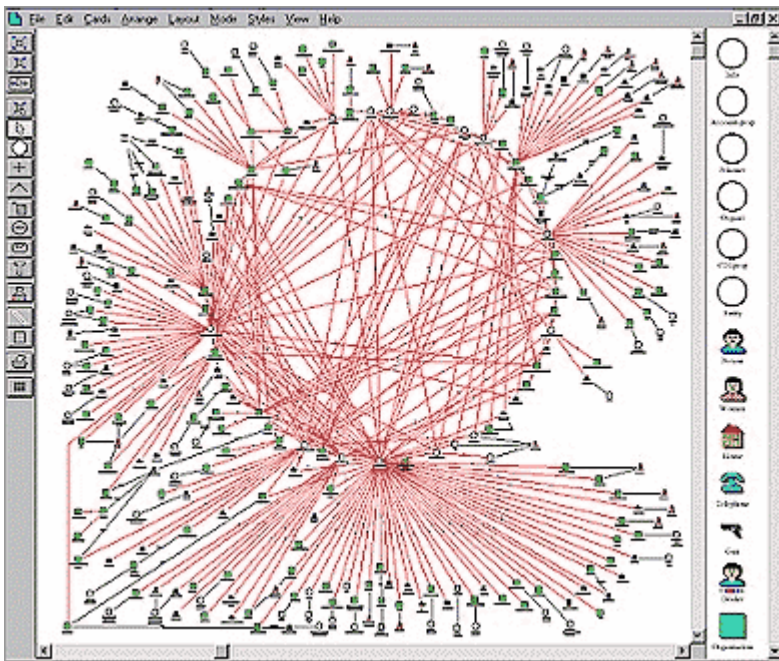
# Vizualizace – investigativní analýza

- osobní kontakty, pojistné podvody



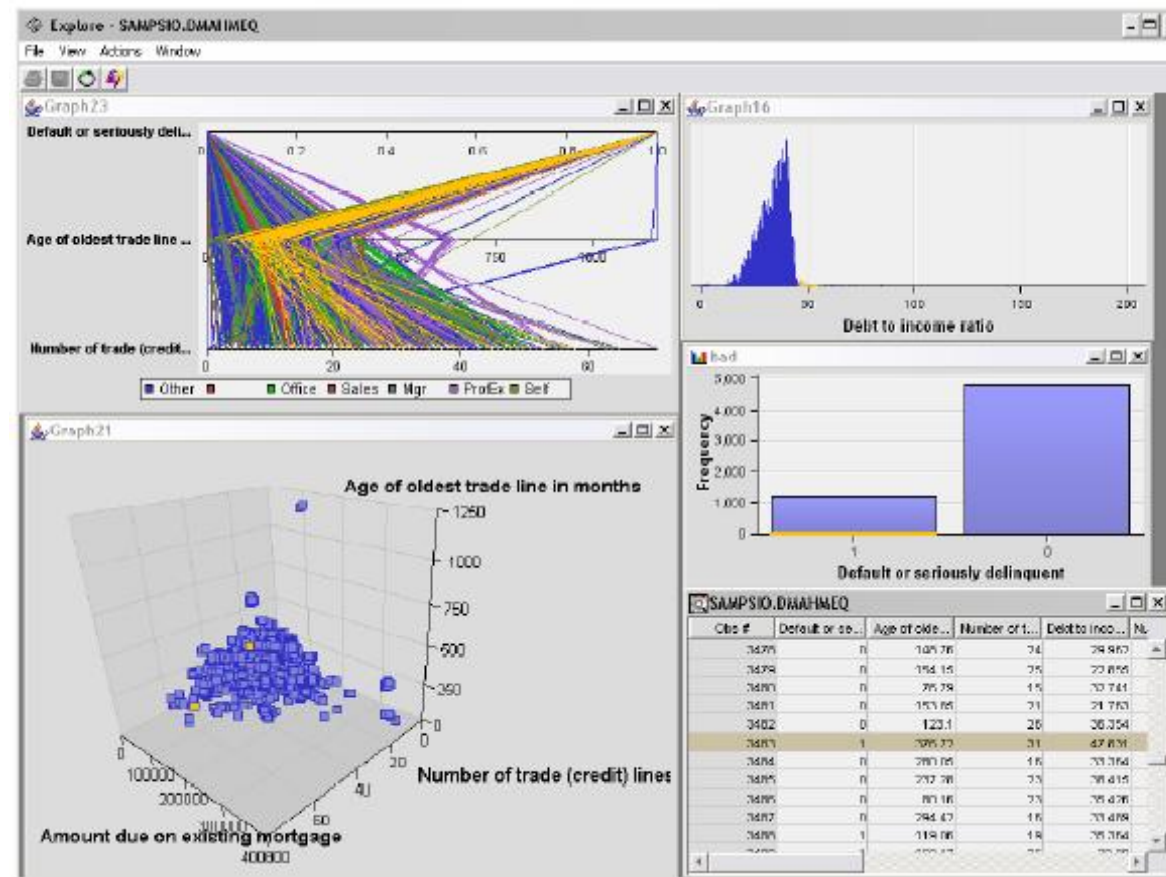
# Vizualizace – investigativní analýza

- Praní špinavých peněz, kriminální gangy



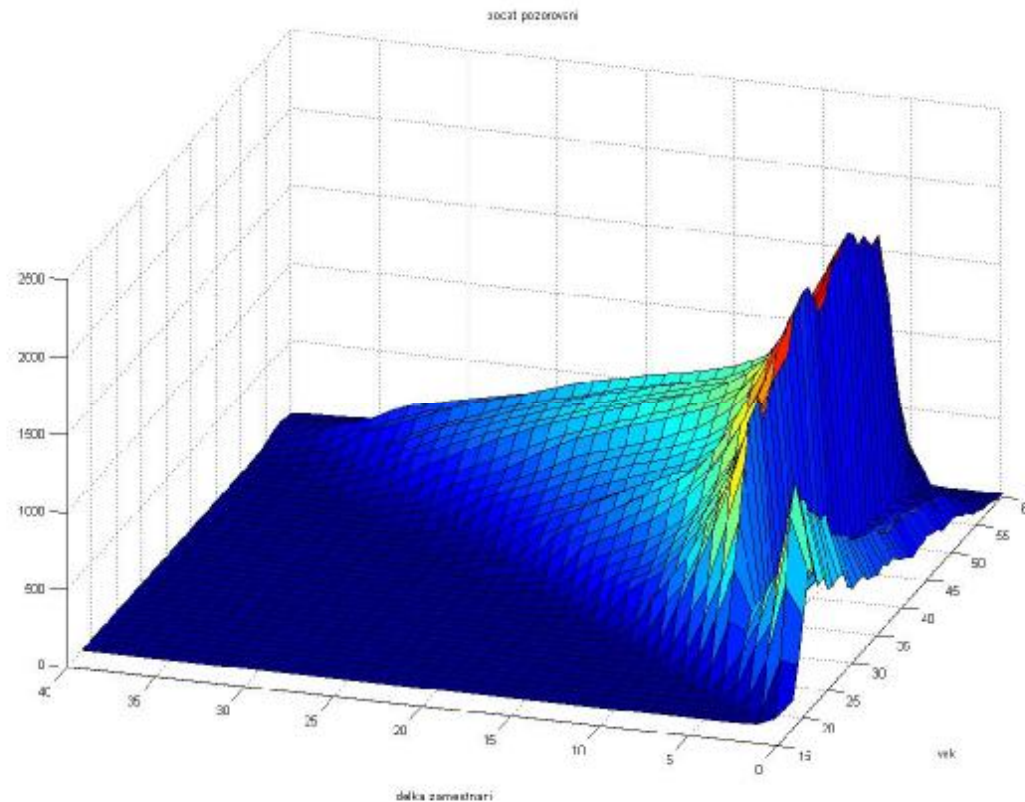
# Vizualizace – portfolio management

- Hledání závislostí v datech:



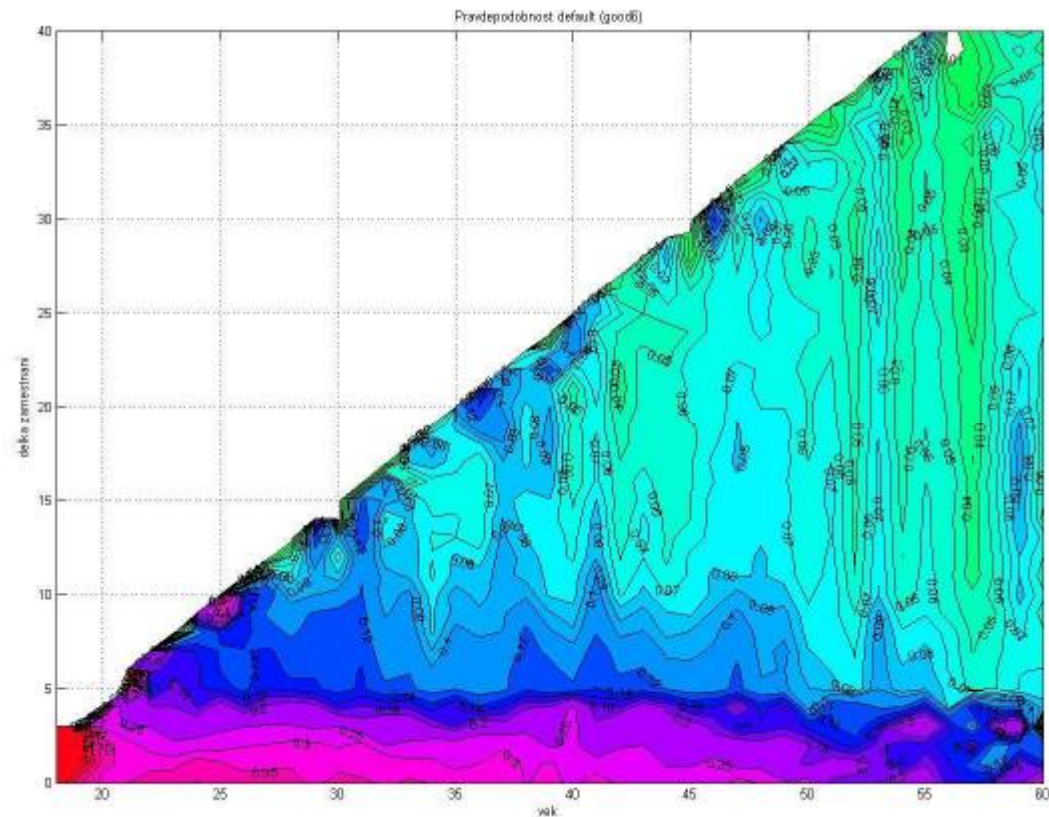
# Vizualizace – portfolio management

- Věk vs. délka zaměstnání



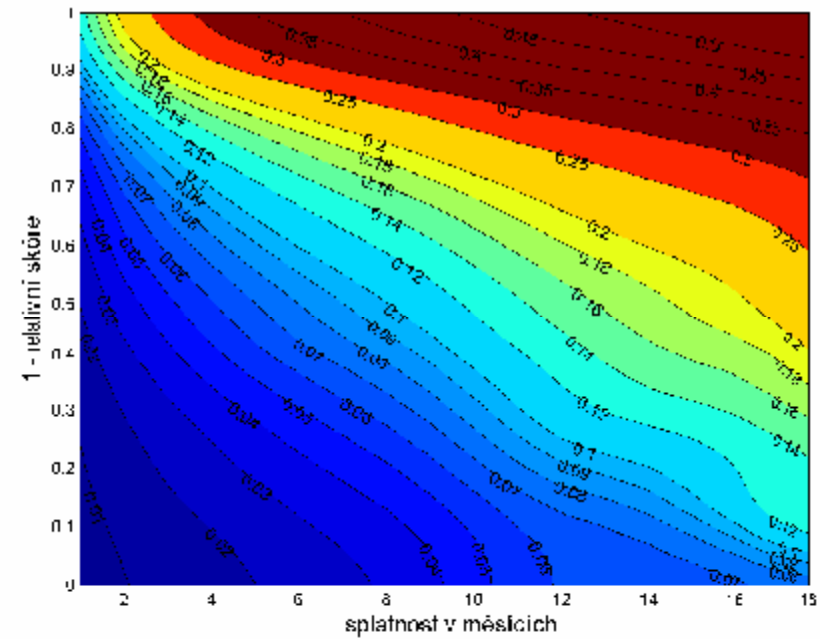
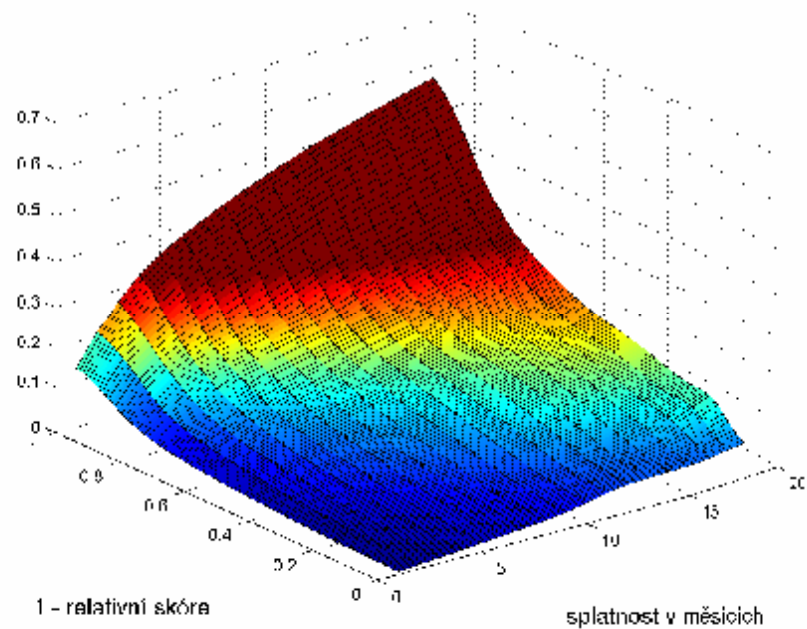
# Vizualizace – portfolio management

