

Vizualizace dat

Vizualizační formy prakticky

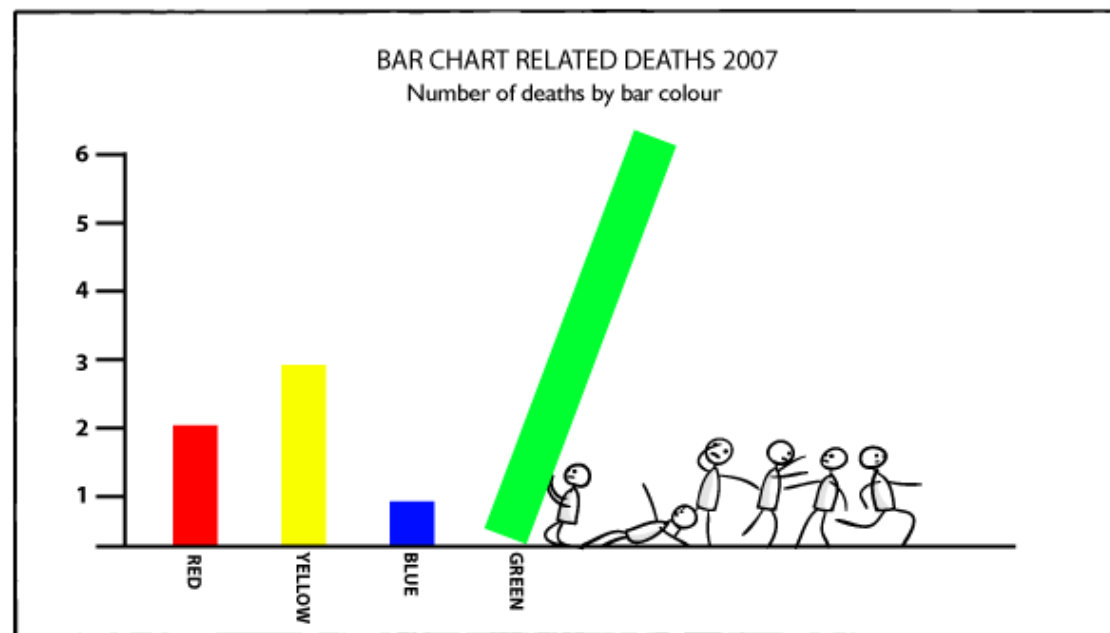
21. 11. 2019

Vizualizační metafory

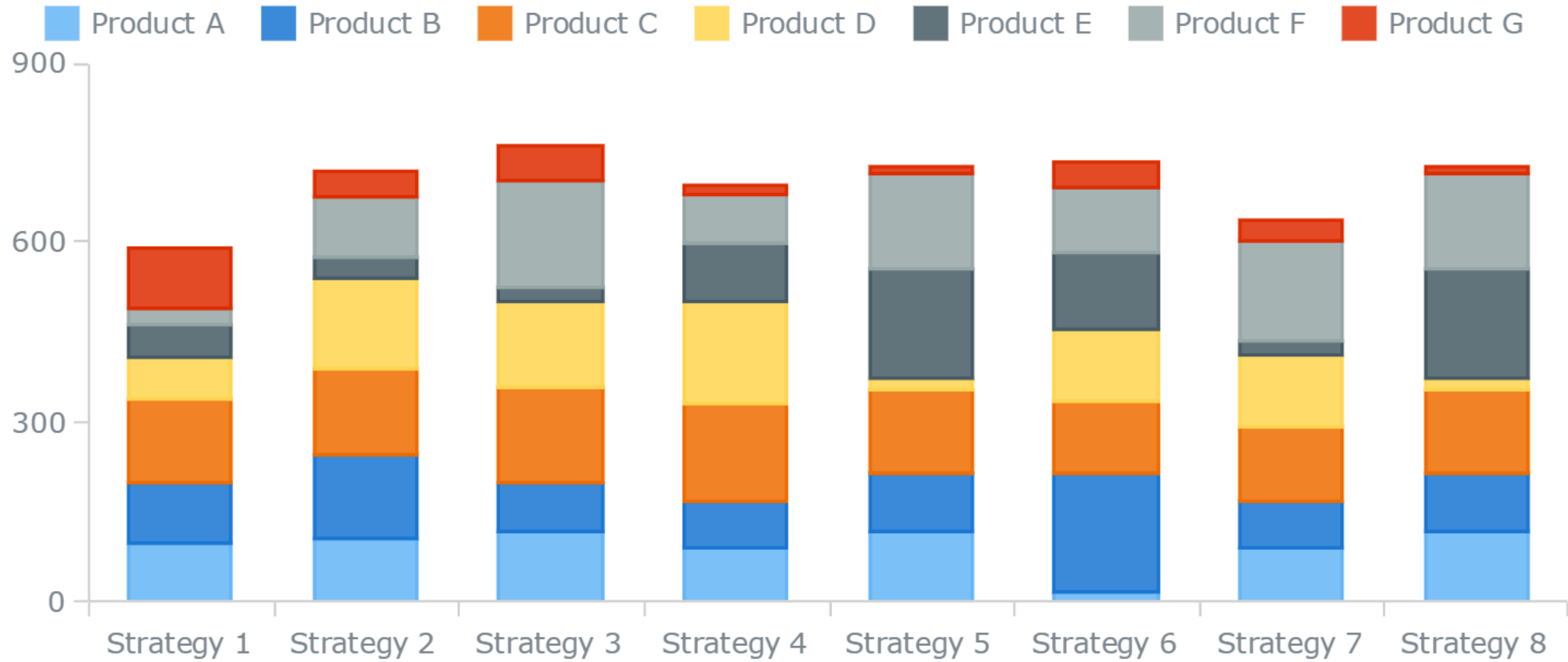
- je jich mnoho
- některé se osvědčily
- některé se ještě hledají
- některé vyvolávají kontroverze
- **dnes**: výběr klasiky, nových přístupů i zajímavostí

Sloupcové grafy

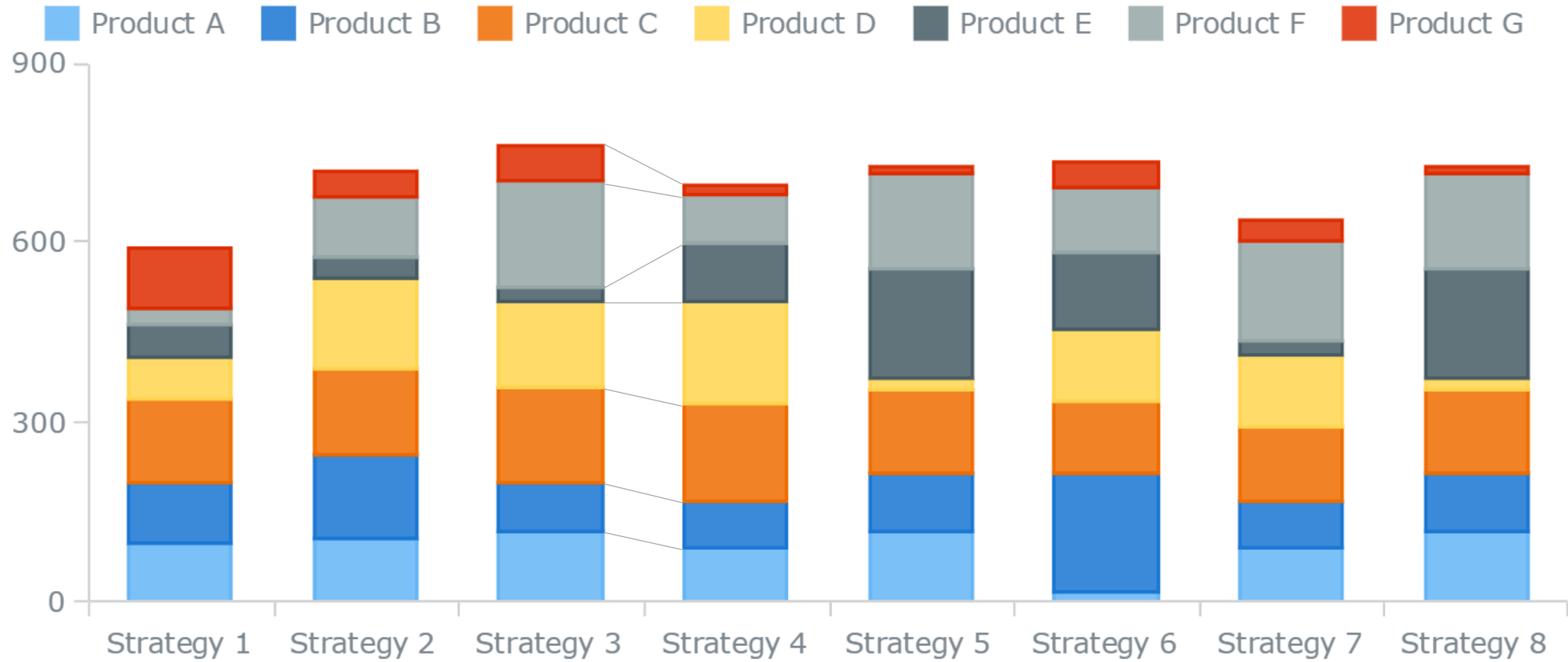
- rozložení souboru dat skrze sloupce
- délka odpovídá hodnotám
- sloupcový klasický
pruhový graf
- skládaný sloupcový
100% skládaný sloupcový



Compare sales strategy

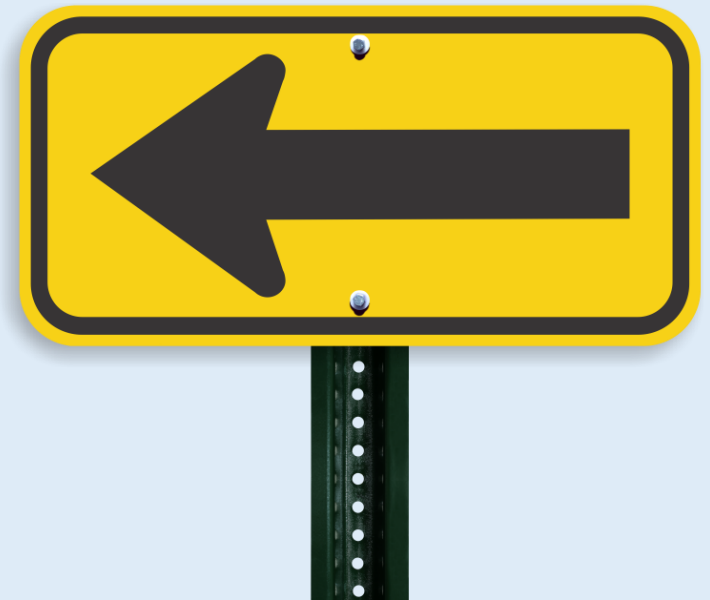


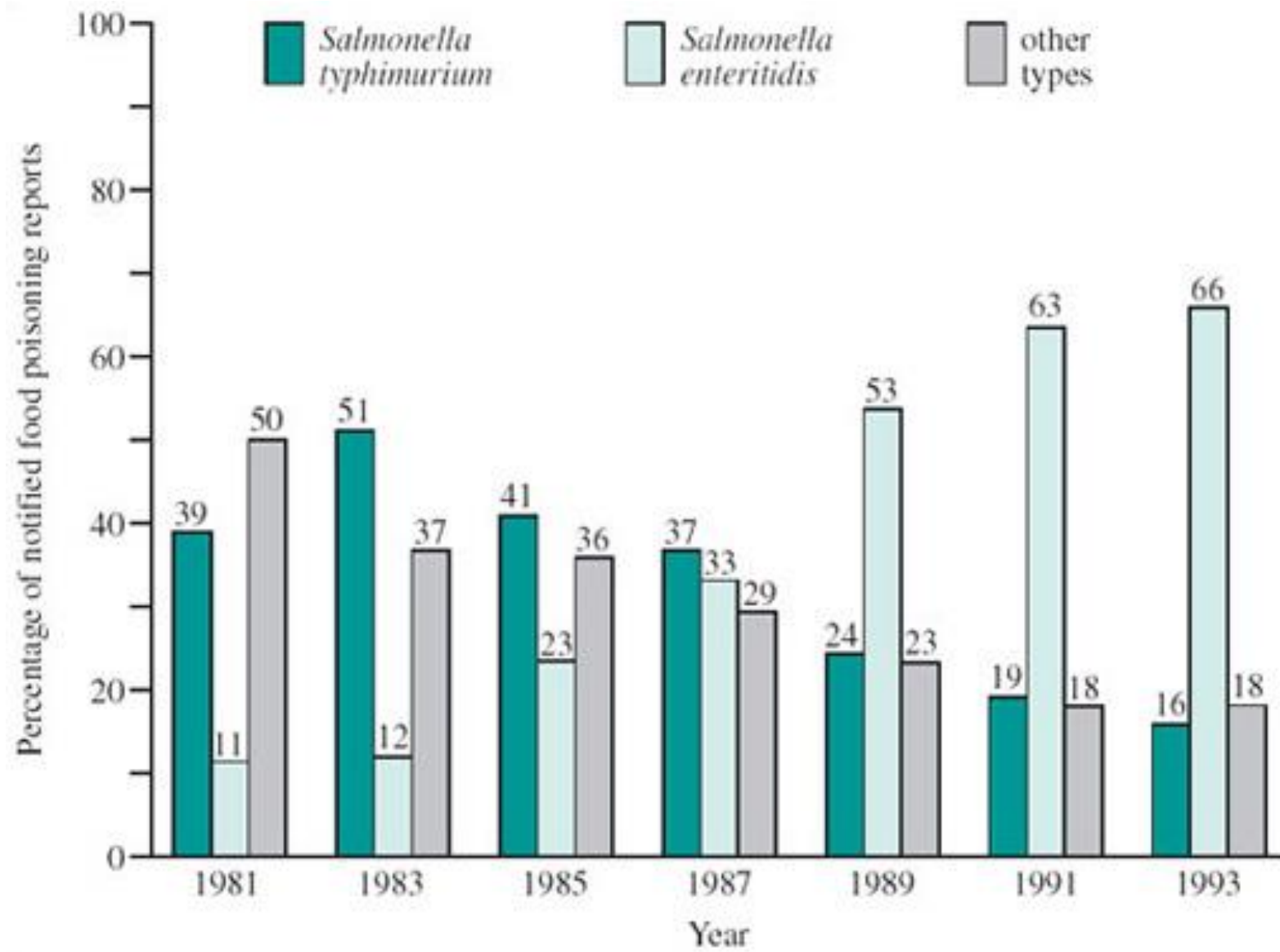
Compare sales strategy



Pár tipů ke sloupcům

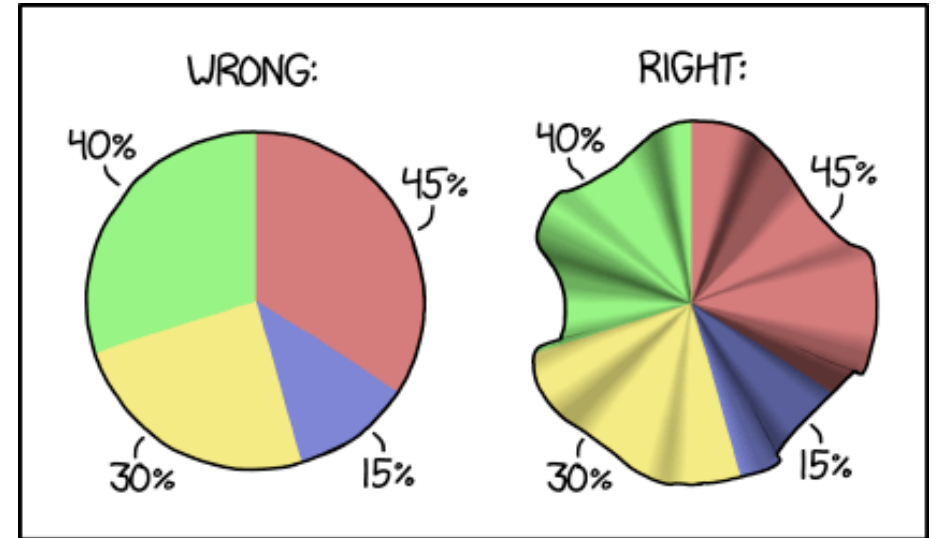
- mezery mezi sloupci
- příliš dlouhá legenda
- ...





Koláčové grafy

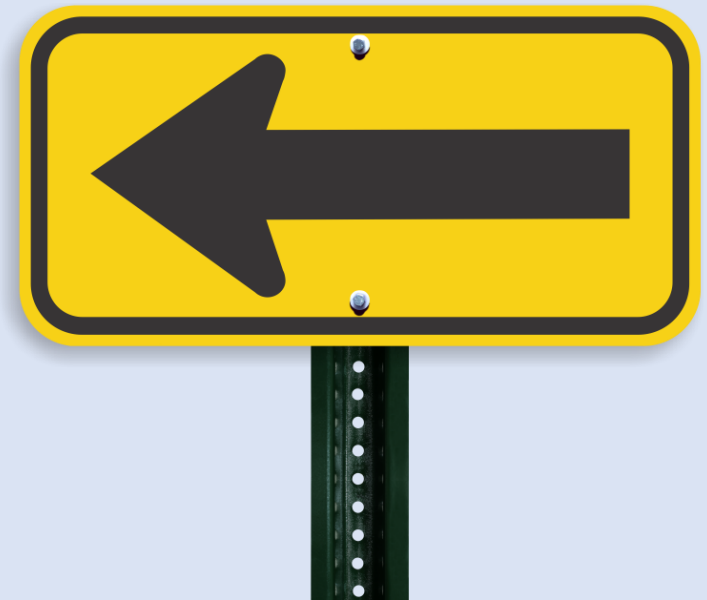
- kontroverzí forma
- obrovská výhoda: 100 %
- má své velké limity
- někdo na něj nedá dopustit
- hádáme se o něm již přes sto let
- [článek o historii koláčového grafu](#)



HOW TO MAKE A PIE CHART IF YOUR PERCENTAGES DON'T ADD UP TO 100

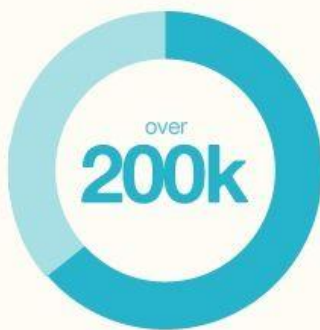
Pár tipů ke koláčům

- omezený počet výsečí
- „rozložení“ nevyužívat
- doporučená posloupnost výsečí
- ...