- Věk vs. délka zaměstnání vs. default



# Vizualizace – portfolio management

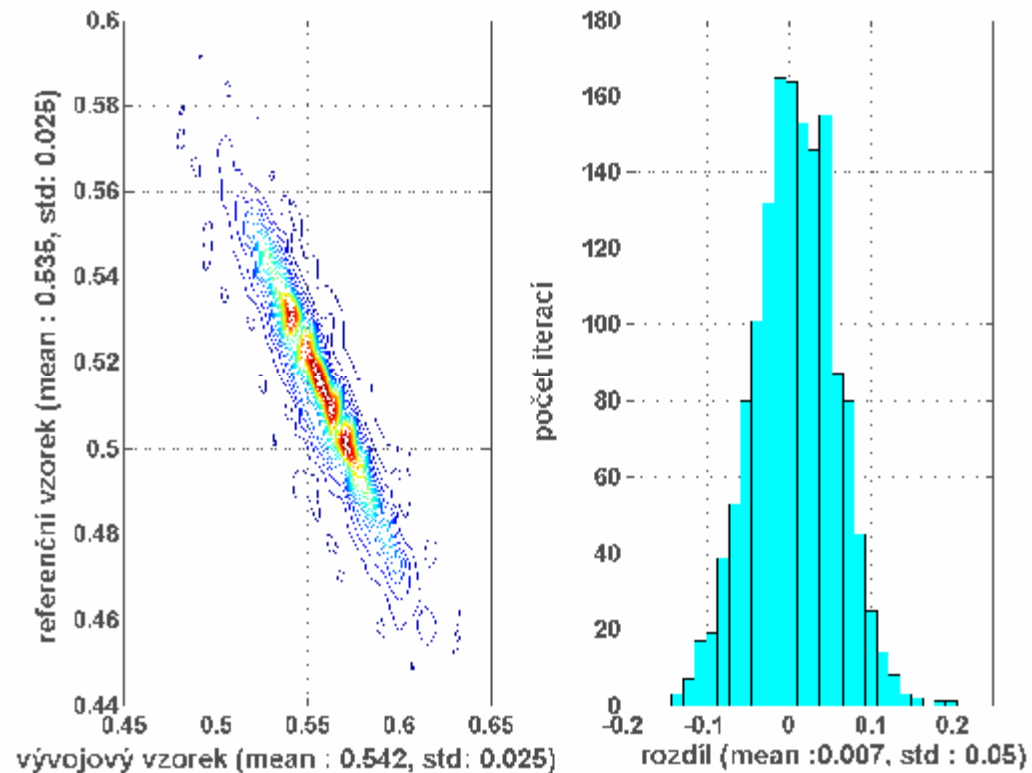
- Distribuční funkce doby do defaultu





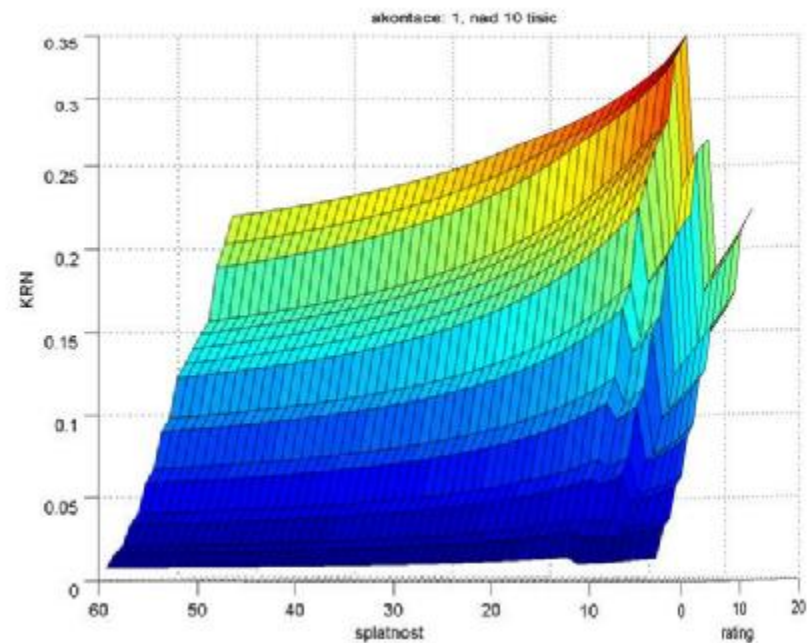
# Vizualizace – portfolio management

- Test stability scóringové funkce



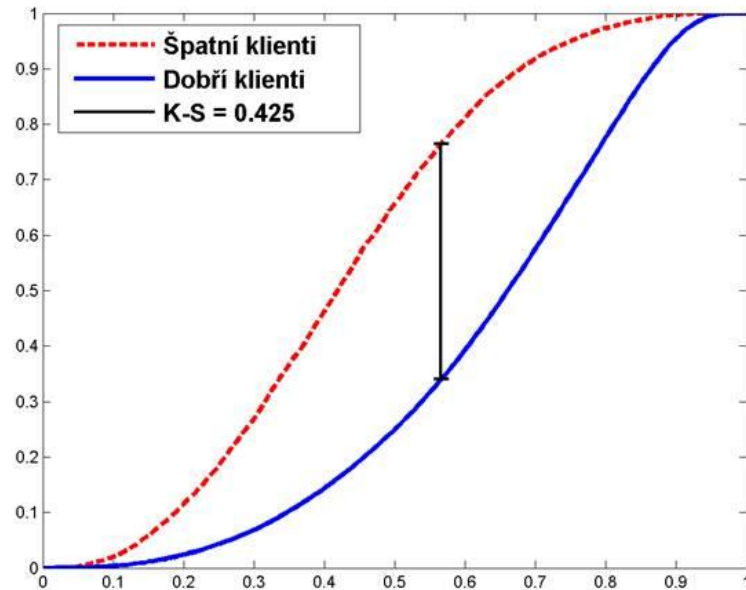
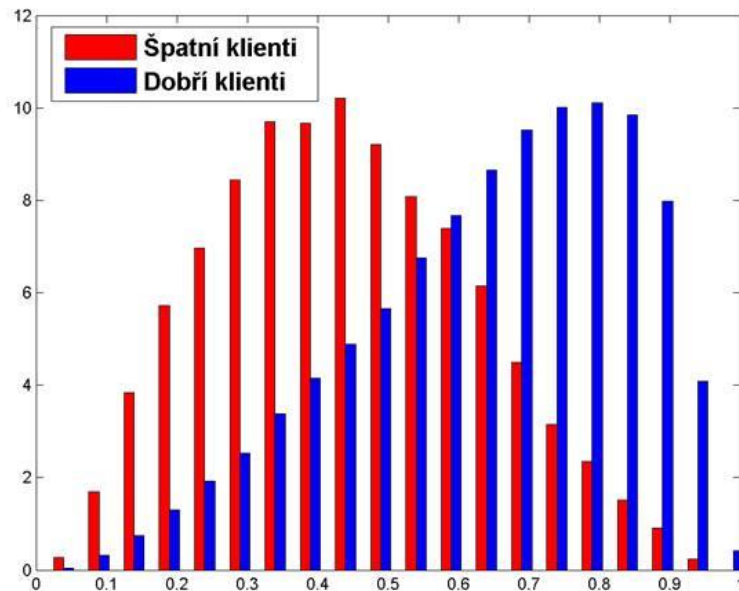
# Vizualizace – portfolio management

- Vizualizace kreditních rizikových nákladů (KRN)



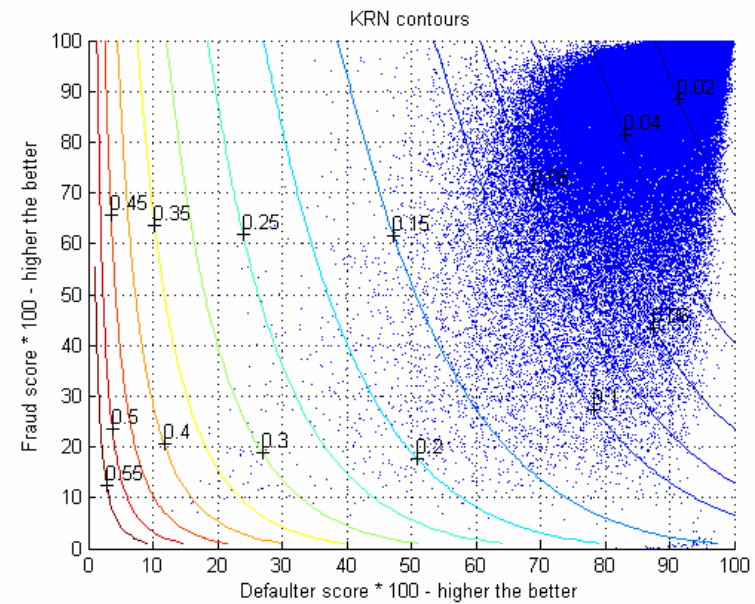
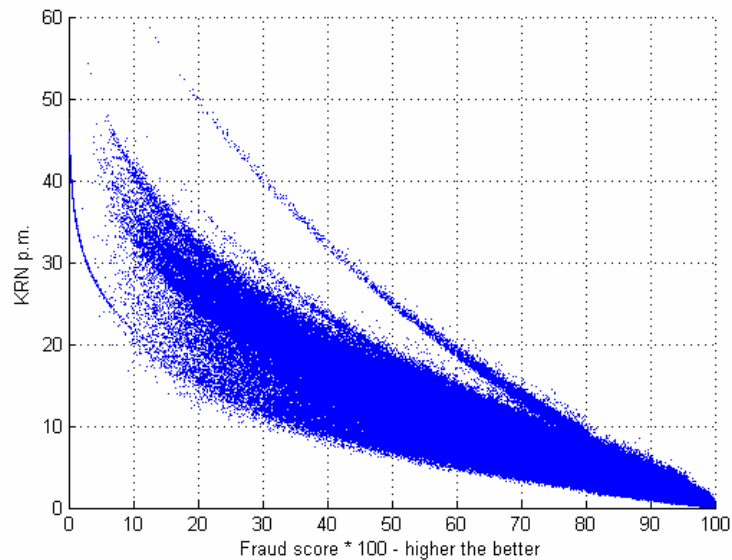
# Vizualizace – portfolio management

□ Histogram, Distribuční funkce



# Vizualizace – portfolio management

## □ Bodové grafy

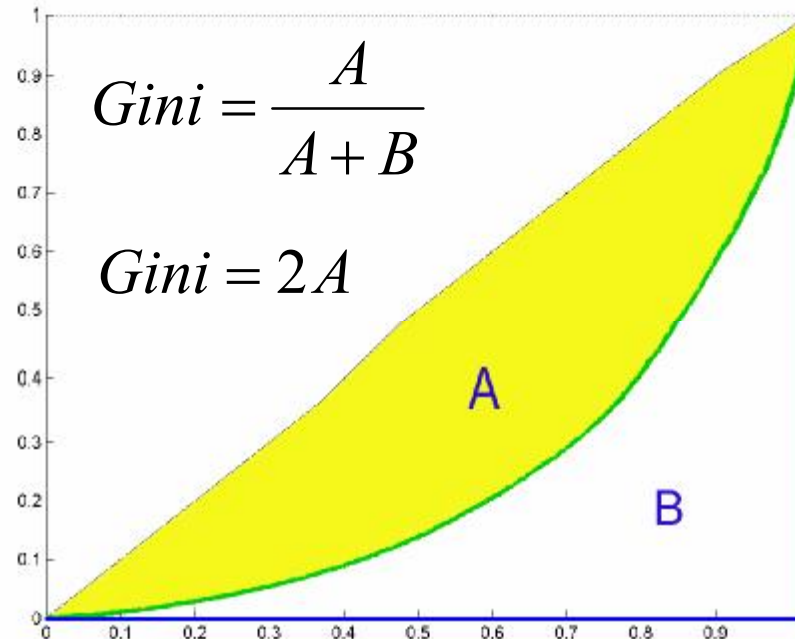


# Vizualizace – portfolio management

□ Lorenzova křivka

$$c\_stat = A + \frac{1}{2}$$

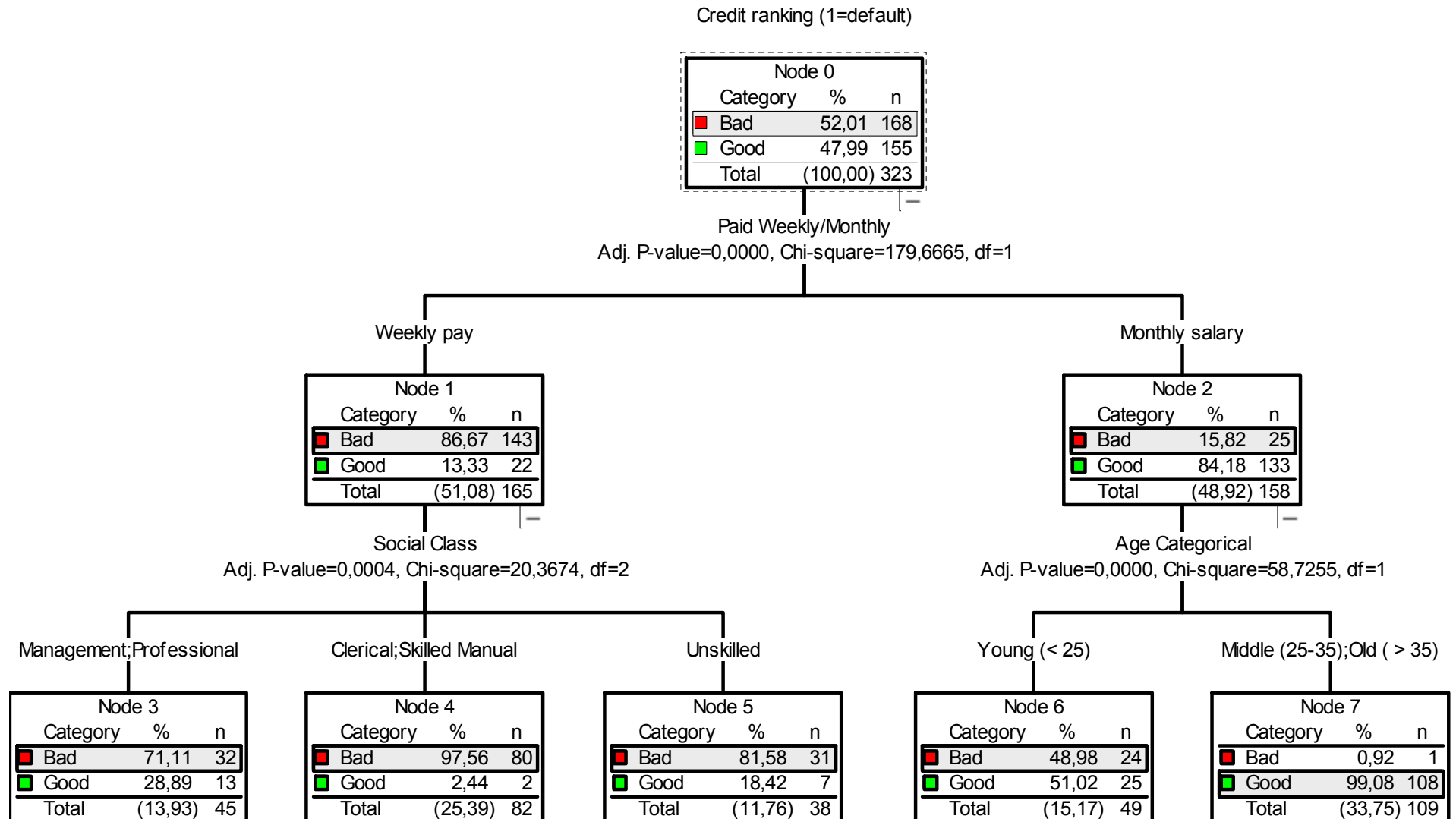
$$c\_stat = \frac{1}{2}(1 + Gini)$$



**C-statistika = pravděpodobnost, že náhodně vybraný dobrý klient má lepší skóre než náhodně vybraný špatný klient.**

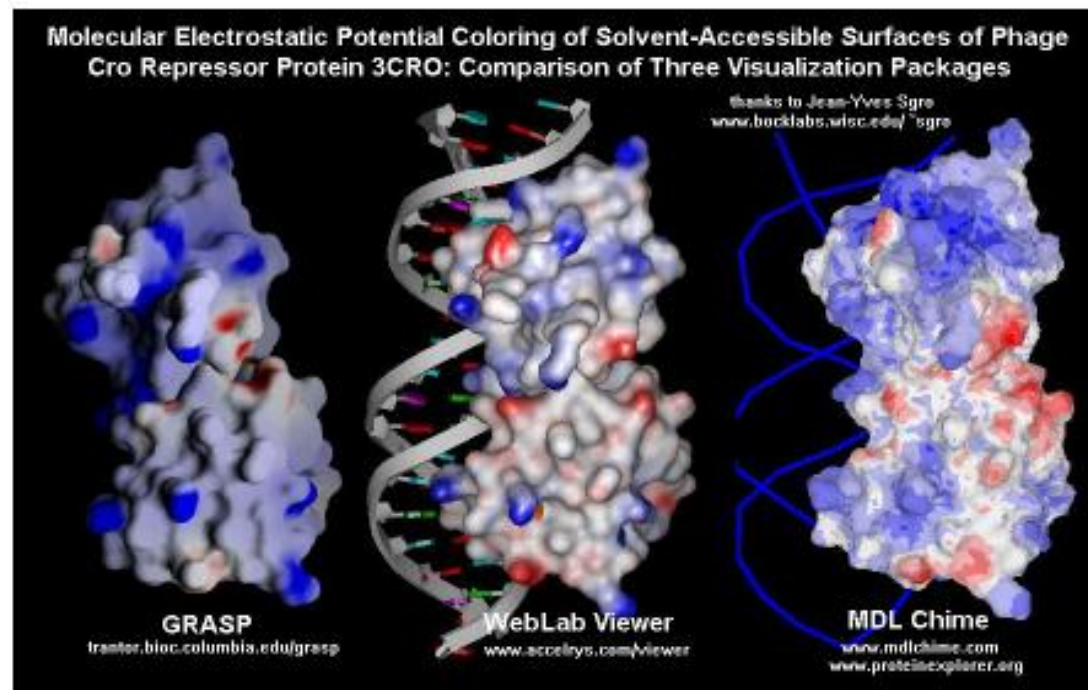
**Nebo... = pravděpodobnost, že aktuálně nabývá sledovaný ukazatel vyšších hodnot než nabýval na referenčním období.**

# Vizualizace - dendrogram

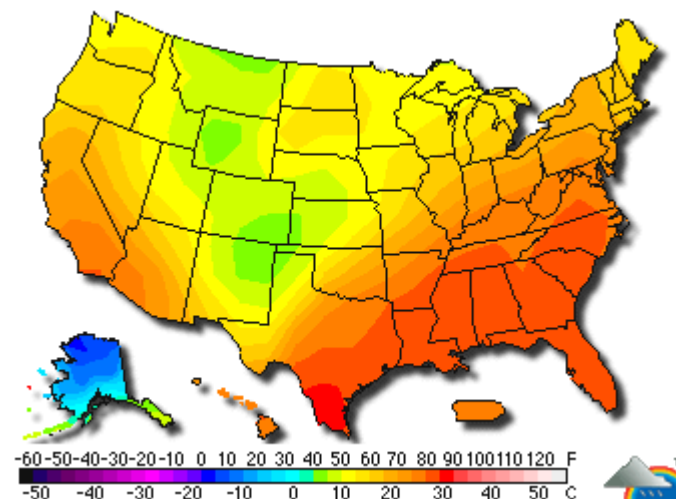
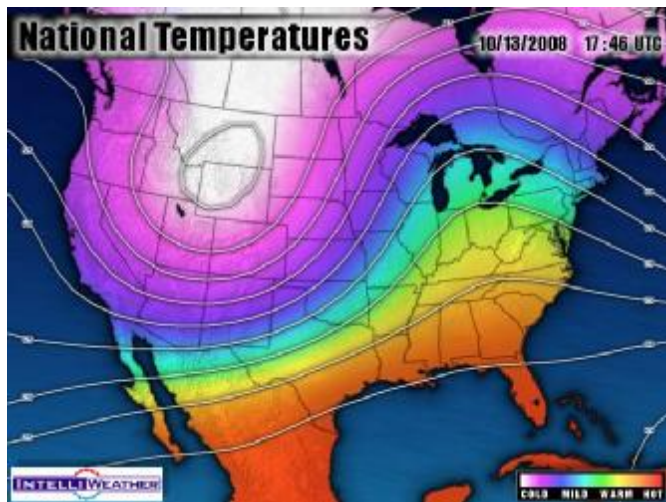
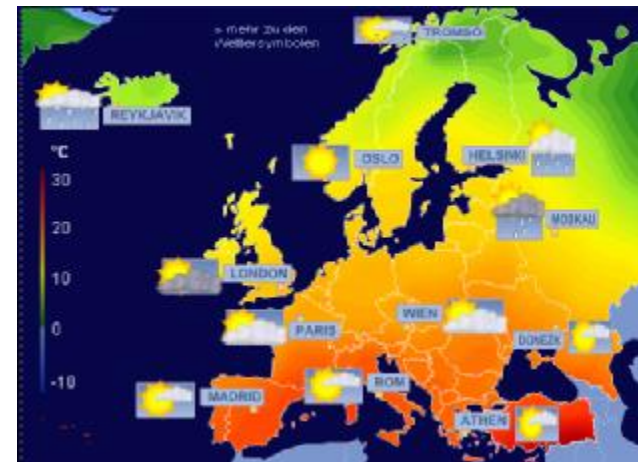