Koblihové grafy

- americký donut
- problematická forma v každém případě
- střed znemožňuje určité formy čtení



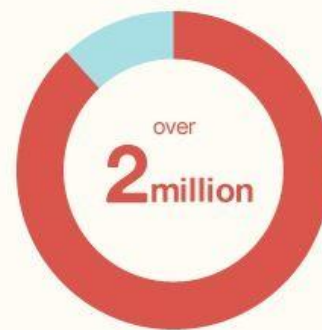
CarWool! users

Already using CarWool! to simplify their car buying experience. We've got the love letters to prove it.



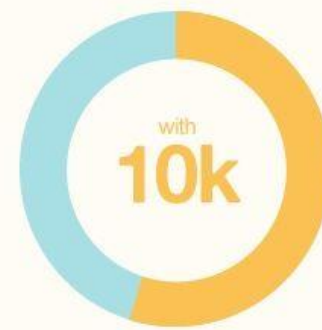
Saved per user

Because we all have better things to do than spend our weekends on a car lot.



Phone calls avoided

Because only your mother should be able to interrupt dinner.

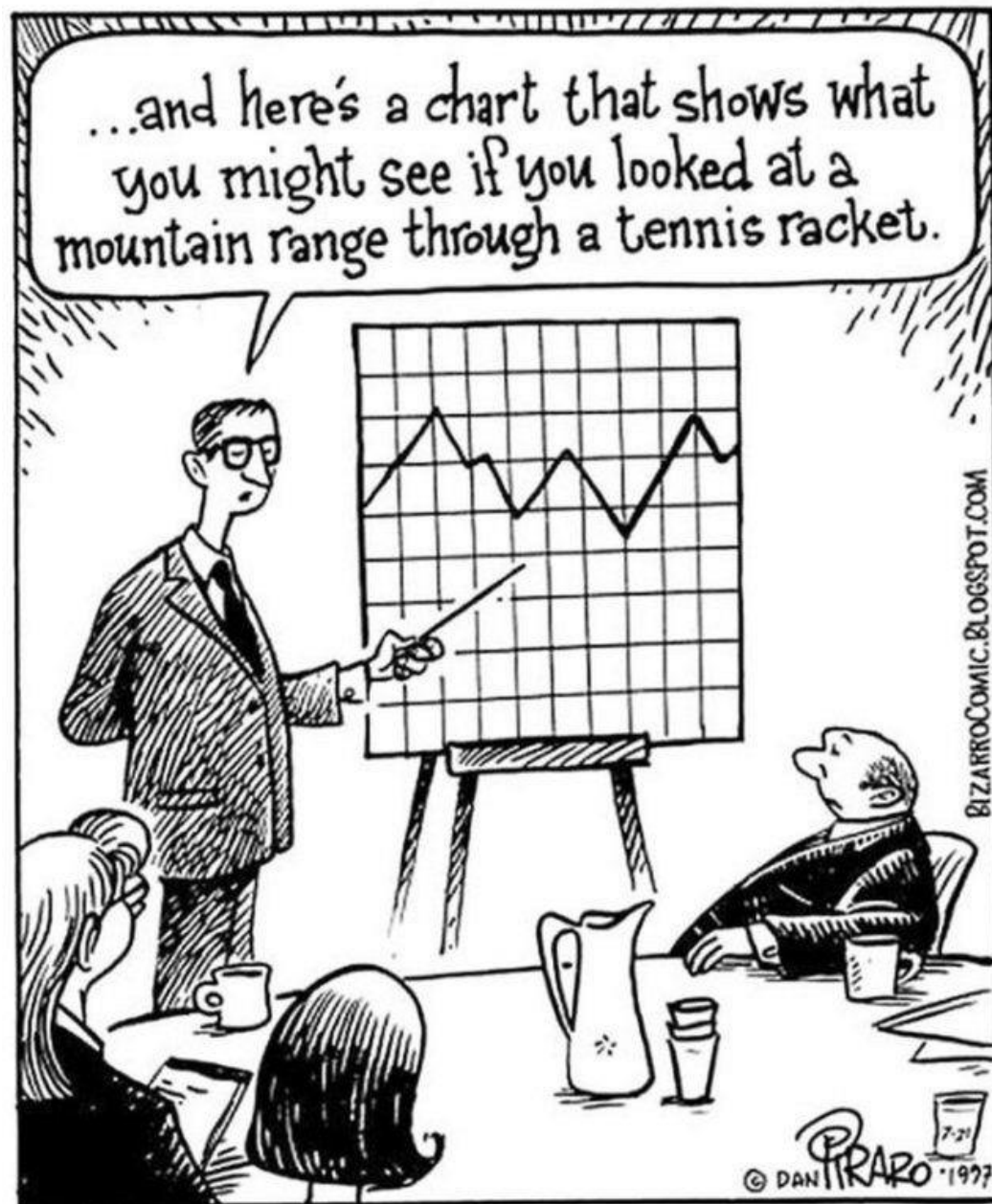


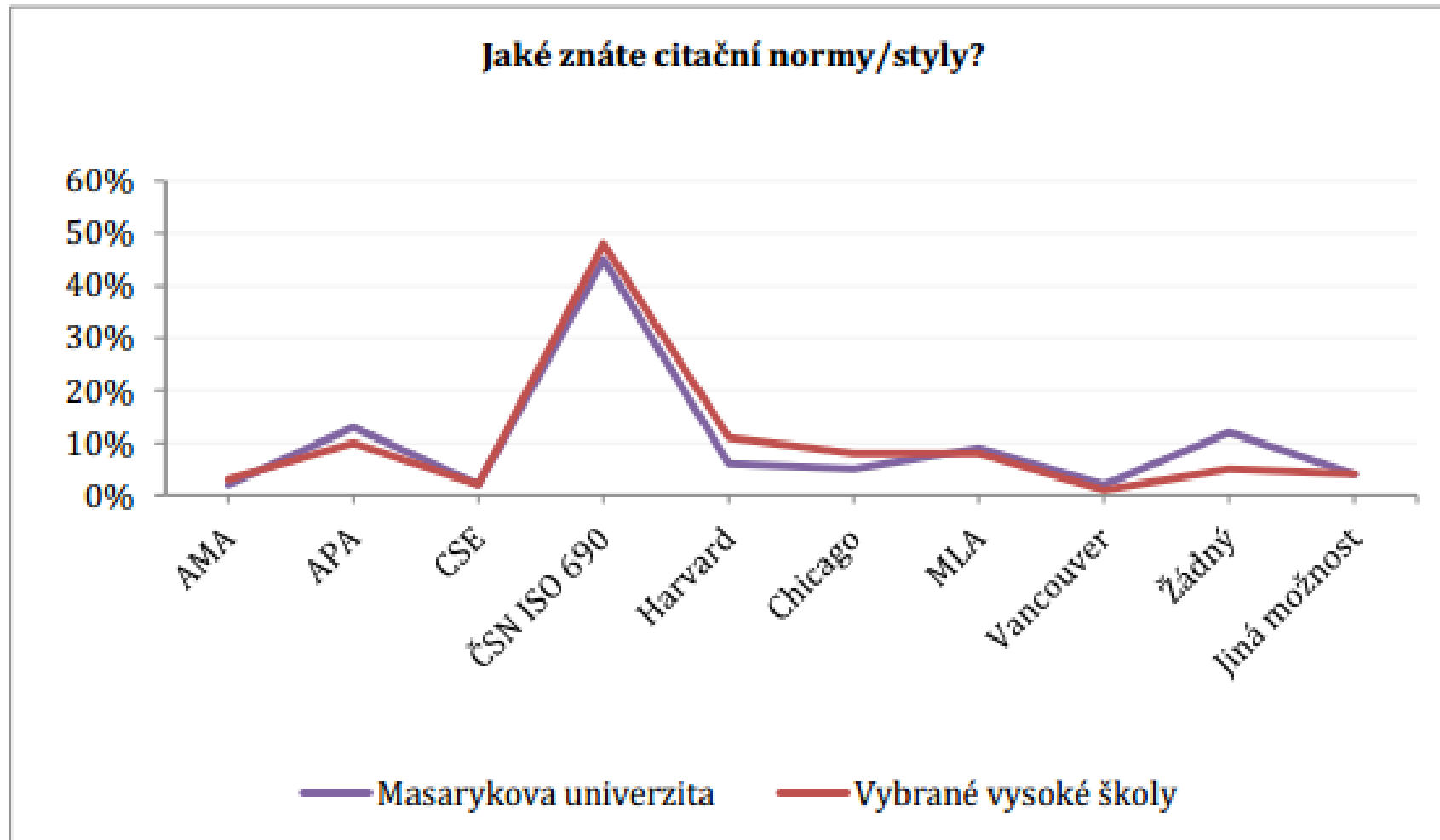
Dealer partners

Ready to communicate with you through the interwebs the way that Al Gore originally intended.

Spojnicové grafy

- trendy v čase
- pouze pro spojité hodnoty!

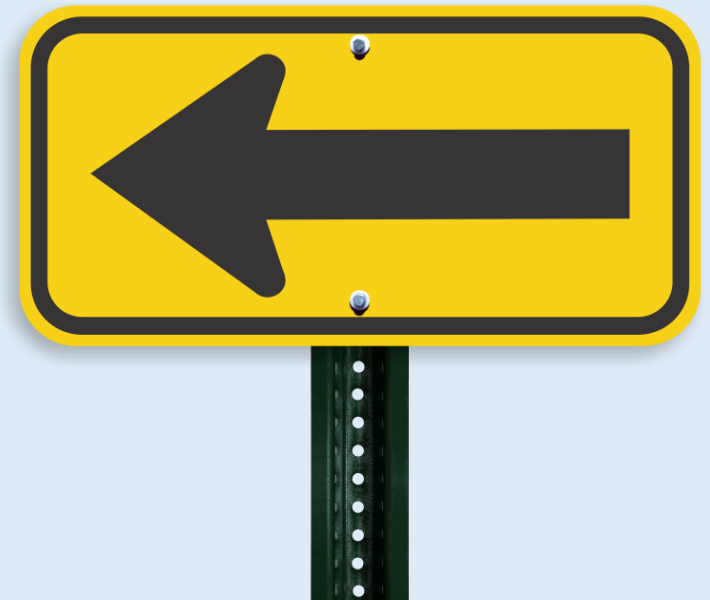




Graf č. 39: Znalost citačních norem - Masarykova univerzita, vybrané vysoké školy

Pár tipů ke spojnicím

- limit počtu linií
- síla spojnic
- ...

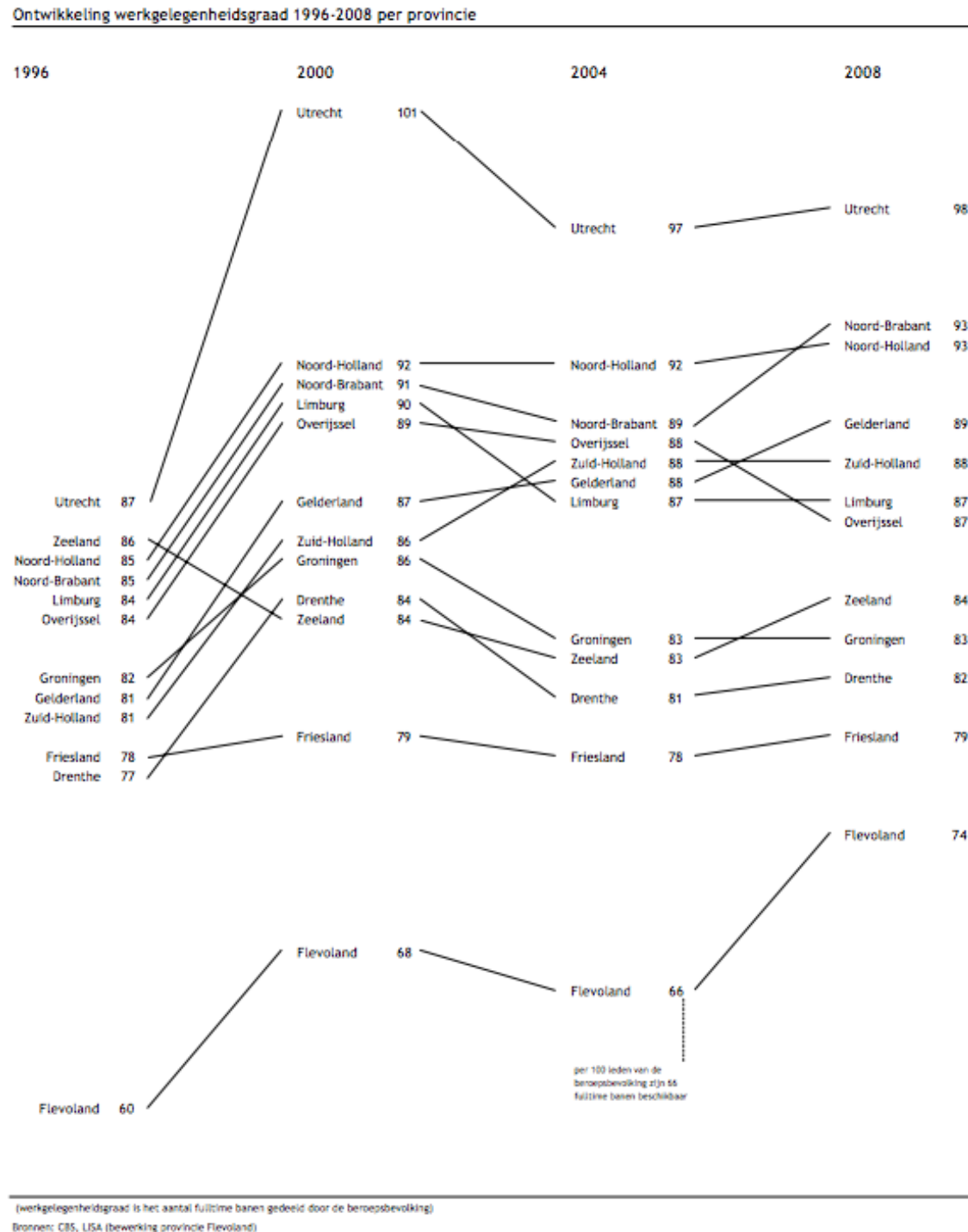


Slopegraph

- v podstatě jen osekáný spojnícový
- velkým propagátorem je Tufte
- dává smysl pro určité případy
- pouze dvě hodnoty na ose x



Current Receipts of Government as a Percentage of Gross Domestic Product, 1970 and 1979



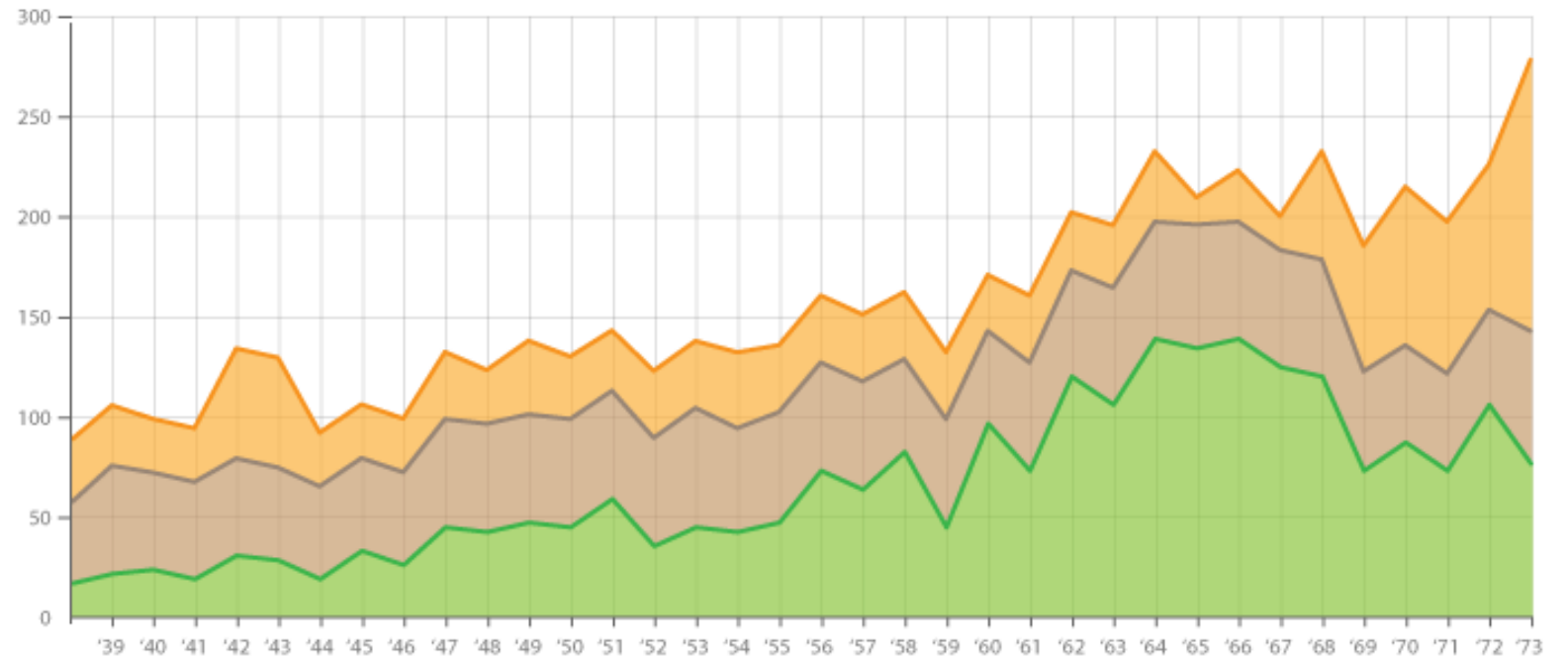
Sparkline

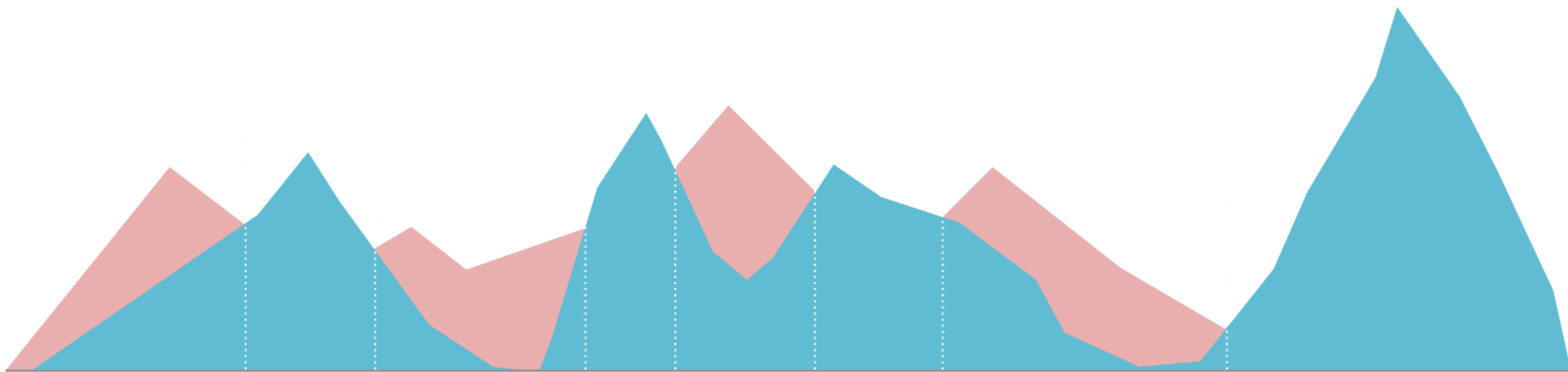


- v Excelu jako *minigrafy*
- velmi malé grafy většinou bez os a legendy
- původně pouze spojnicové
- umožňují rychlou analýzu trendů
- doplňují tabulková data
- [zase Tufte](#)