# Vizualizace v biologii/chemii

- 3D zobrazení proteinu



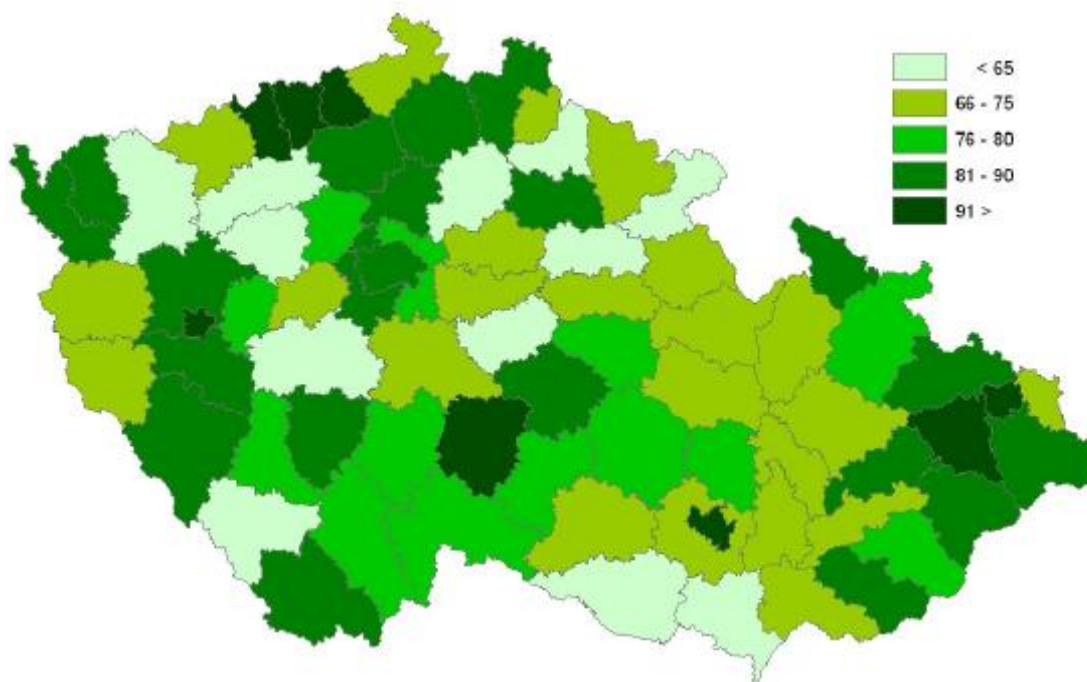
# Meteo-vizualizace





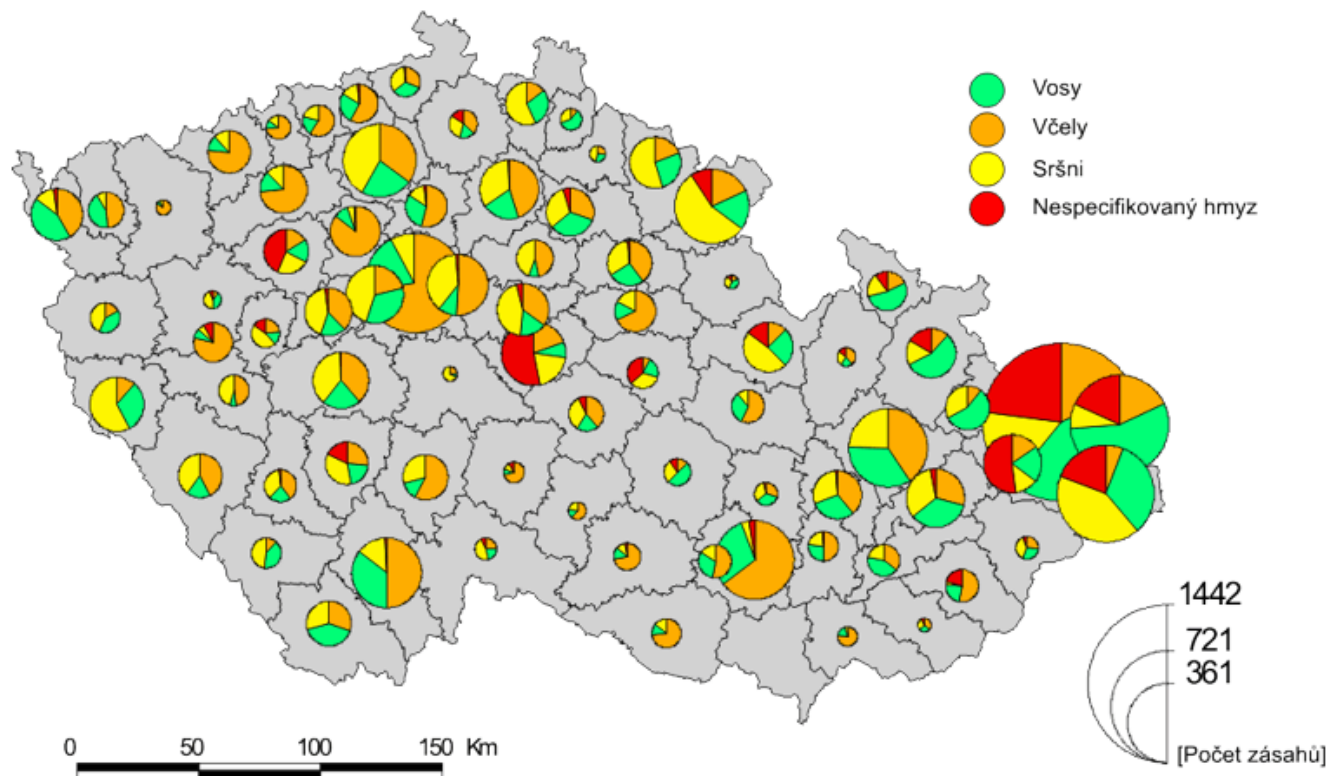
# Kartogram

- Obce s počtem 500 a více obyvatel s vysokorychlostním připojením k internetu, podle okresů (%), k 31.12.2006

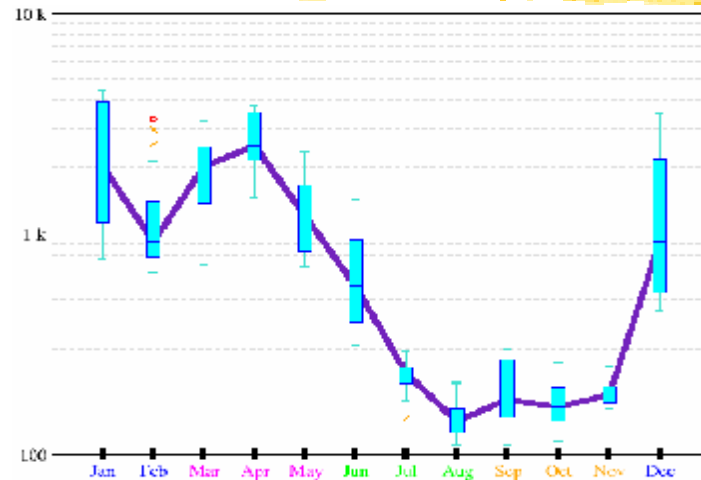


# Kartodiagram

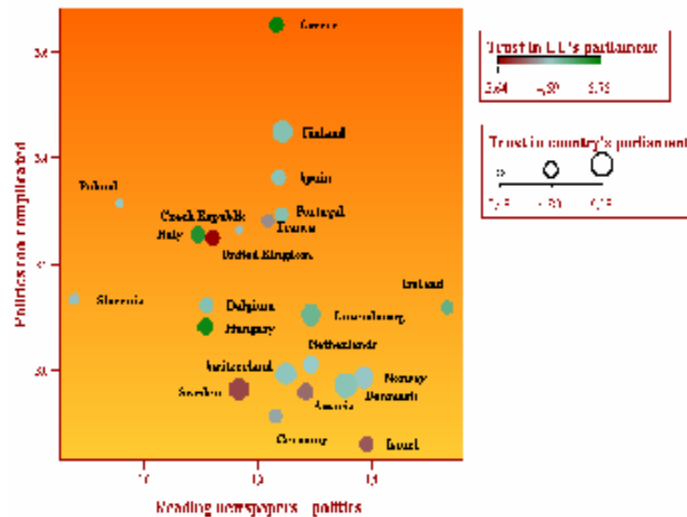
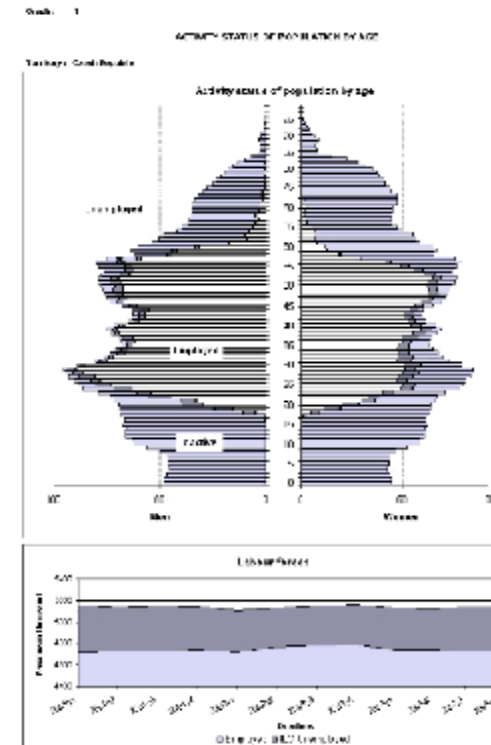
## ZÁSAHY JEDNOTEK PO PROTI HMYZU v okresech České republiky v letech 1997-2000



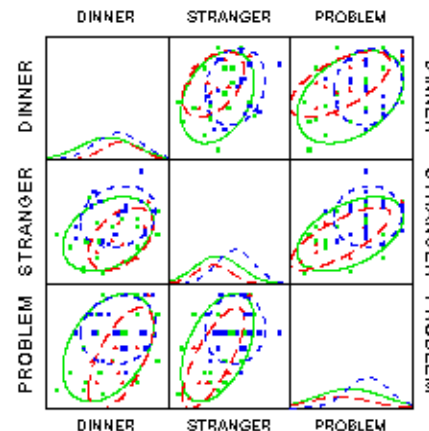
# Grafy –další typy



Zdroj: plasma-gate.weizmann.ac.il/ Grace/gallery



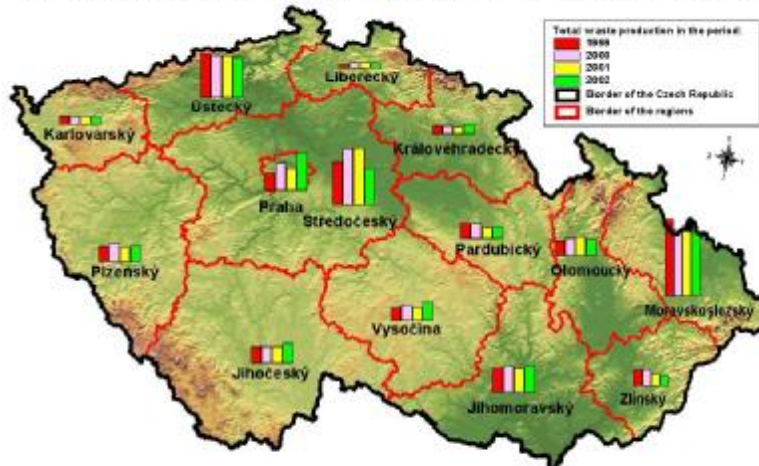
Social Competence Measures Across Setting



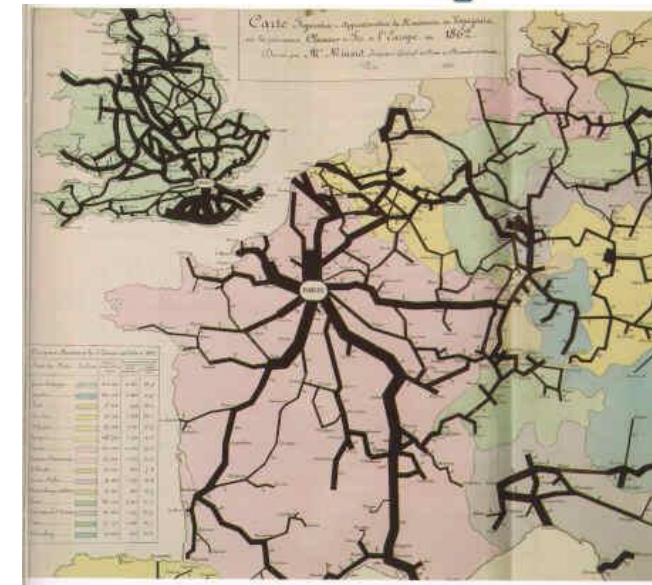
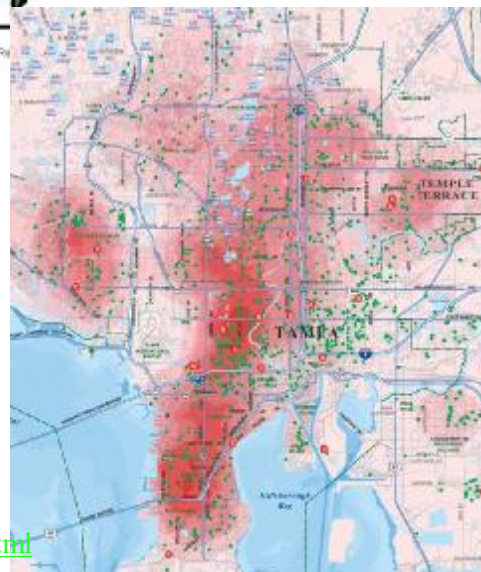
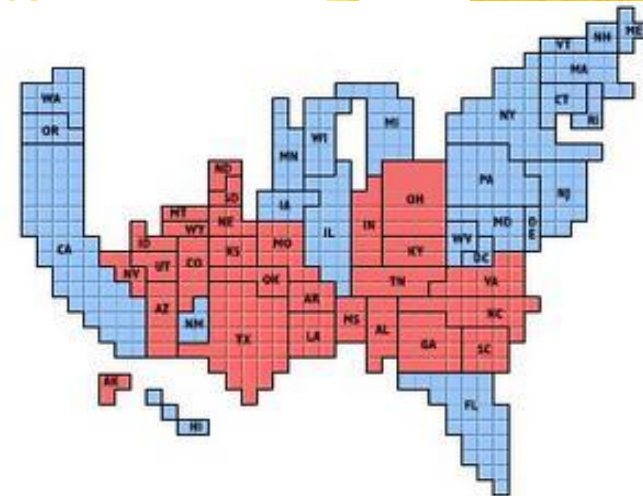
SETTING  
 ■ Center  
 ▲ Sitter  
 ● Parent

# Geografická data

Overall waste production in the regions of the Czech Republic

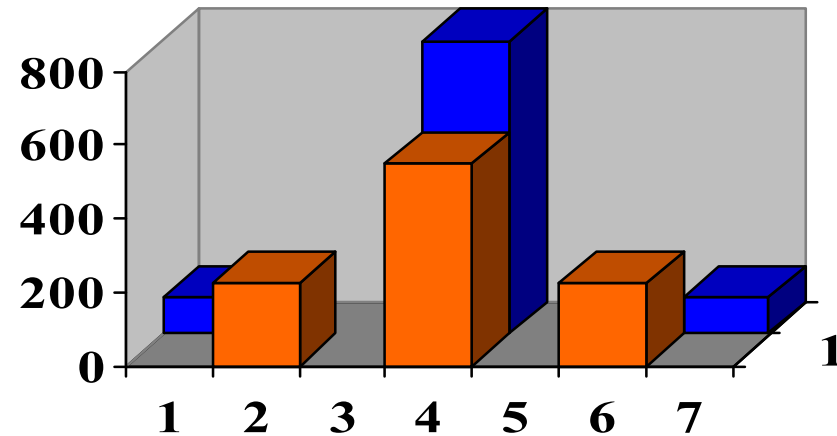


Source: Regionální rozvoj a agentura Ústeckého kraje, a.s. (Regional Development Agency of the Ústí Region)