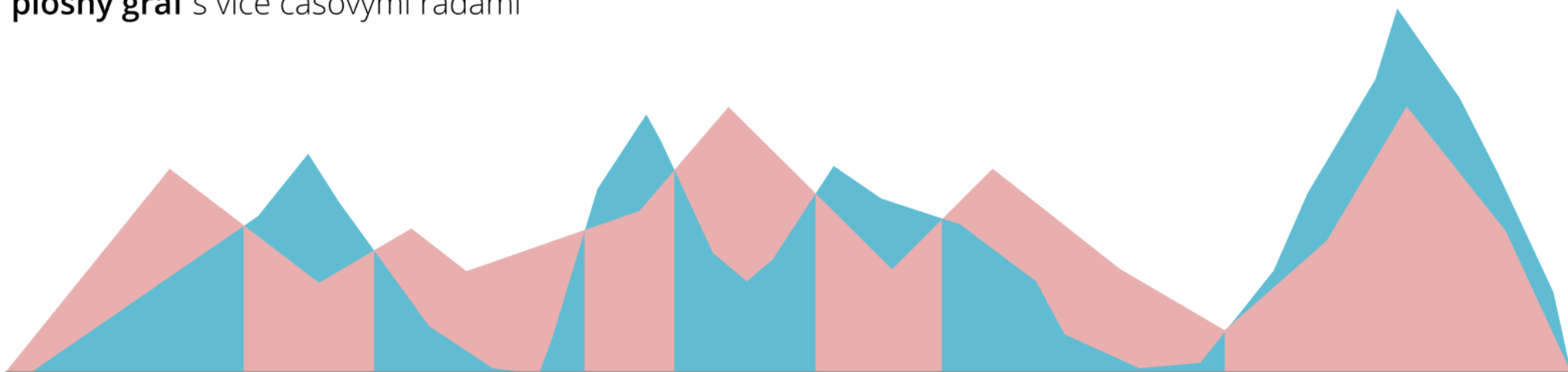
Plošný graf

- *area chart*
- jako spojnicový
- skládaný plošný

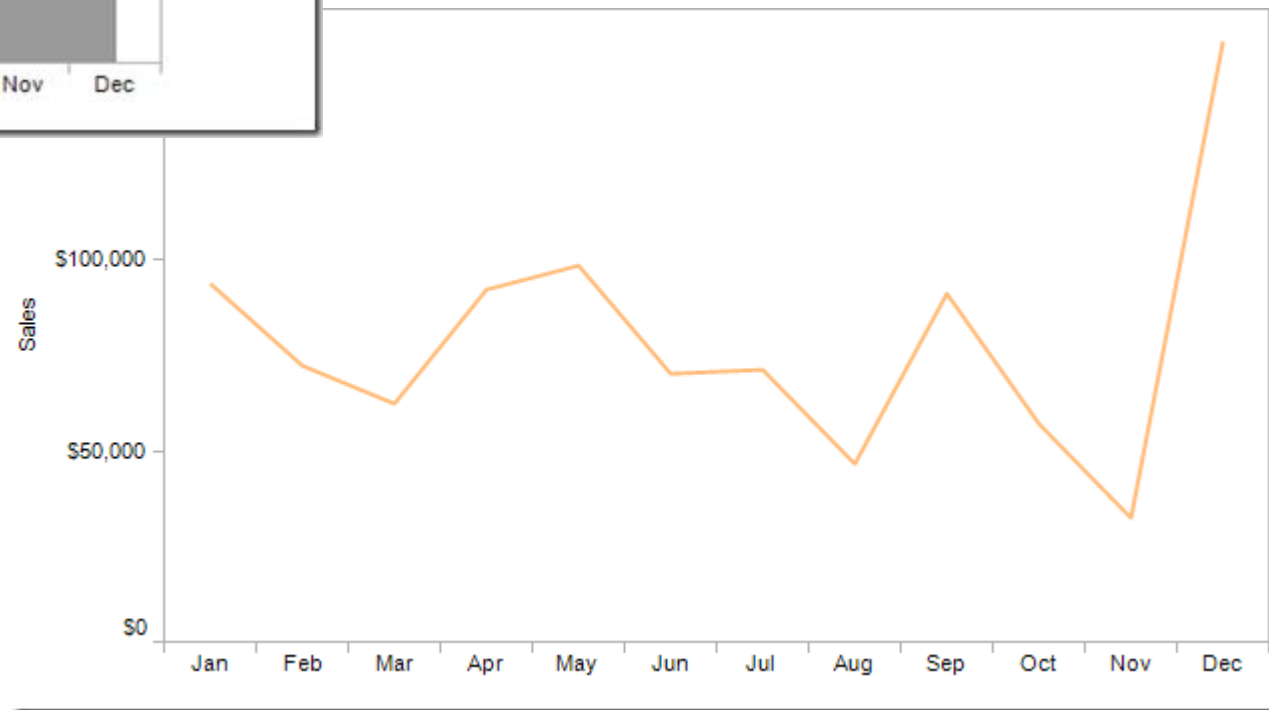
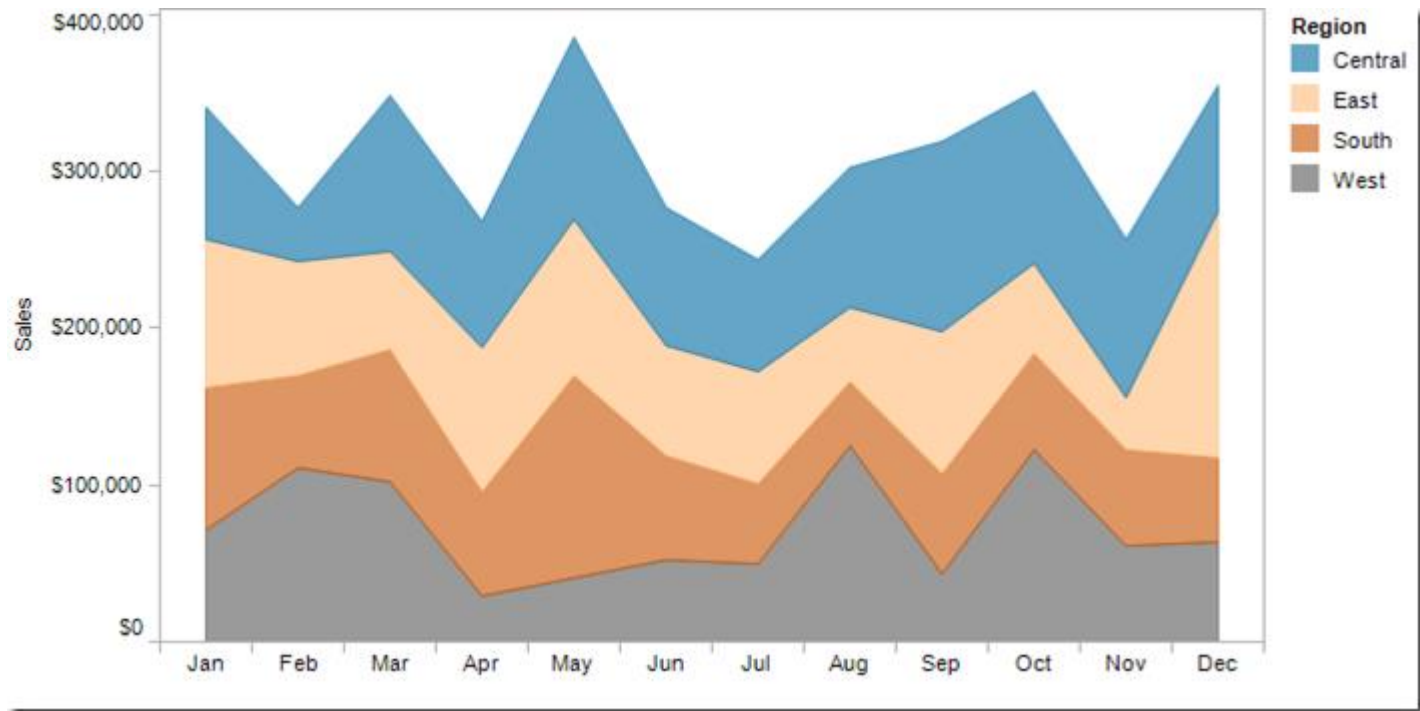




plošný graf s více časovými řadami

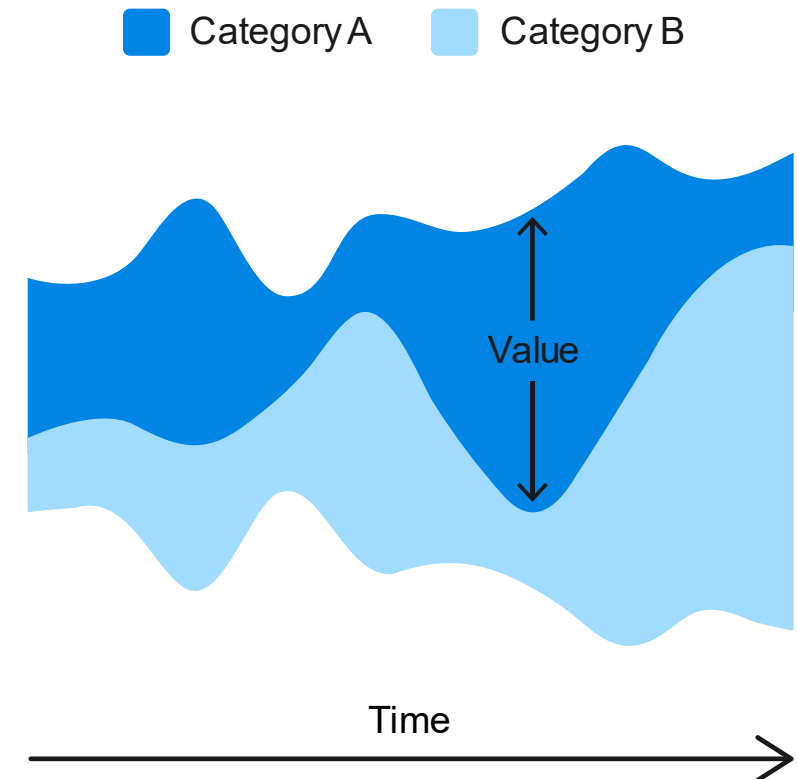


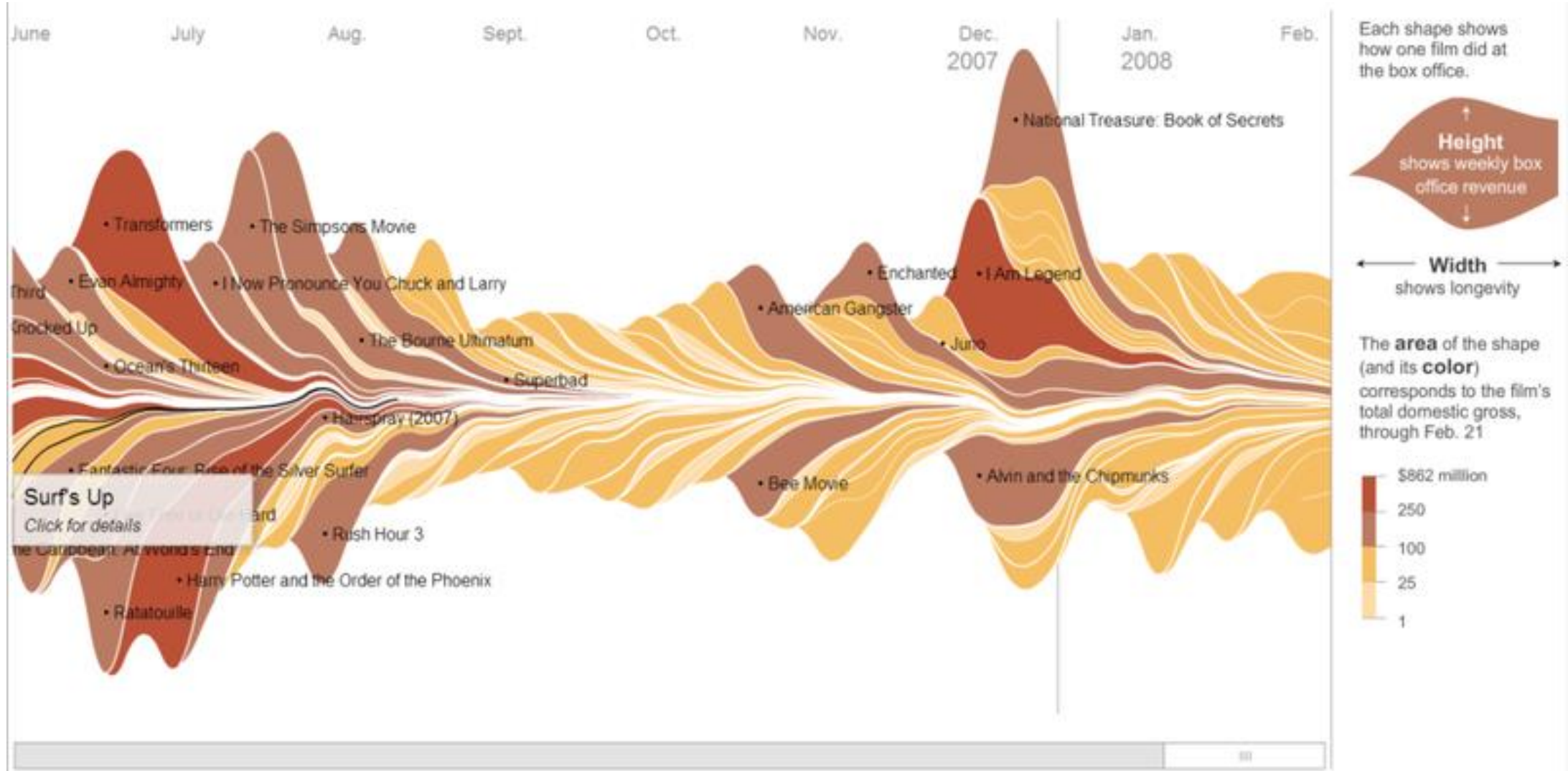
braided graf



Streamgraph

- variace skládaného plošného
- vhodný pro větší data
- rychle se ale zahltí
- spíše popularizační forma
- také jako [*theme river*](#)

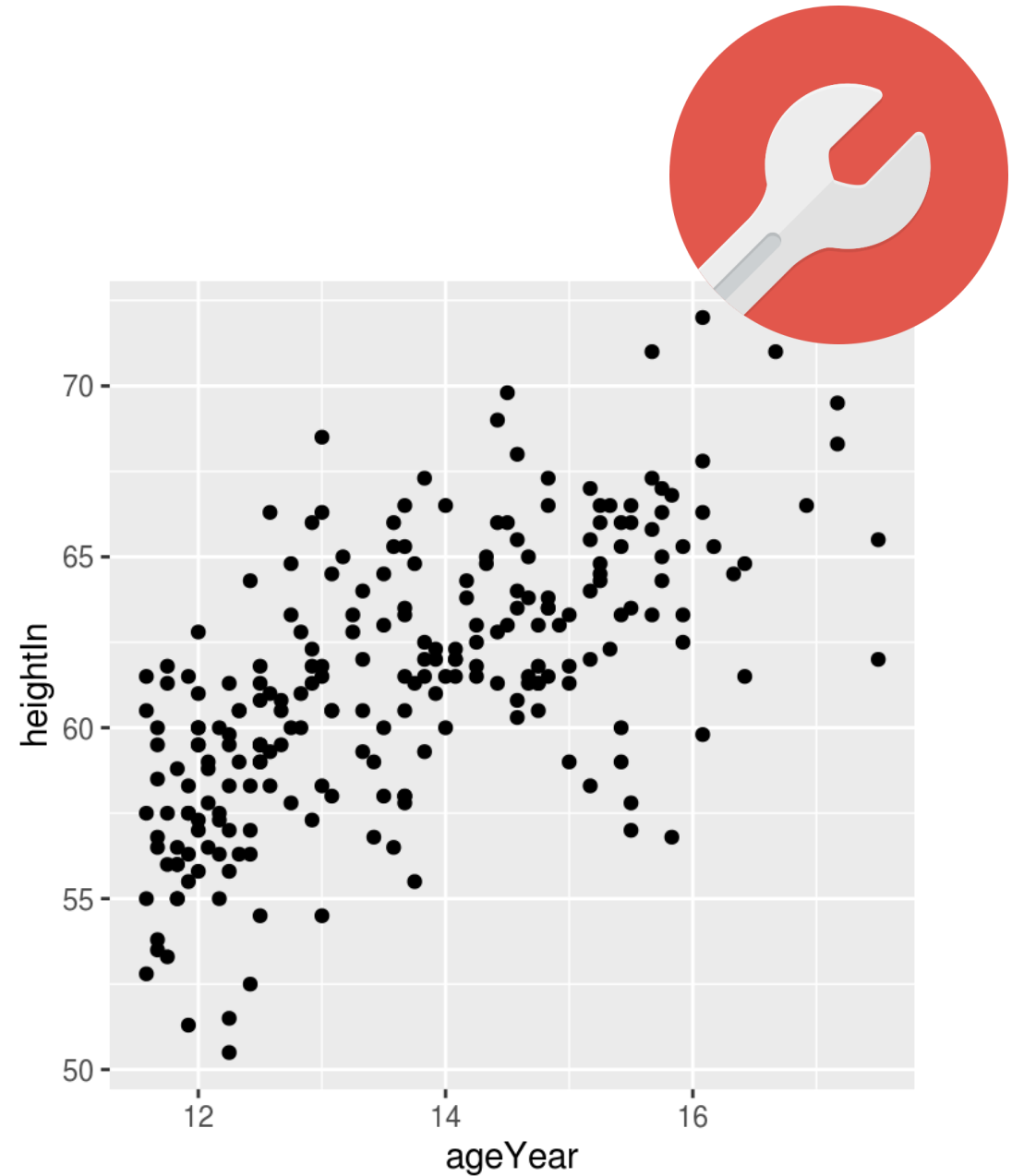


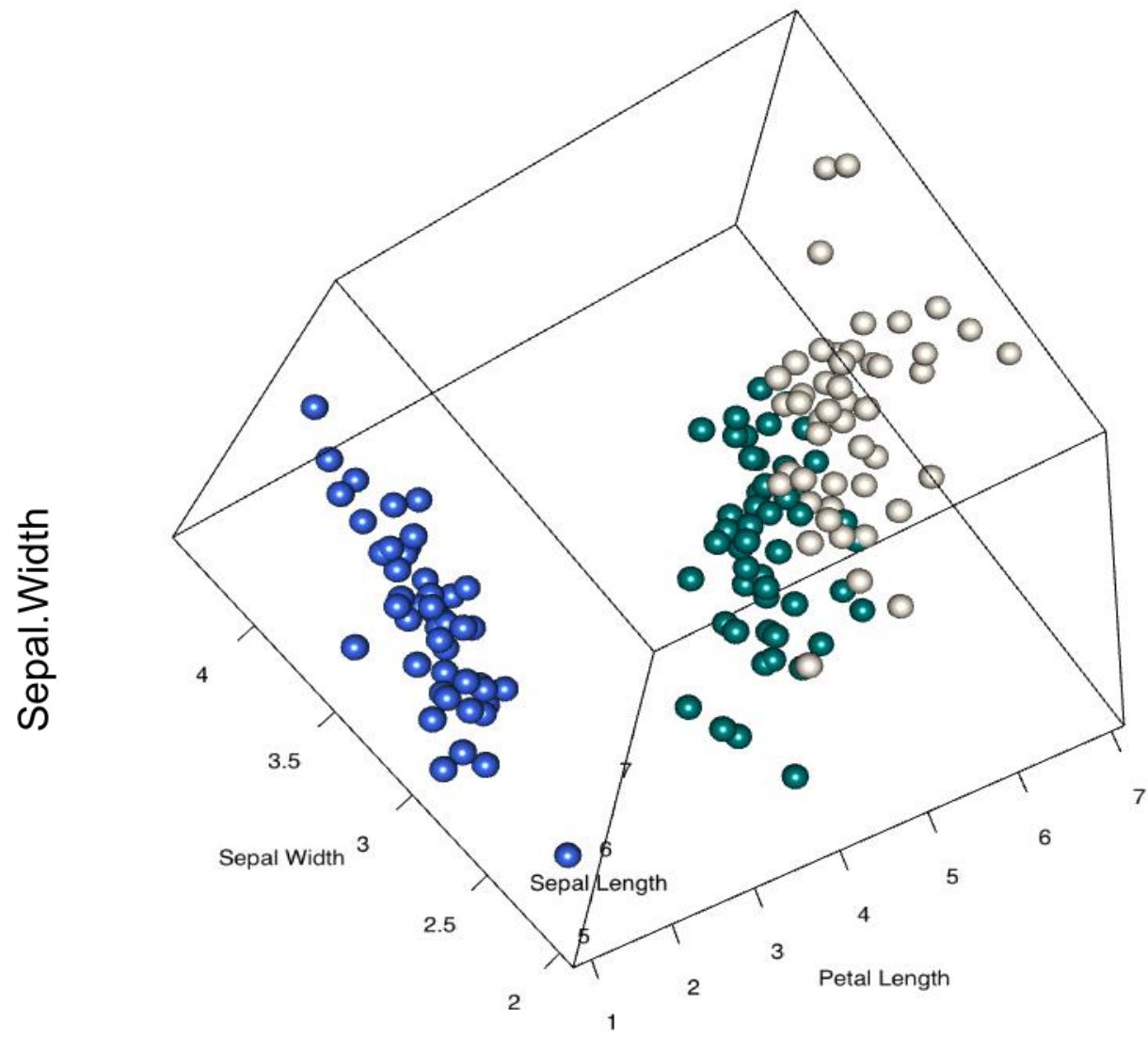
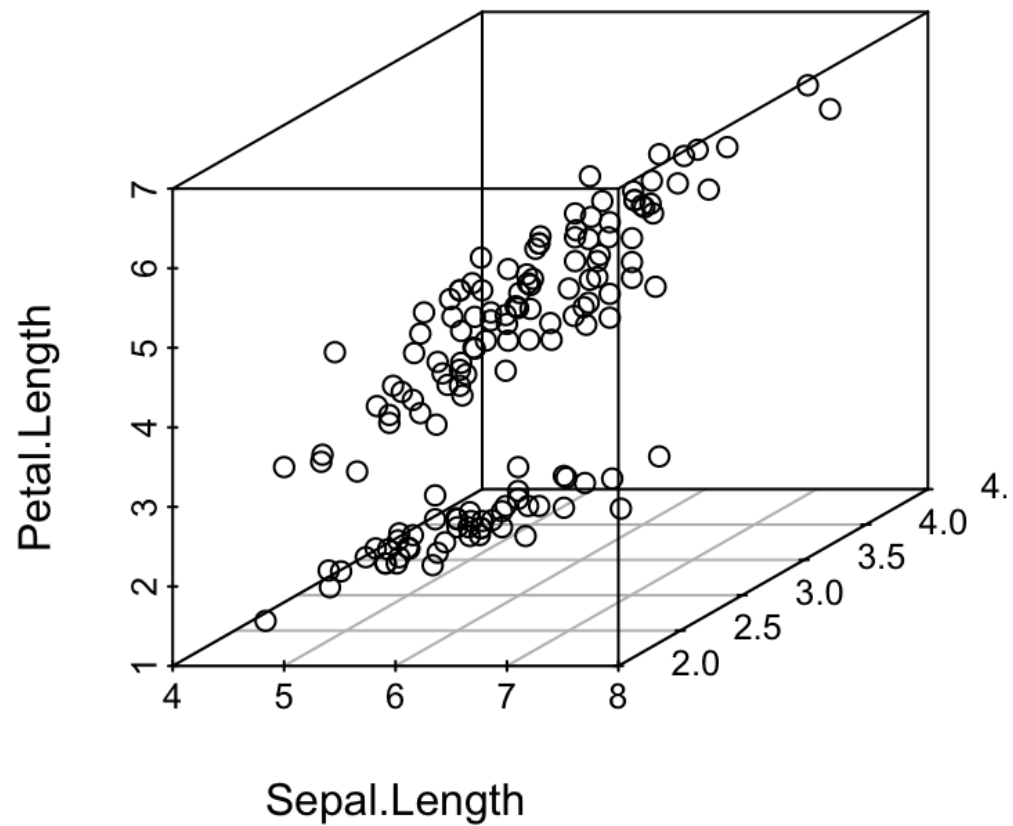


Korelační diagram

- scatterplot, bodový graf
- hodnoty dvou proměnných
- *podmíněné formátování?*
- v Excelu náročné
- *RAWGraphs*

+ velikost bodů = bubble chart

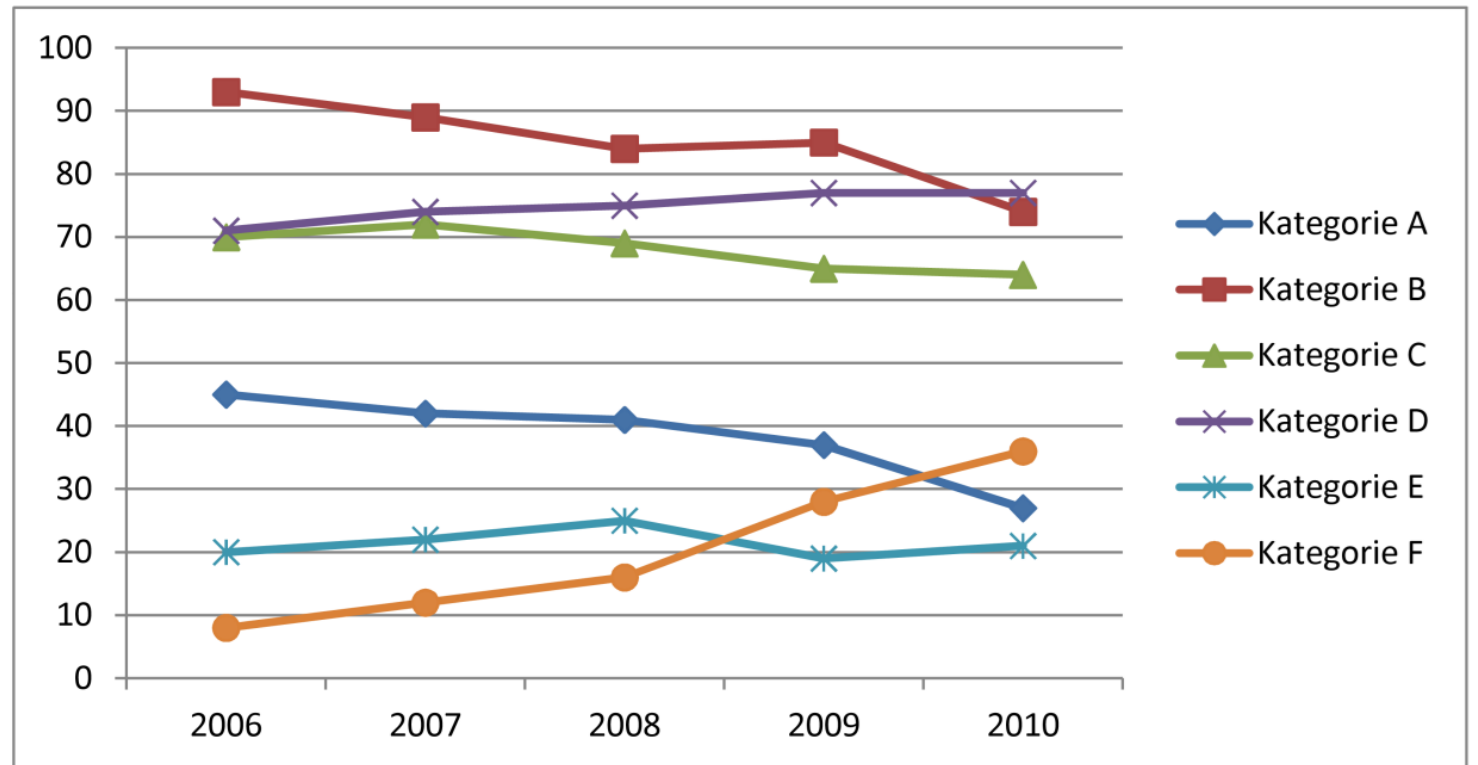


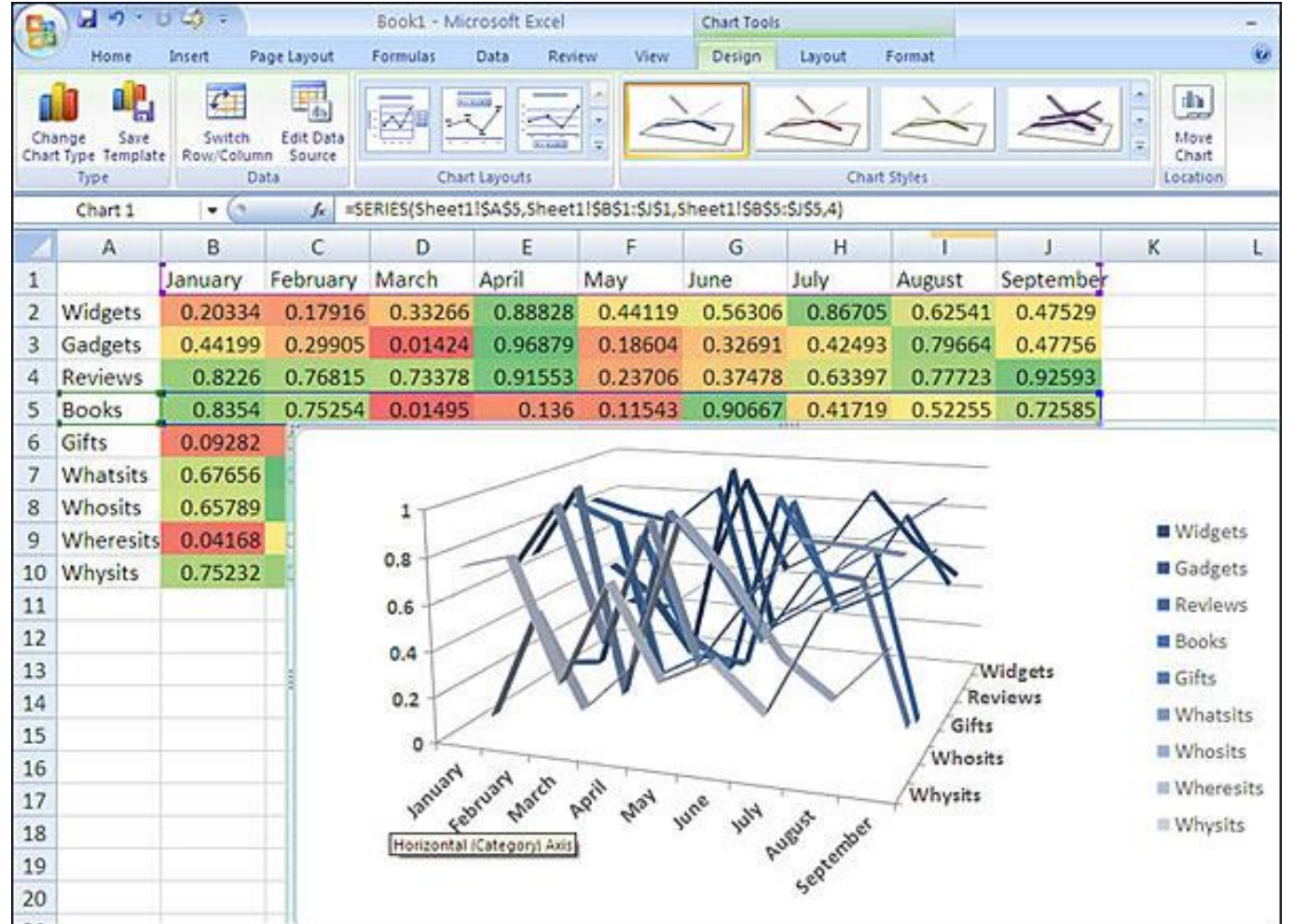
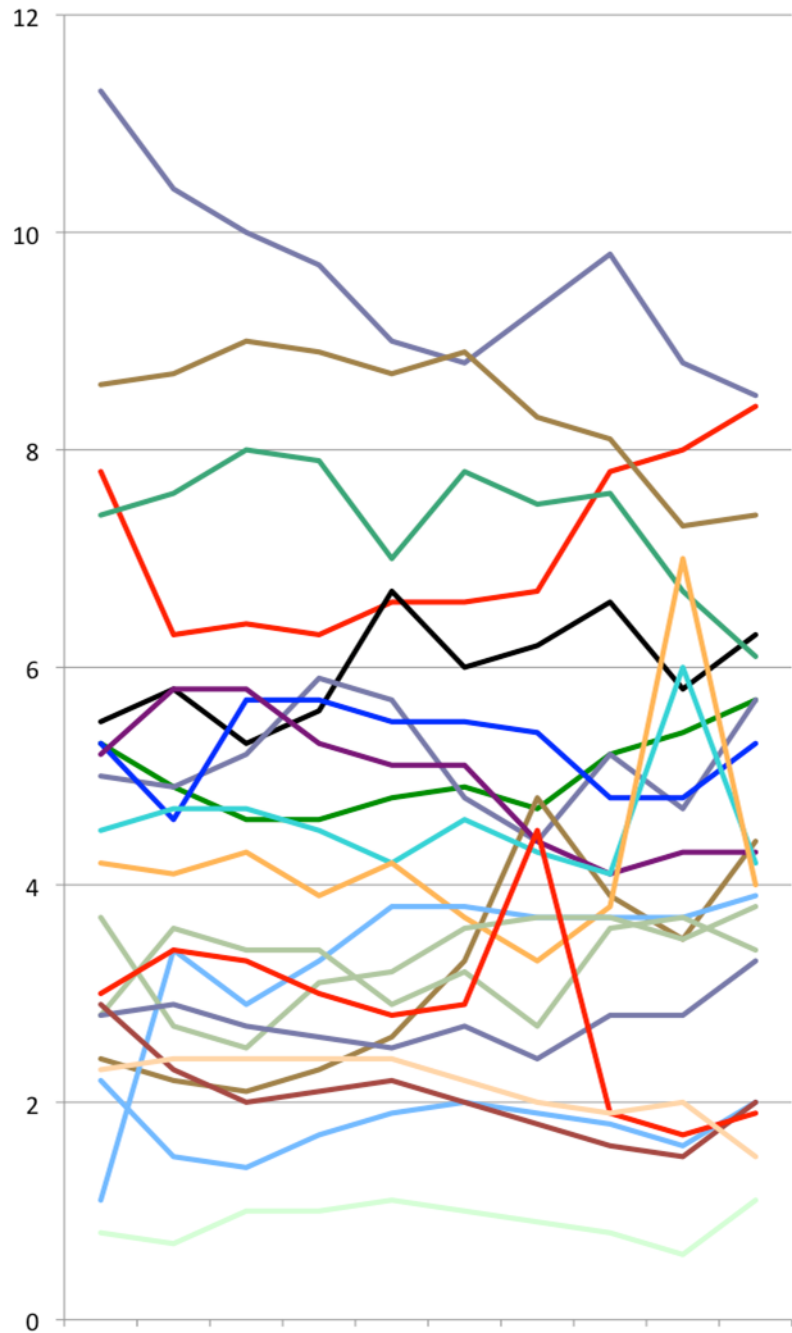