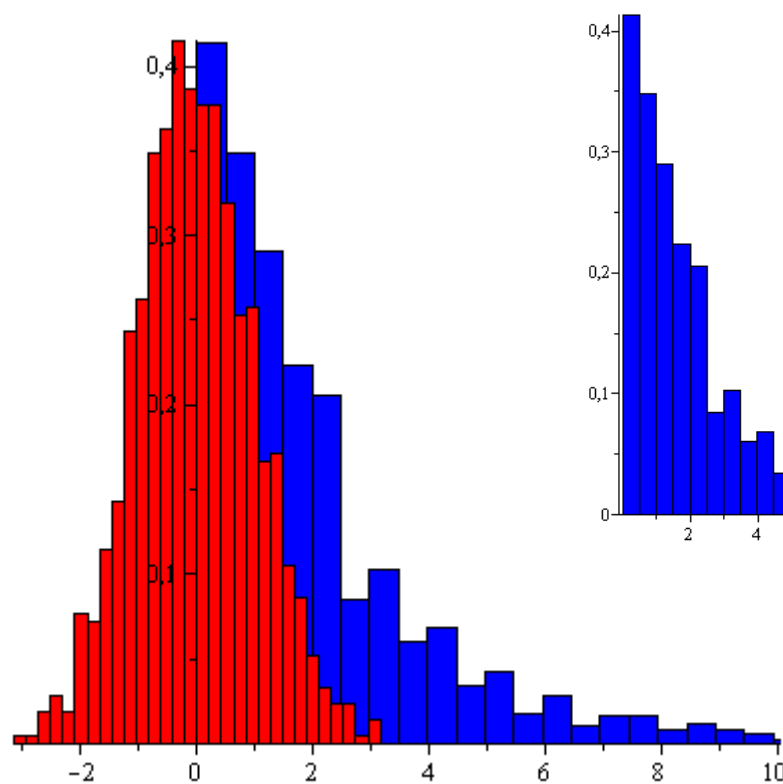
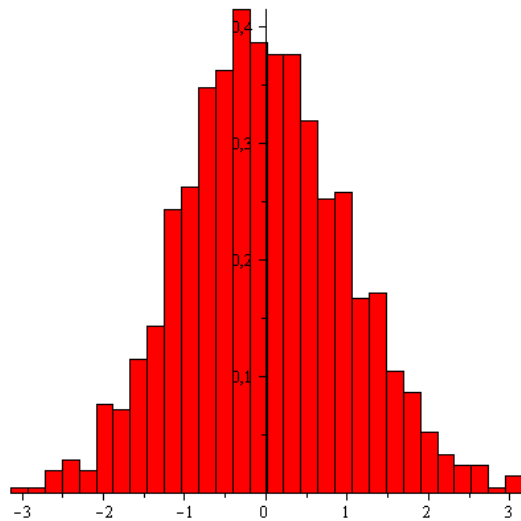
# Nehodnotit jen na základě popisných charakteristik

- shodný průměr (4)
- shodný medián (4)
- shodný modus (4)
- shodný rozptyl (1,8)
- shodná šikmost (0)
- shodný rozsah (1000)
- různá špičatost