Potíže s mnoha daty

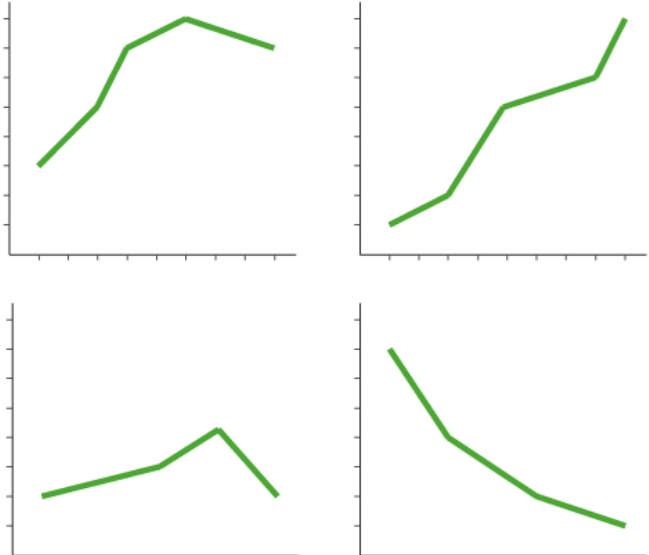
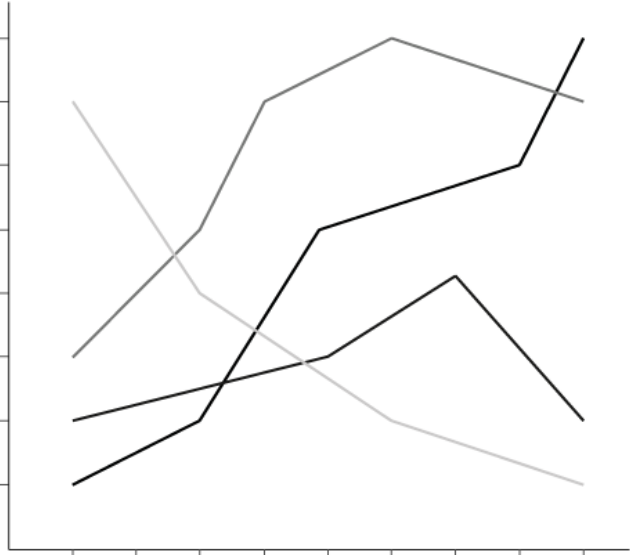
- máme mnoho dat, mnoho záznamů
 - tradiční formy na to nestačí
 - naráží na své limity
 - co dál?
-
- *small multiples*
 - specifické formy – horizon chart

	2006	2007	2008	2009	2010
Kategorie A	45	42	41	37	27
Kategorie B	93	89	84	85	74
Kategorie C	70	72	69	65	64
Kategorie D	71	74	75	77	77
Kategorie E	20	22	25	19	21
Kategorie F	8	12	16	28	36



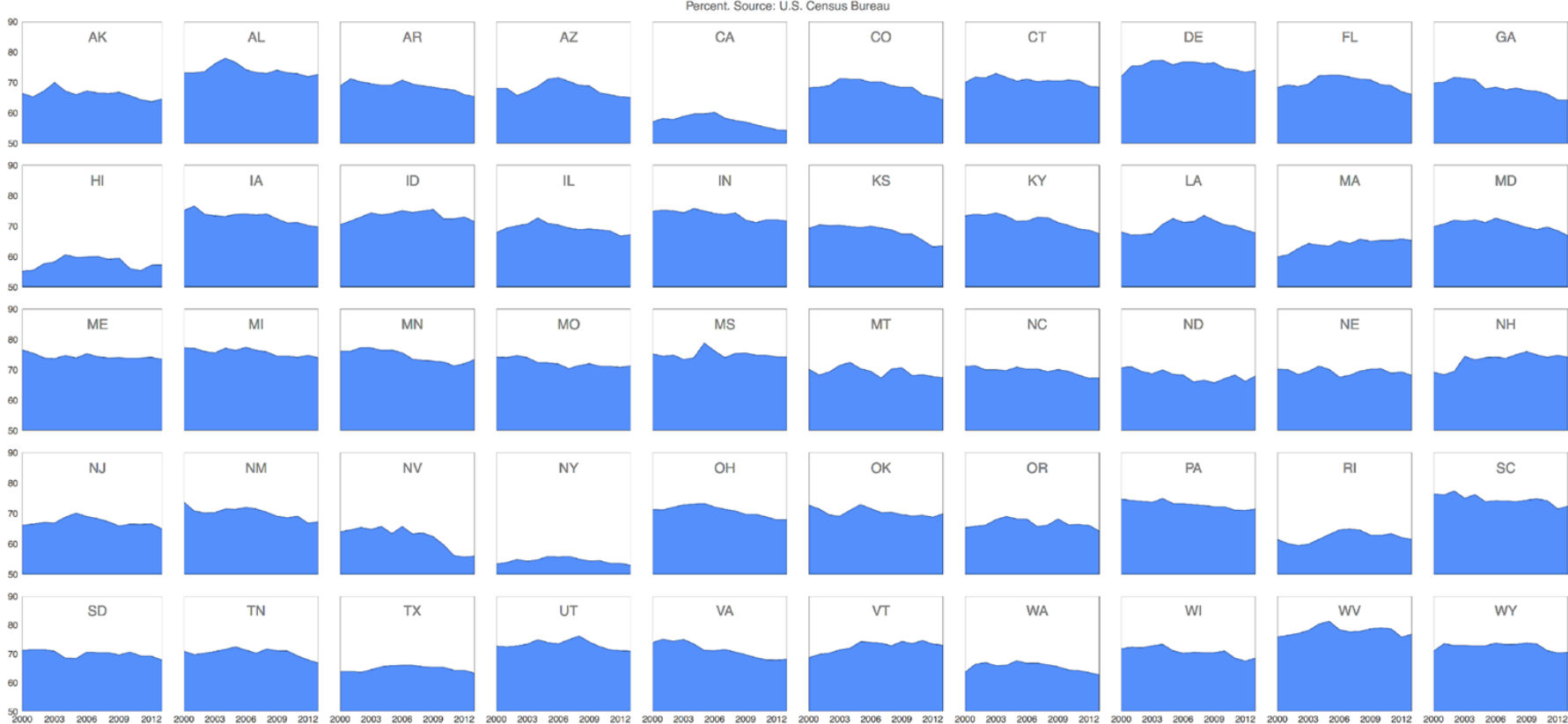


small multiples



small multiples

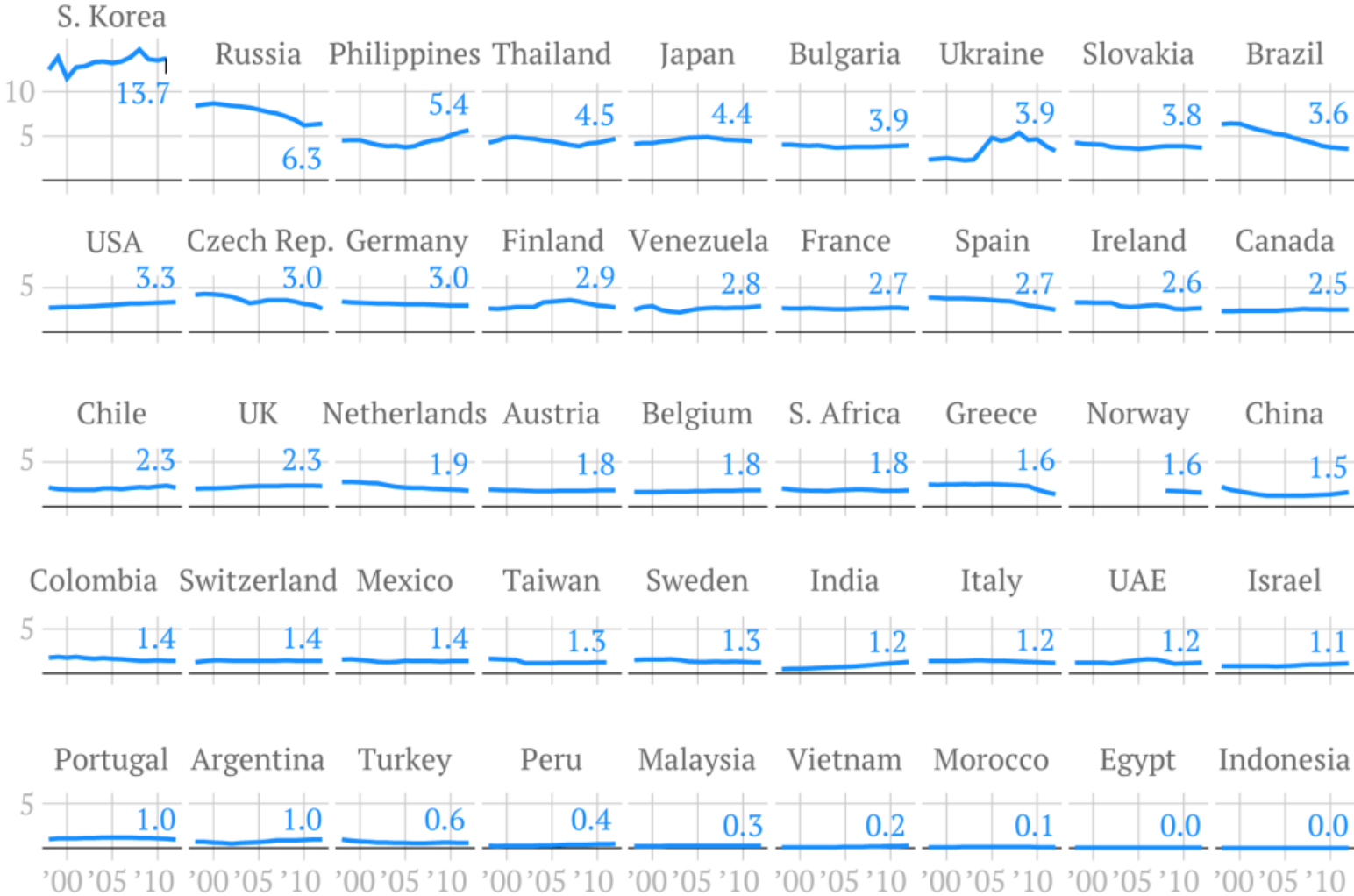
Evolution Of Homeownership Rates Across The United States, 2000–2013



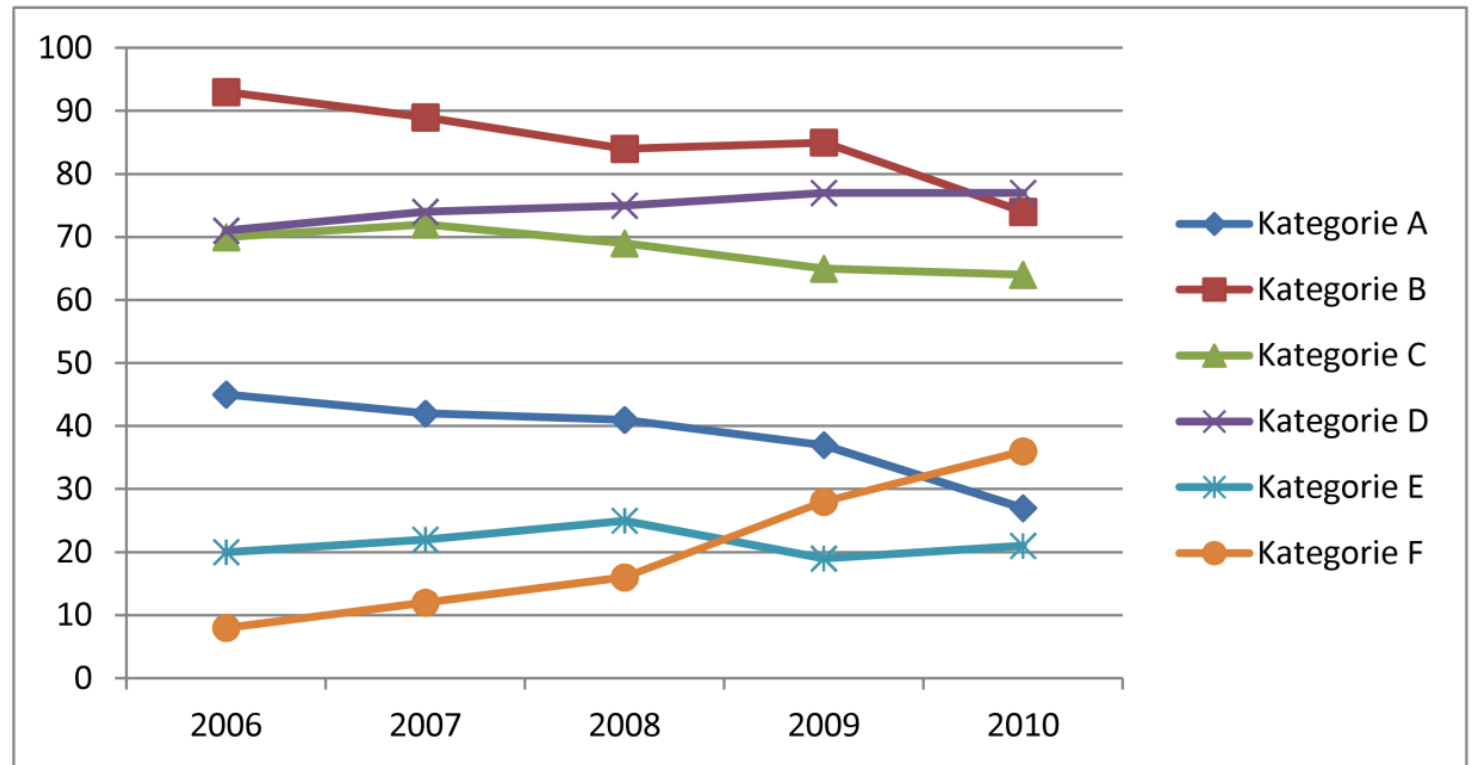
small multiples

The average amount of liquor consumed by a person of drinking age

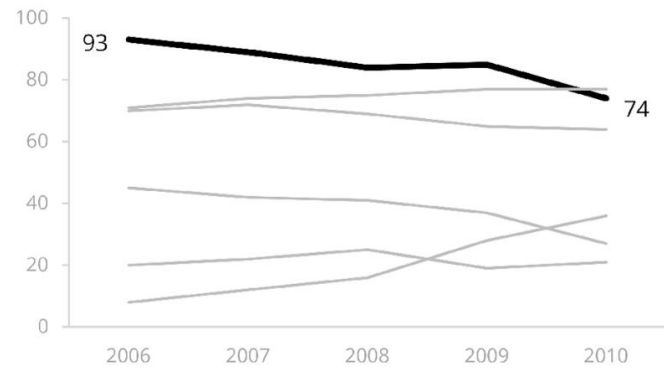
Shots per week of any spirit



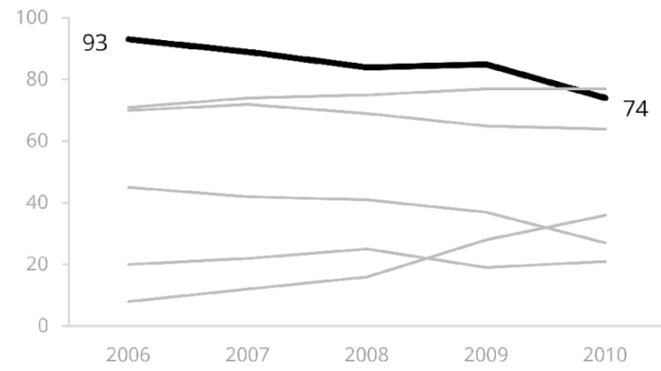
	2006	2007	2008	2009	2010
Kategorie A	45	42	41	37	27
Kategorie B	93	89	84	85	74
Kategorie C	70	72	69	65	64
Kategorie D	71	74	75	77	77
Kategorie E	20	22	25	19	21
Kategorie F	8	12	16	28	36



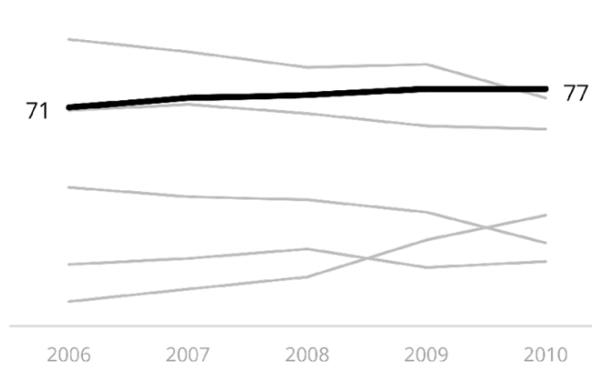
Varianta B



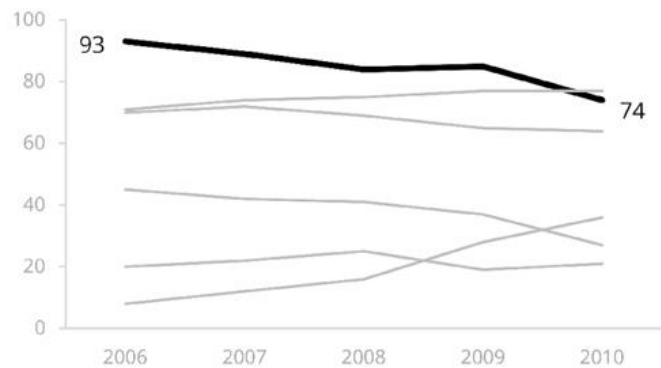
Varianta B



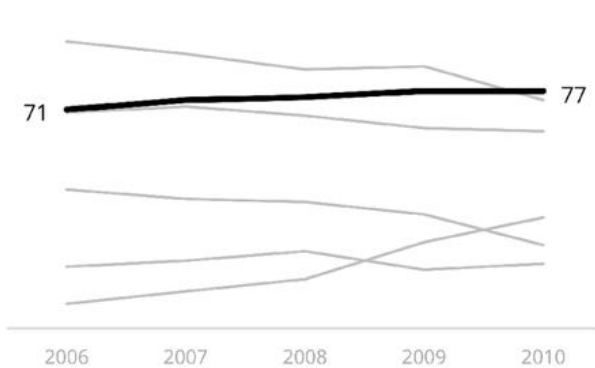
Varianta D



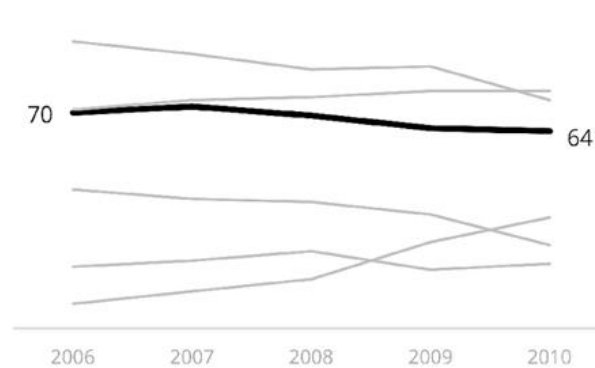
Varianta B



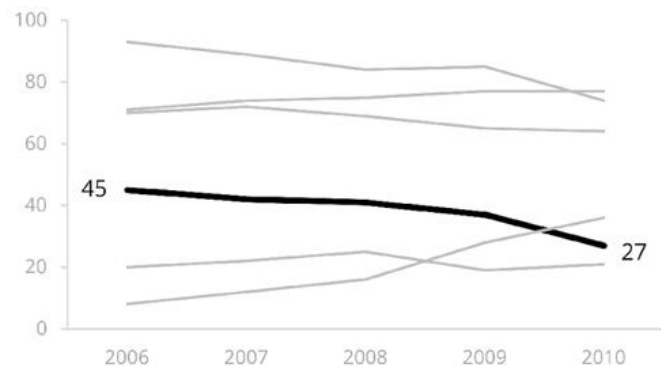
Varianta D



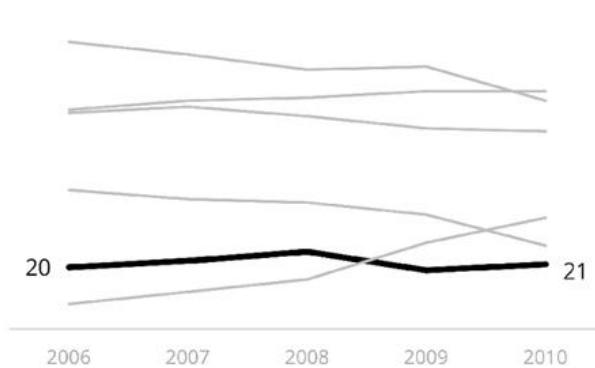
Varianta C



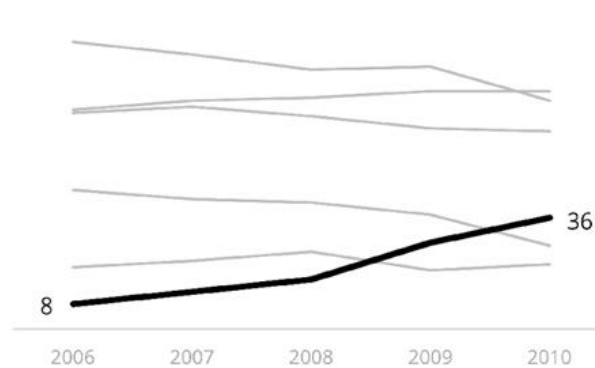
Varianta A



Varianta E

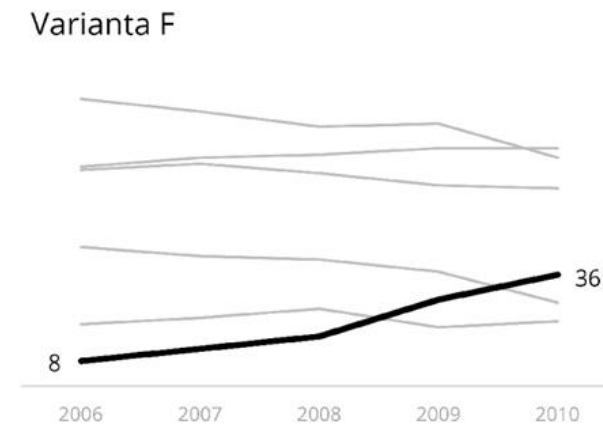
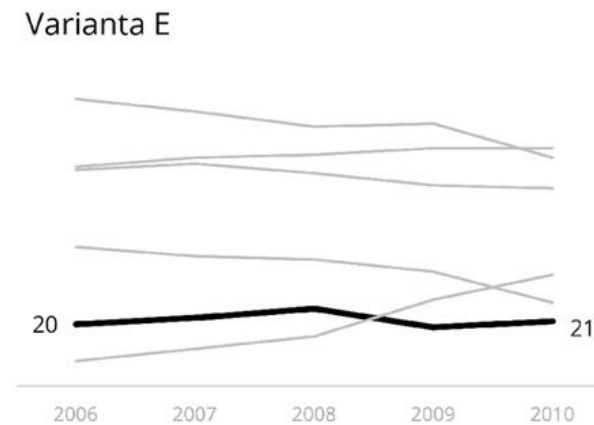
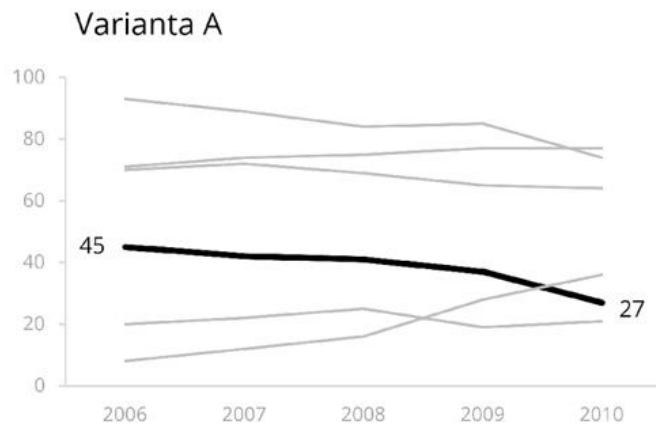
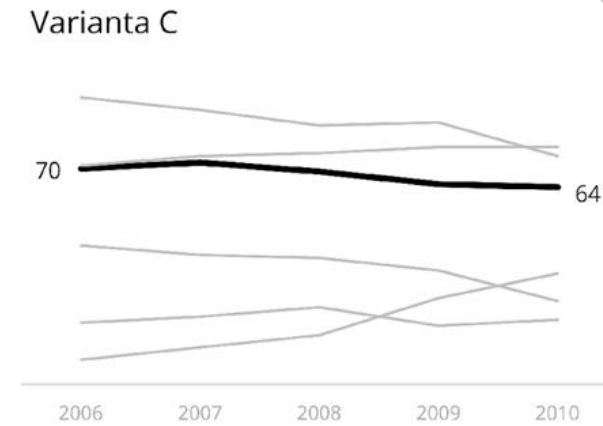
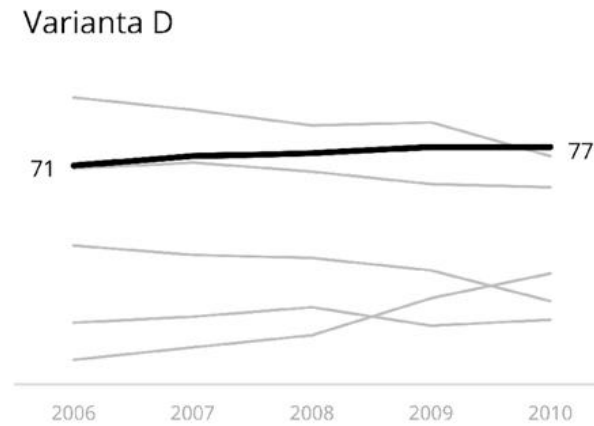
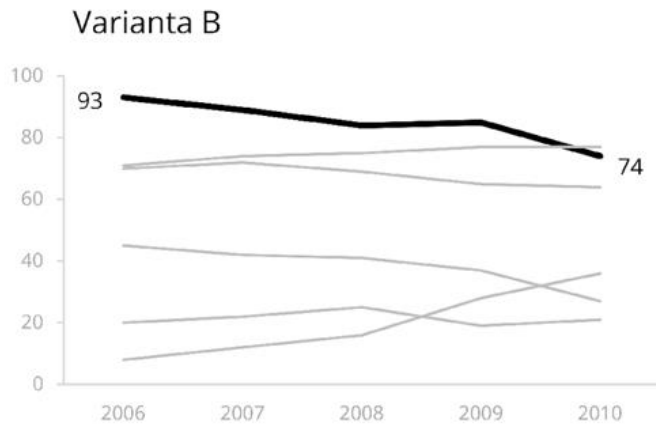


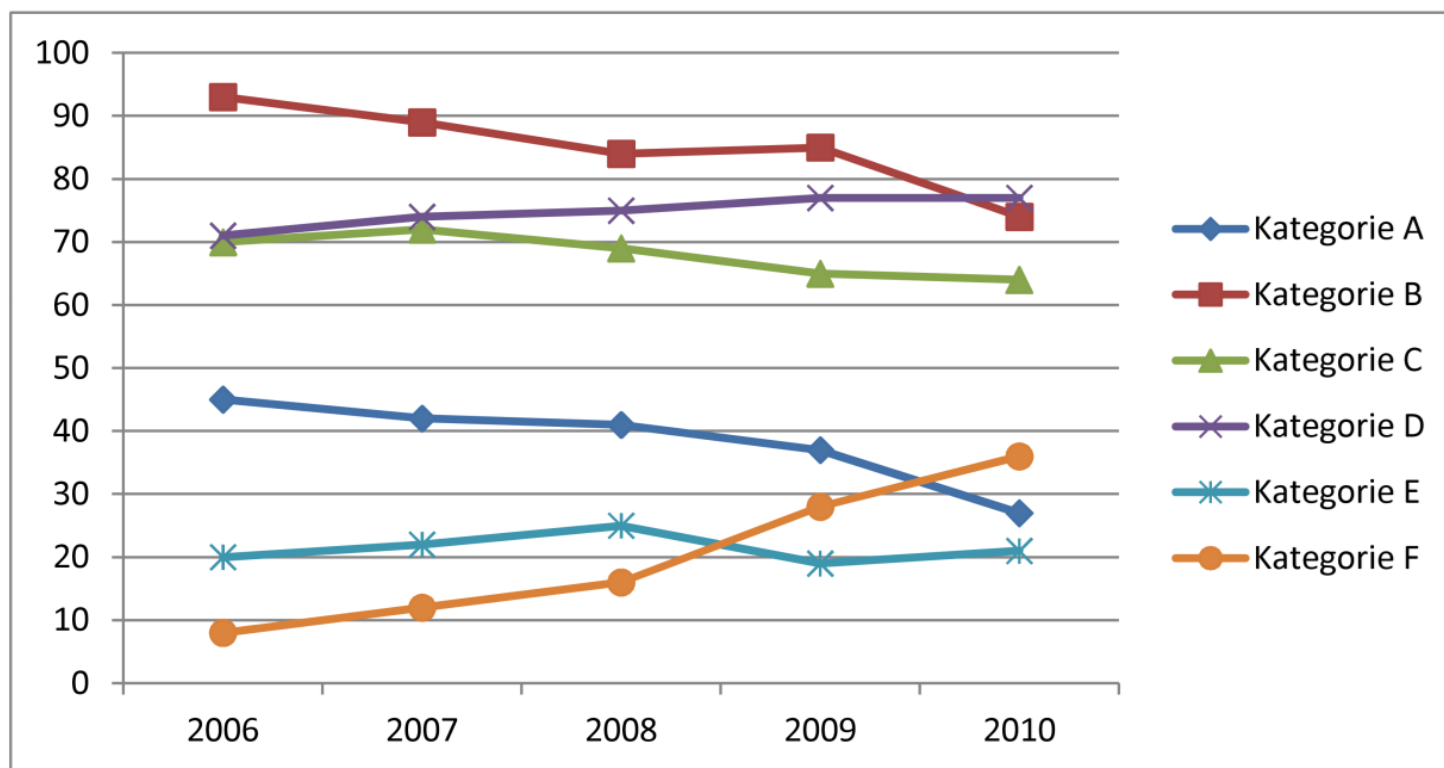
Varianta F



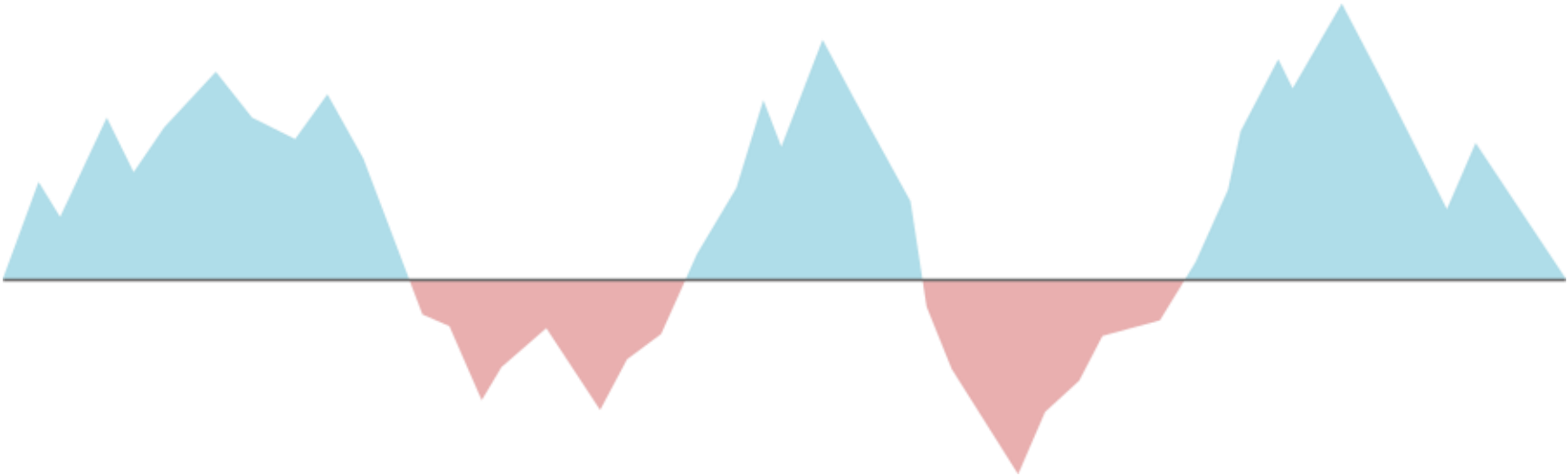
Varianty A a E nejsou ekonomicky perspektivní

Zatímco varianta F, kterou jsme po spuštění výroby považovali za neperspektivní, na trhu nalézá odbyt, výrobu dlouhodobě průměrných variant A a E doporučujeme zastavit.