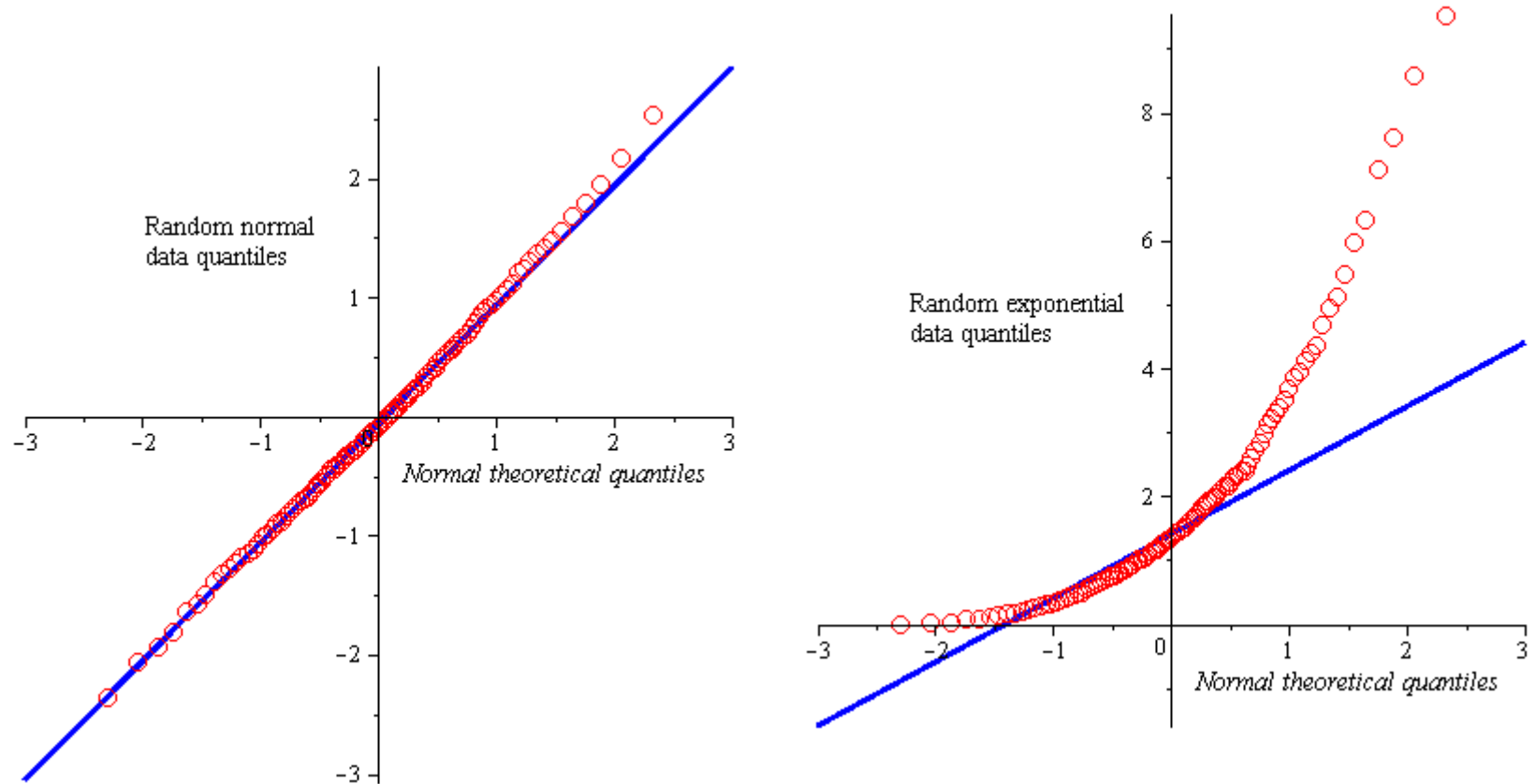
# Q-Q plot

- Histogramy normálního(červeně) a exponenciálního(modře) rozdělení



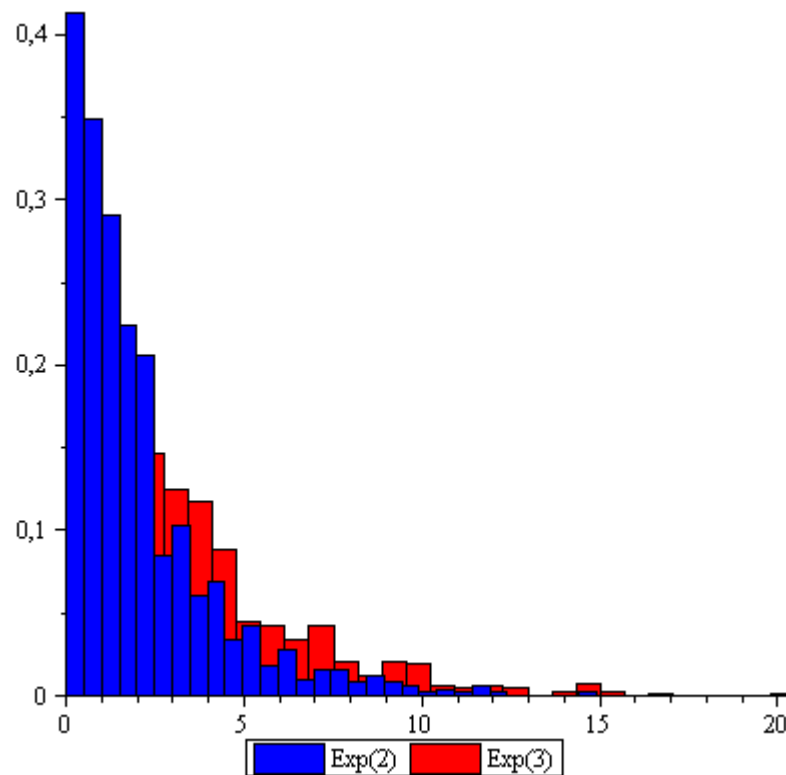
# Q-Q plot

- Q-Q plot pro normálně a exponenciálně rozdělených dat



# Q-Q plot

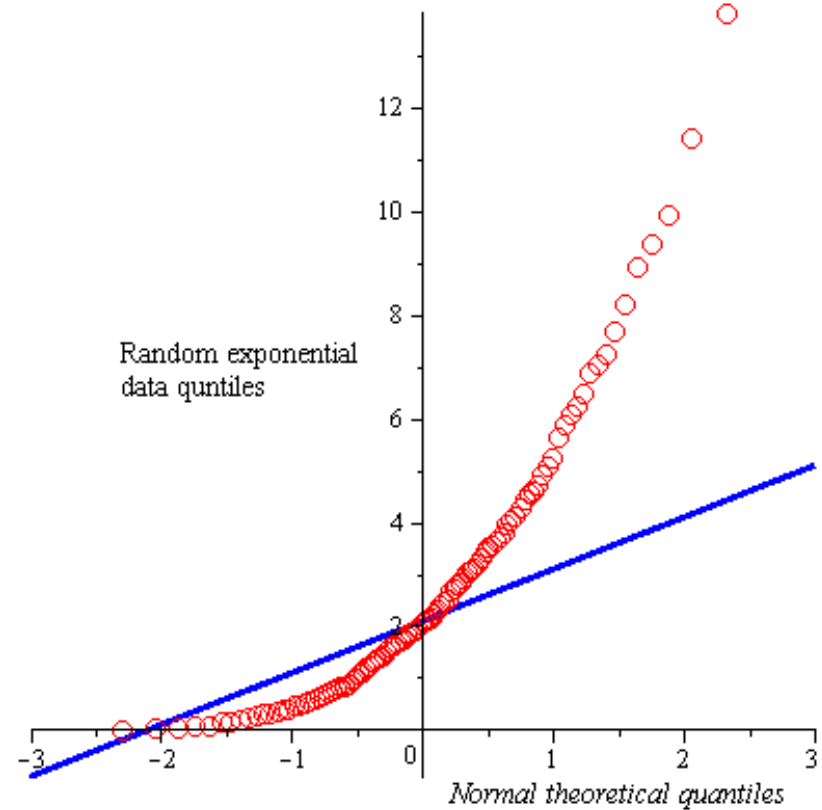
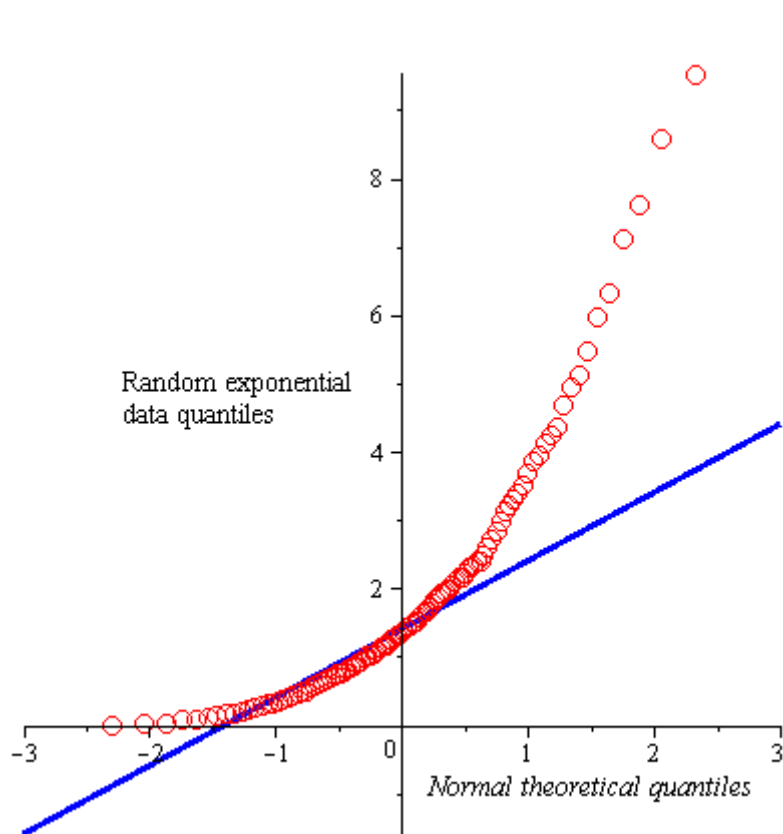
- Histogramy dvou exponenciálně rozdělených dat





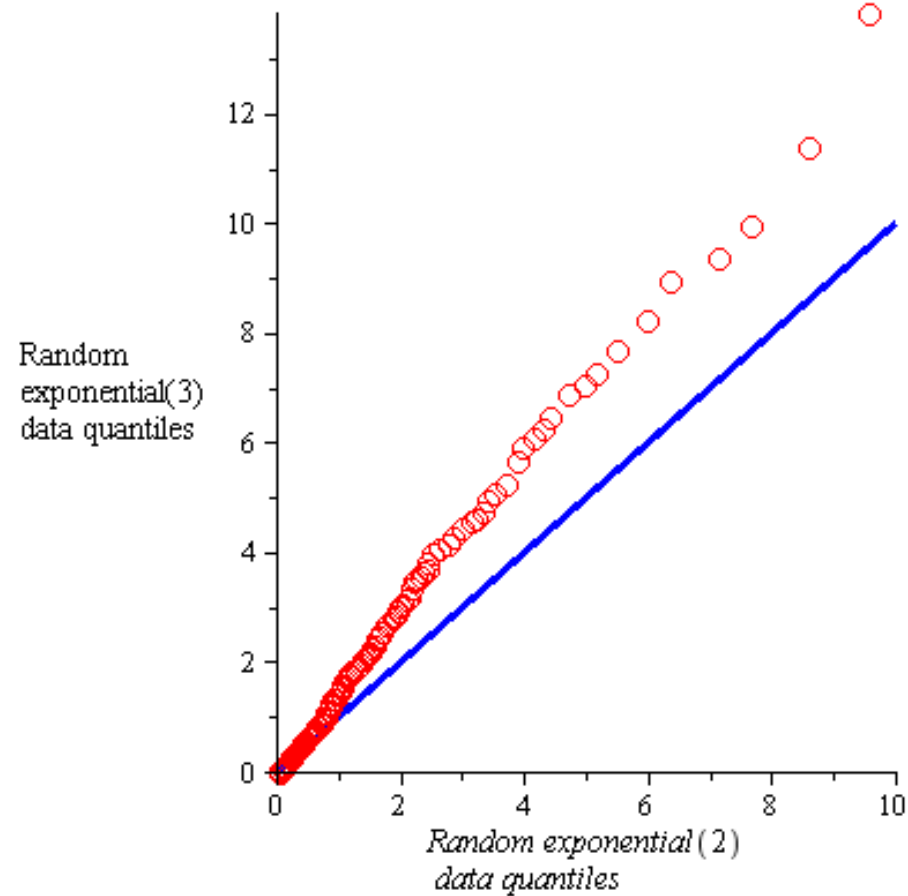
# Q-Q plot

- Q-Q plot: normální vs. exponenciální data



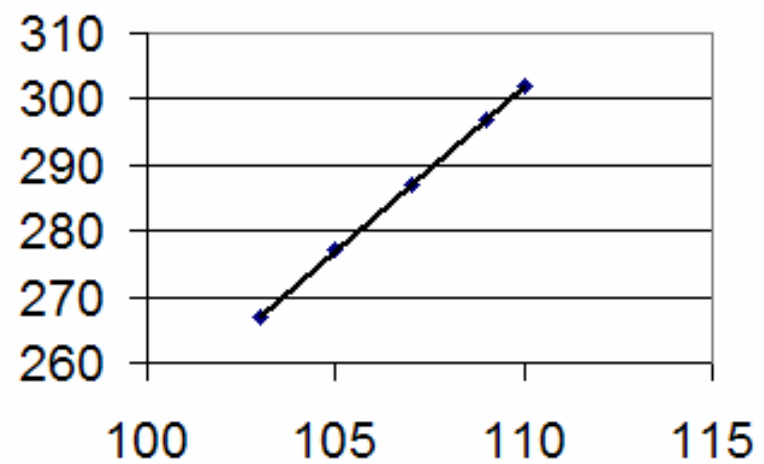
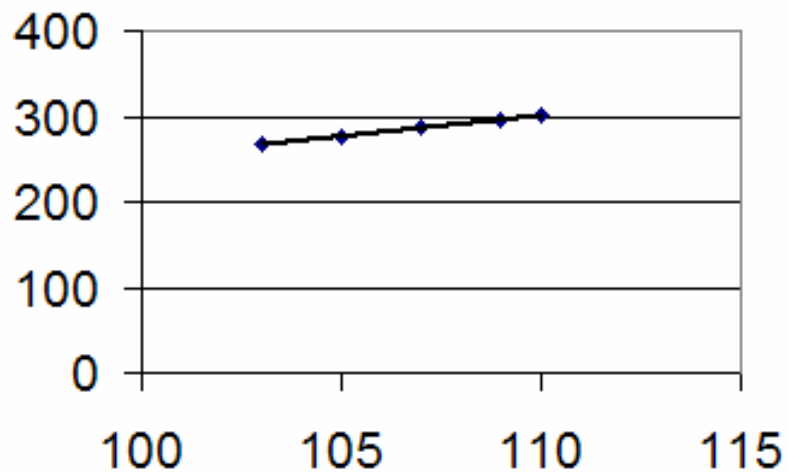
# Q-Q plot

- Exponenciální vs. exponenciální data



# Měřítko

□ která přímka roste strměji?



x	y
103	567
105	577
107	587
109	597
110	602

# Měřítko



- Pohled tvůrce grafu:
  - Zvýraznění trendu – pozitivní výsledky.
  - Potlačení trendu – negativní výsledky.
  
- Pohled uživatele grafu:
  - Grafy bez uvedeného měřítka jsou silně podezřelé.
  - Nepodléhat podsouvané informaci o růstu/poklesu.