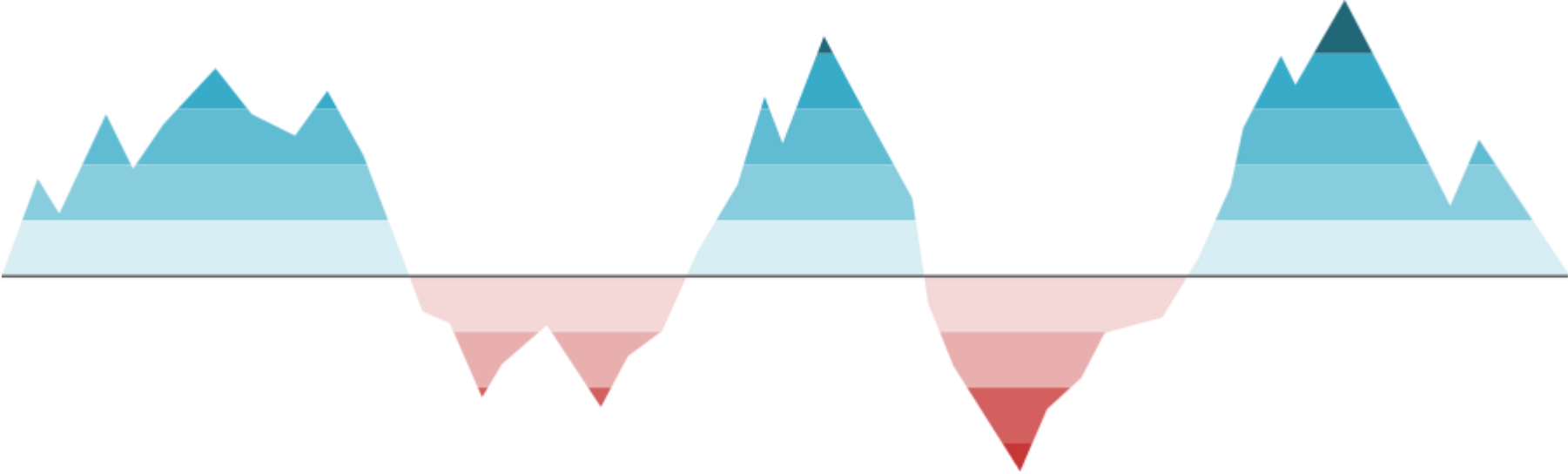




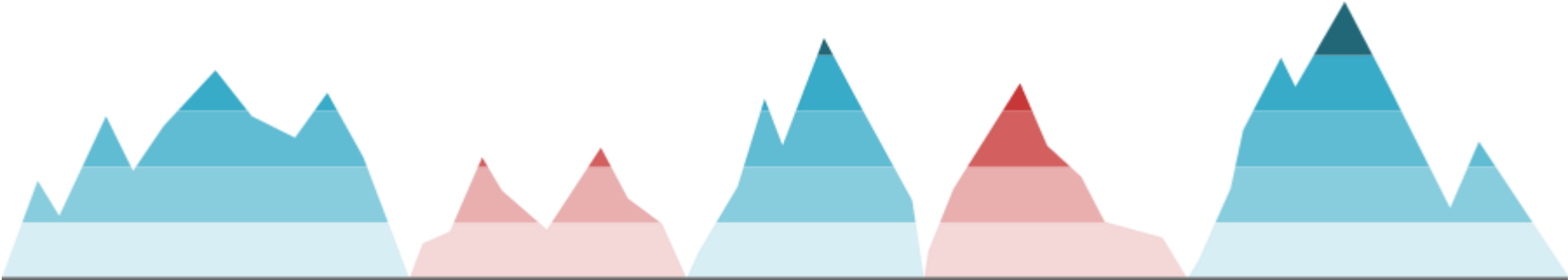
horizon chart



horizon chart



horizon chart



horizon chart



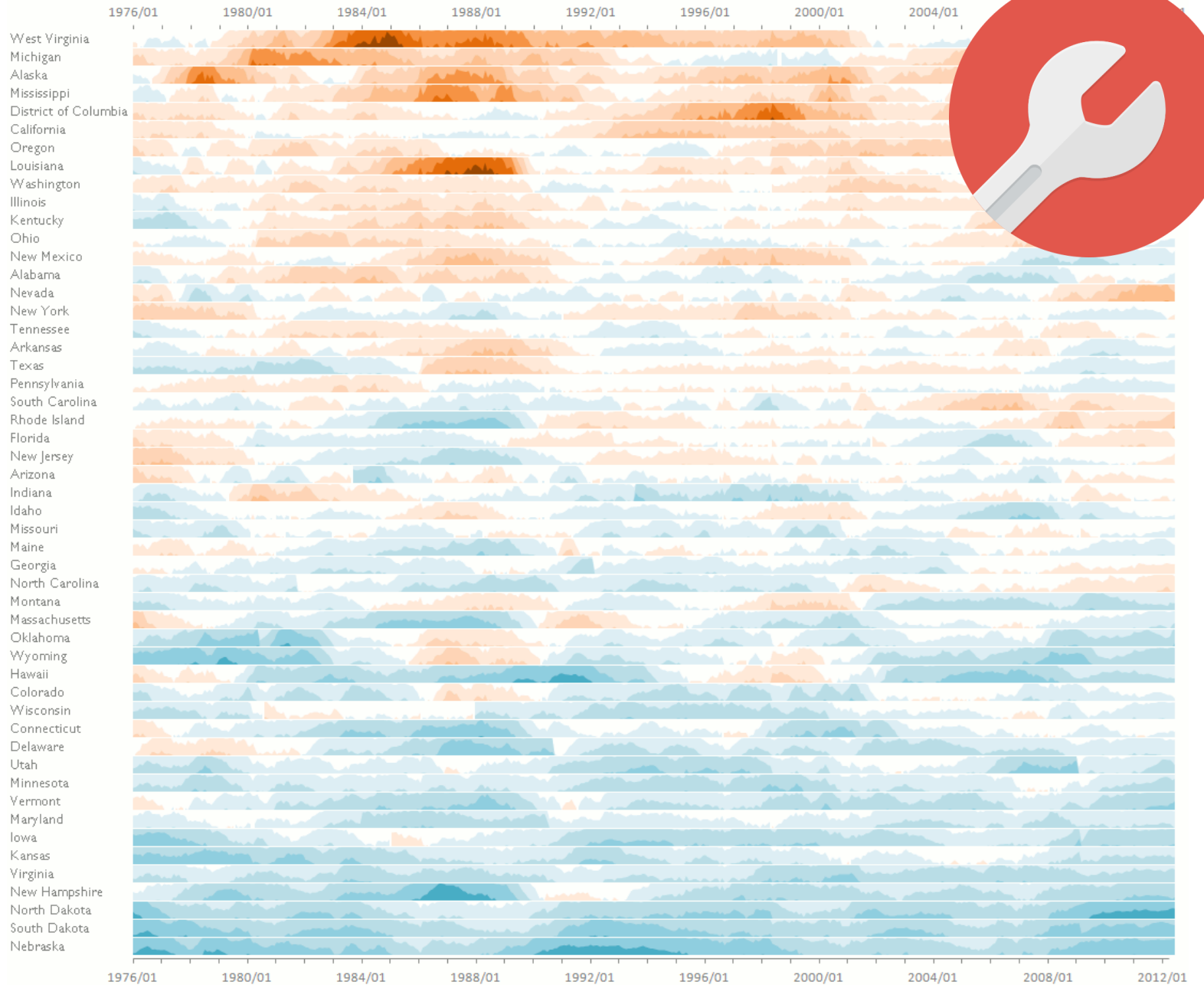
zrcadlené negativní hodnoty



offset negativní hodnoty

horizon chart

Unemployment Rate: Differences to the National Average



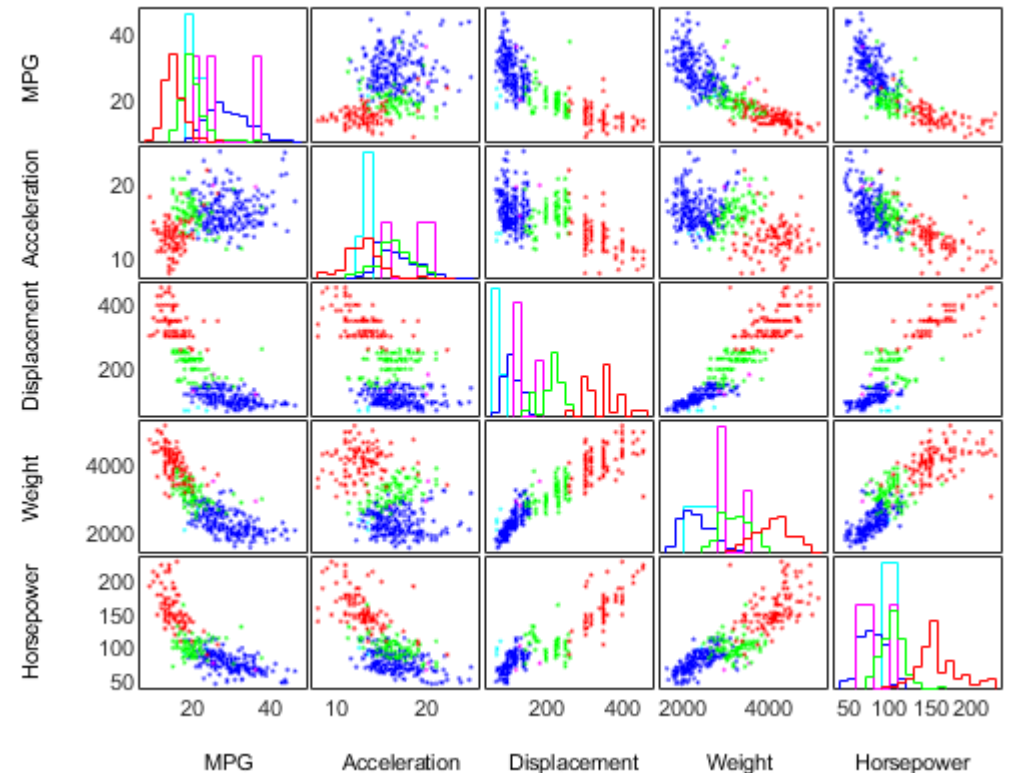
Vizualizace vícerozměrných dat

- vizualizace vícerozměrných dat je velmi náročná
- většina metod je cenná, ale mají své nevýhody

- *radar chart*
- *rovnoběžné souřadnice*
- *chernoff faces*

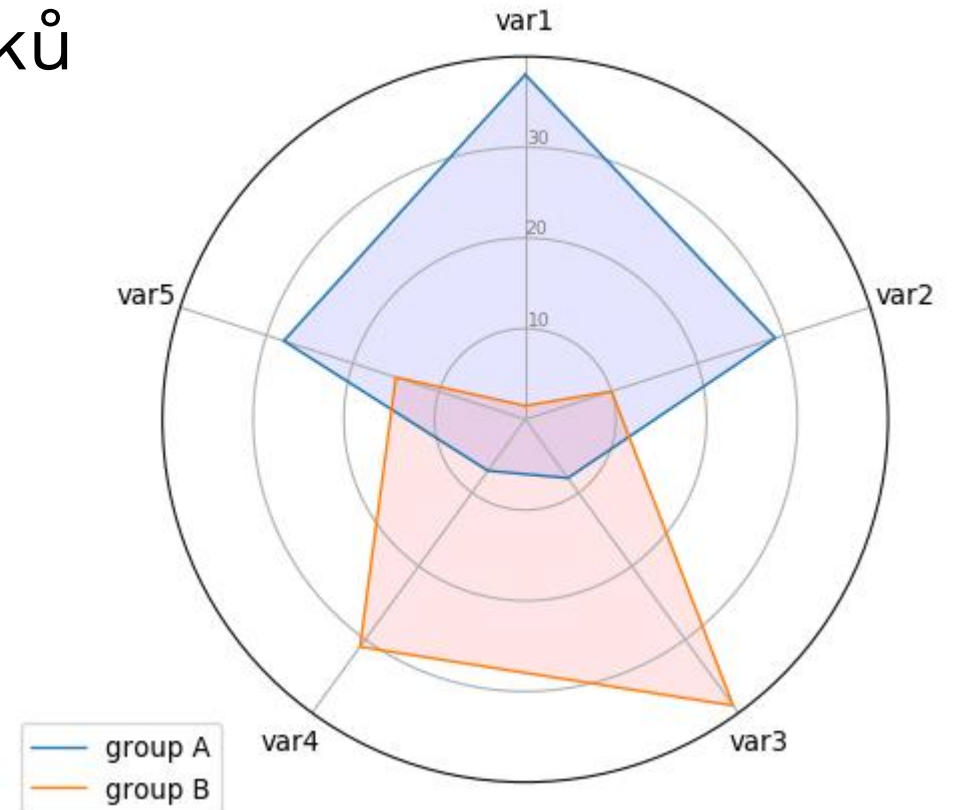
Vícerozměrný korelační diagram

- *multivariate scatter plot*
- matice korelačních diagramů
- čím více proměnných, tím více řádků



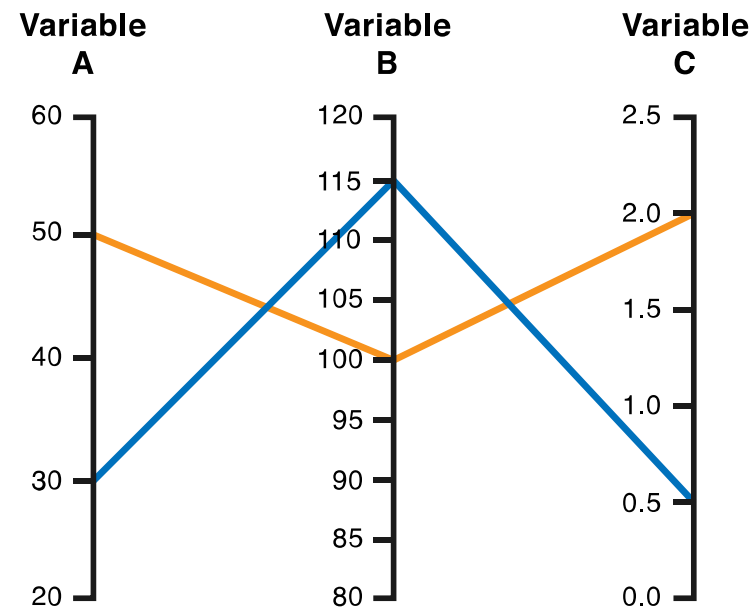
Radar chart

- relativně těžko se čte
- počet proměnných = počet paprsků

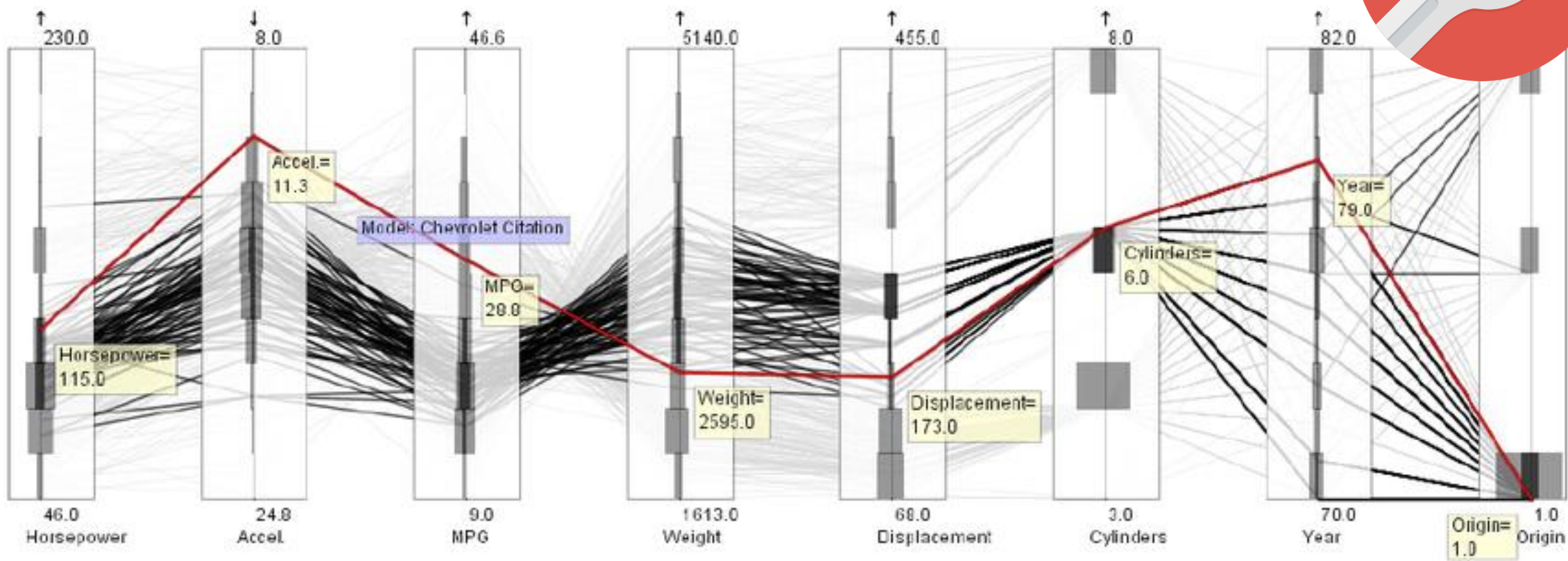


Parallel coordinates

- *rovnoběžné souřadnice*
- vícerozměrná data
- jedna položka s více proměnnými
- dobré k odhalení vztahů mezi p
- obtížné čtení
- *RAWGraphs*



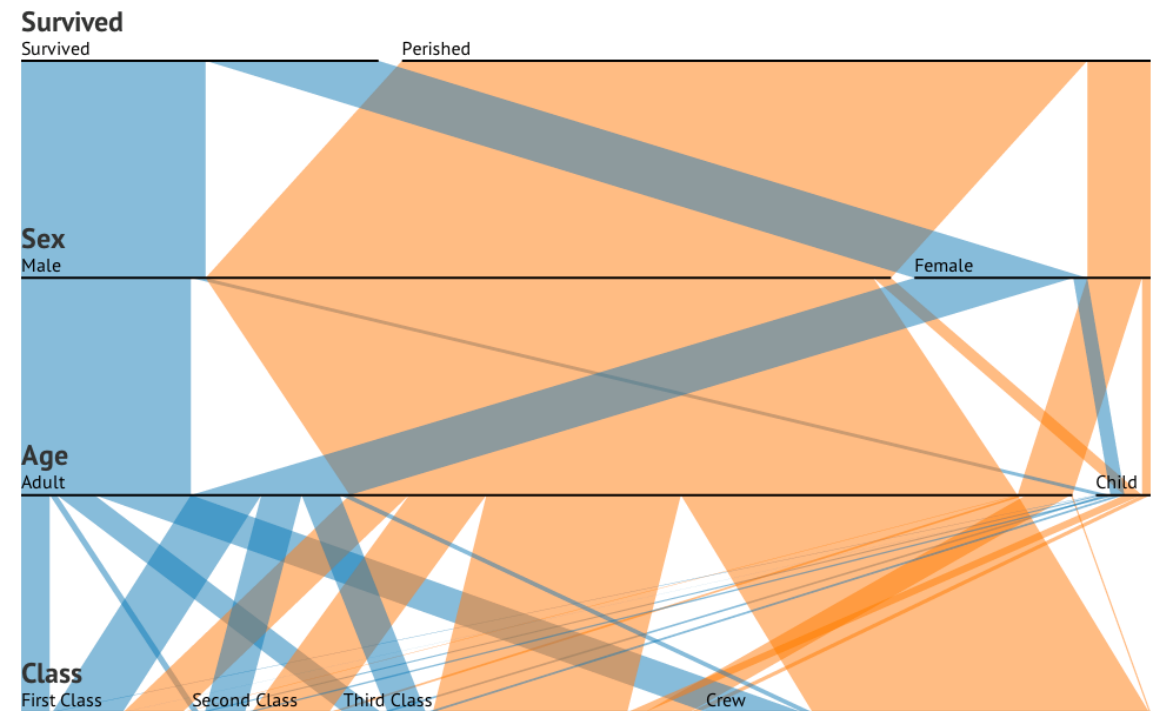
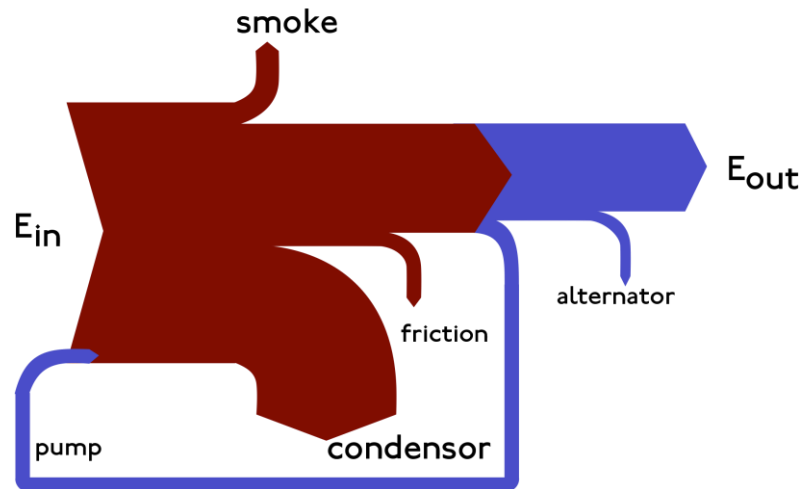
Data			
	Variable A	Variable B	Variable C
Item 1	50	100	2.0
Item 2	30	115	0.5

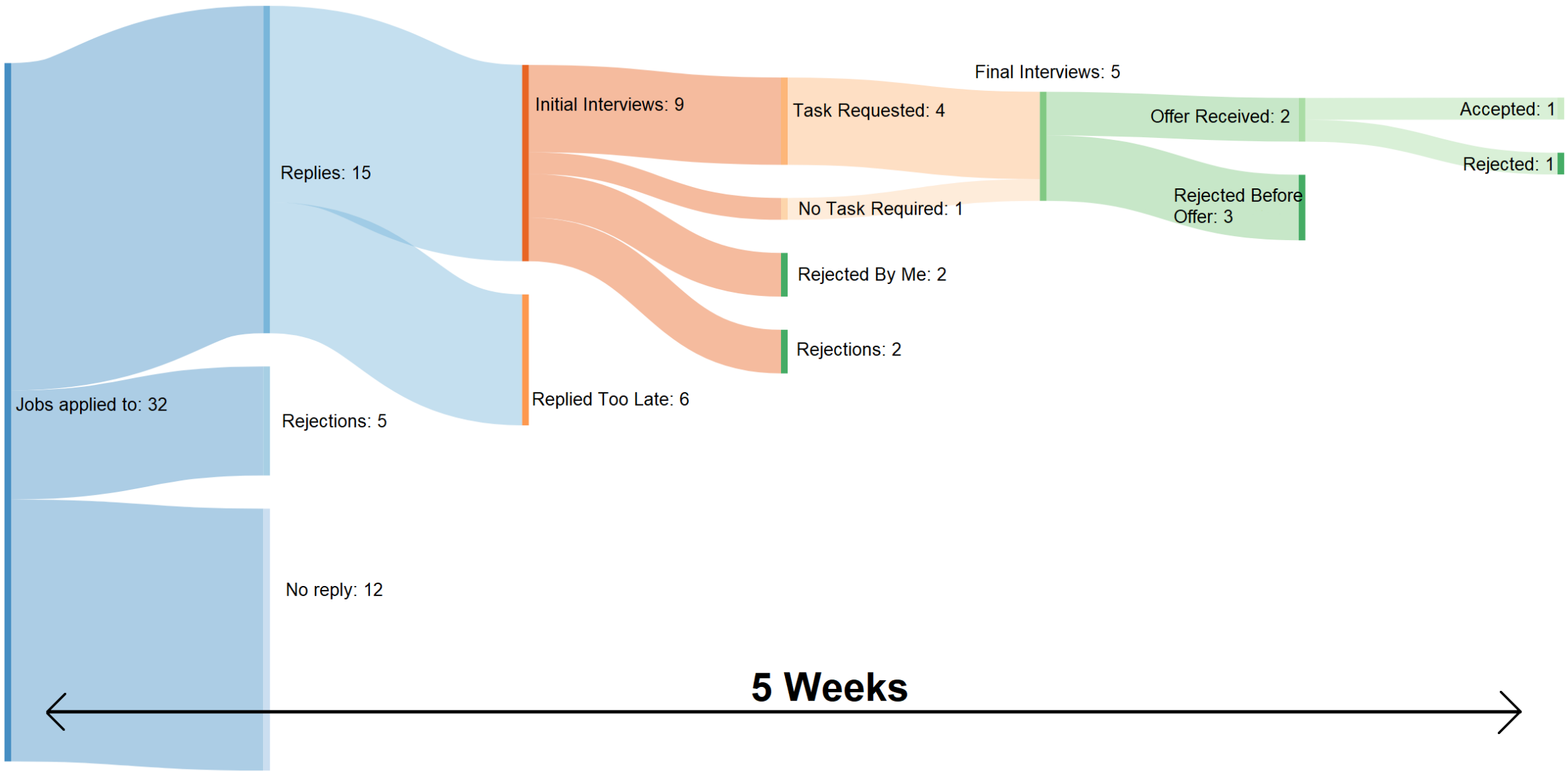


Parallel sets



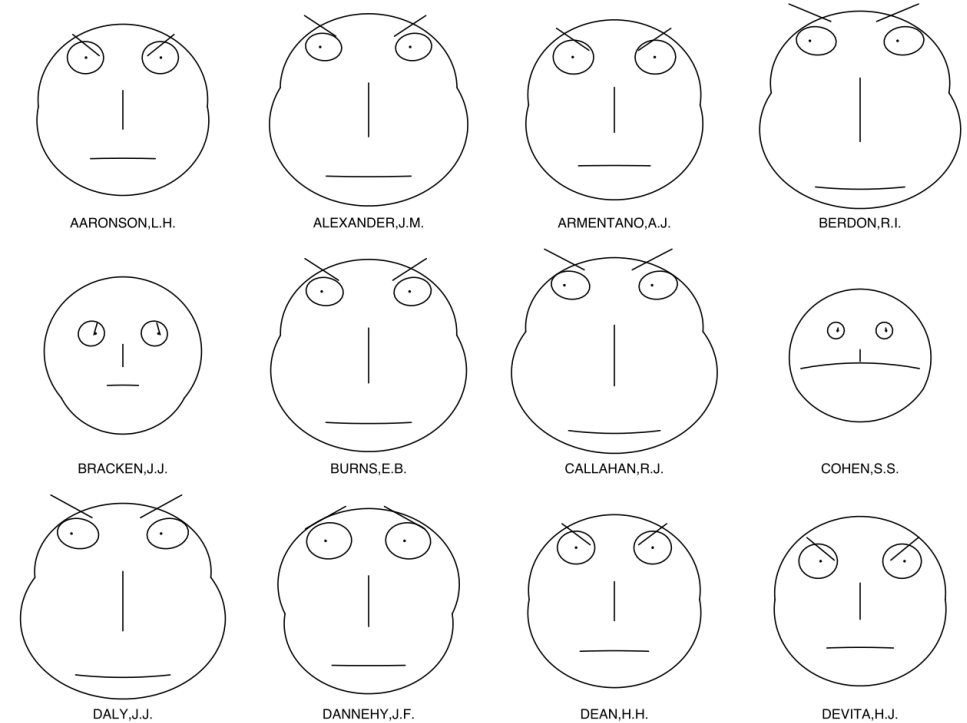
- pro kategorická data, Robert Kosara
- <https://eagereyes.org/parallel-sets>
- alluvial diagrams (RAW)
- *sankey*





Chernoffovy tváře

- 1973, Herman Chernoff
- ikonové grafy
- mozek: *fusiform face area*
- vícerozměrná data
- velmi špatně se čtou
- [v praxi](#) se používají minimálně
- existují pokusy je vylepšit
- [nemusí to být jen obličej](#)



Chernoffovy tváře

- [empathic visualization algorithm](#) (EVA)
- lidé velmi dobře čtou tváře
- finanční situace 4 firem
- do které investujete?



Chernoffovy tváře

„Despite the best effort of academics to find a way to make the faces ‘work’, the usual response to Chernoff faces from the cartographic community is a combination of fascination and loathing.“ – [Martin Elmer](#)

- Předpozornostní povaha tváře?
- [co je s tím za problémy?](#)
- naštvaný může být dobrý v jistých ohledech
- pareidolie – rozpoznávání tváří v mracích atp.

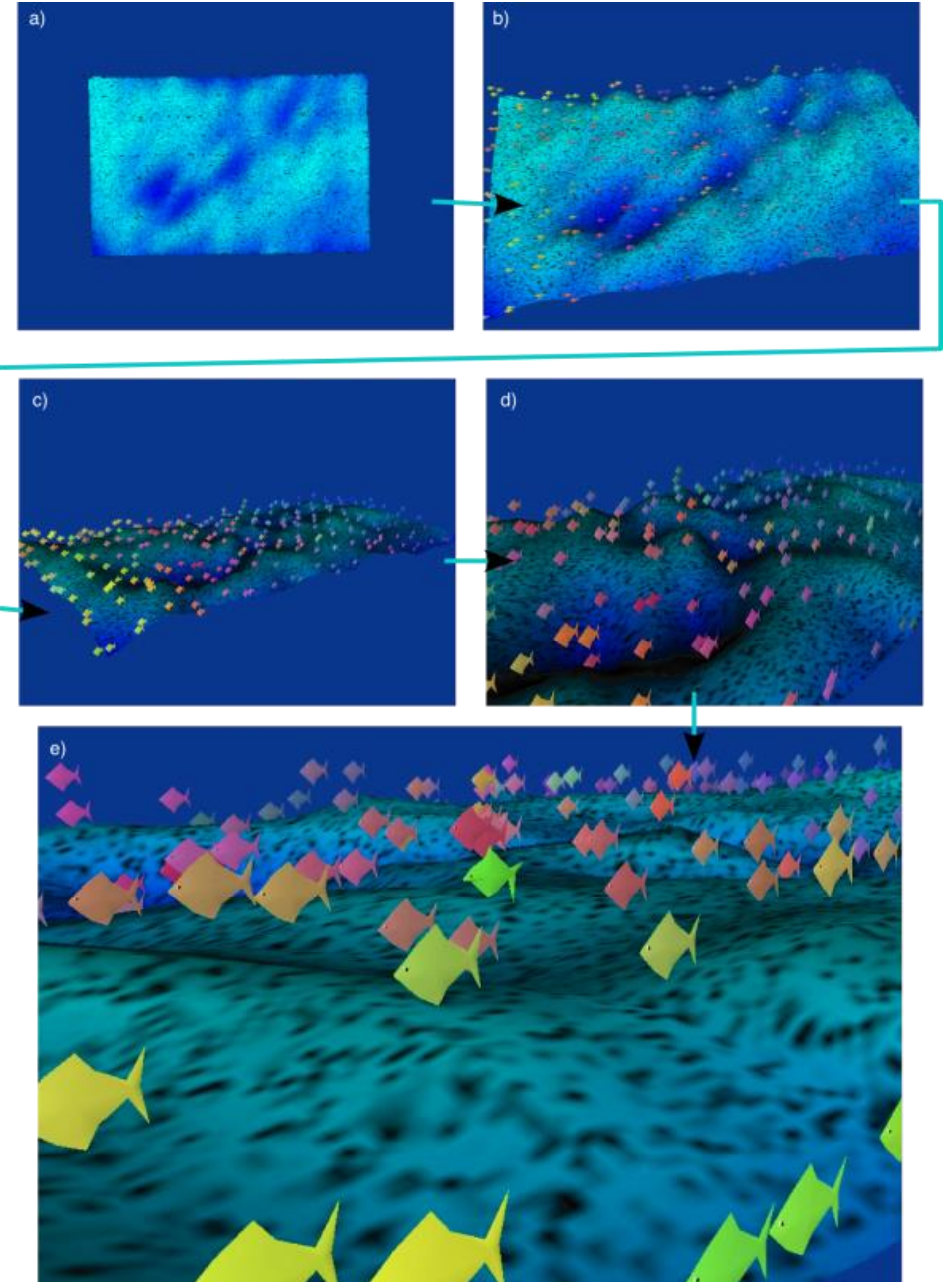
Chernoffovy tváře

- enviromentální metaforý
- Leaf Glyphs



Metaforické 3D glyfy

- přidejme ještě jednu dimenzi
- proč ne, že
- [REEFSOME](#)



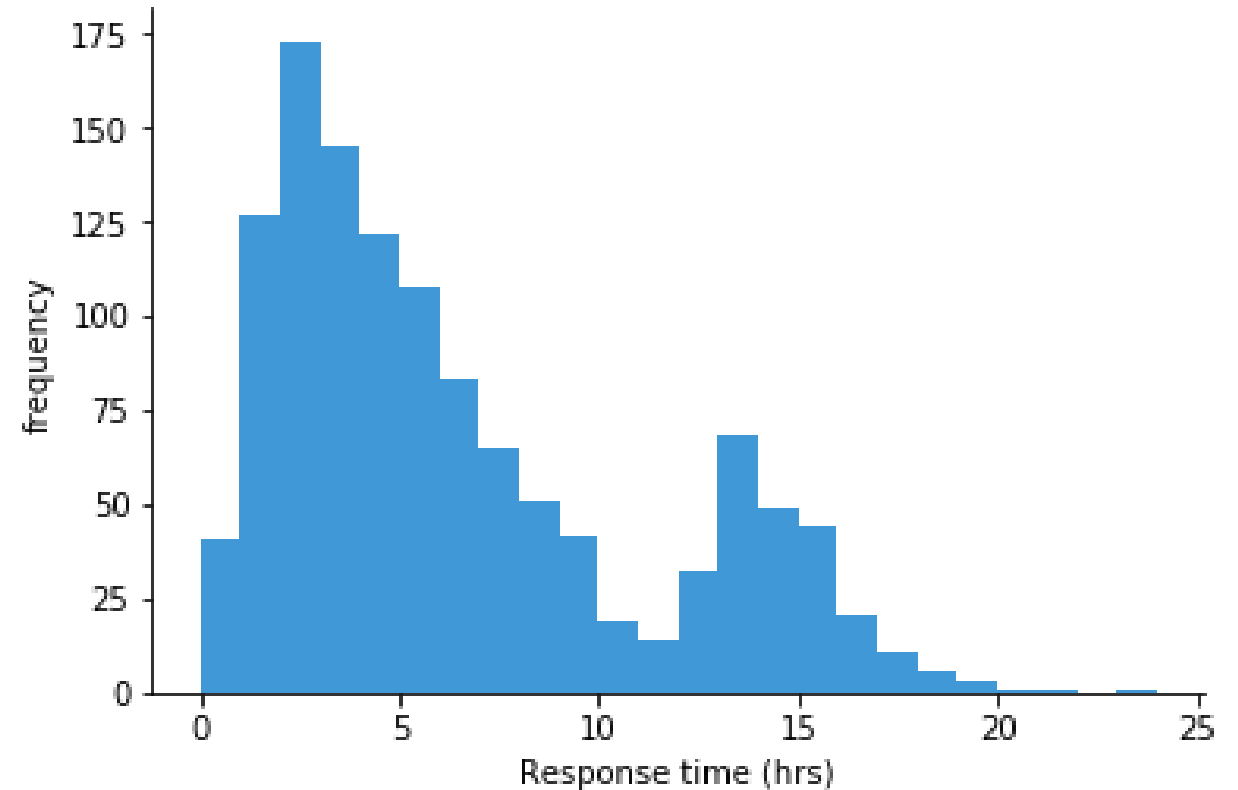
Metaforické 3D glyfy

- [Kanohi 3D data exploration](http://kanohi.com/3d-data-exploration/)
- <http://data-exploration.ch/>



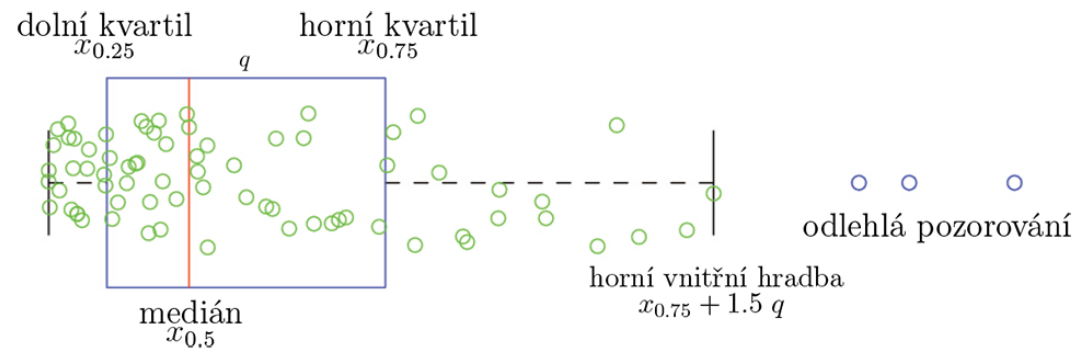
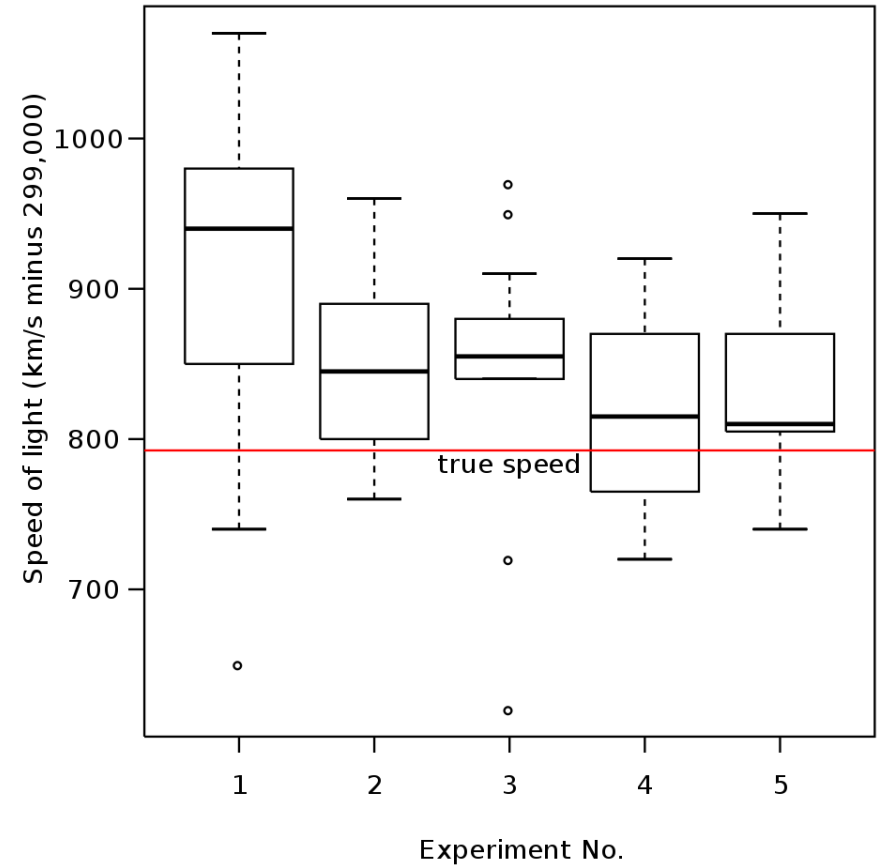
Histogram

- vypadá jako sloupcový graf
- vizualizuje distribuci dat
- na ose x je interval
- na ose y četnost
- v čem je rozdíl?

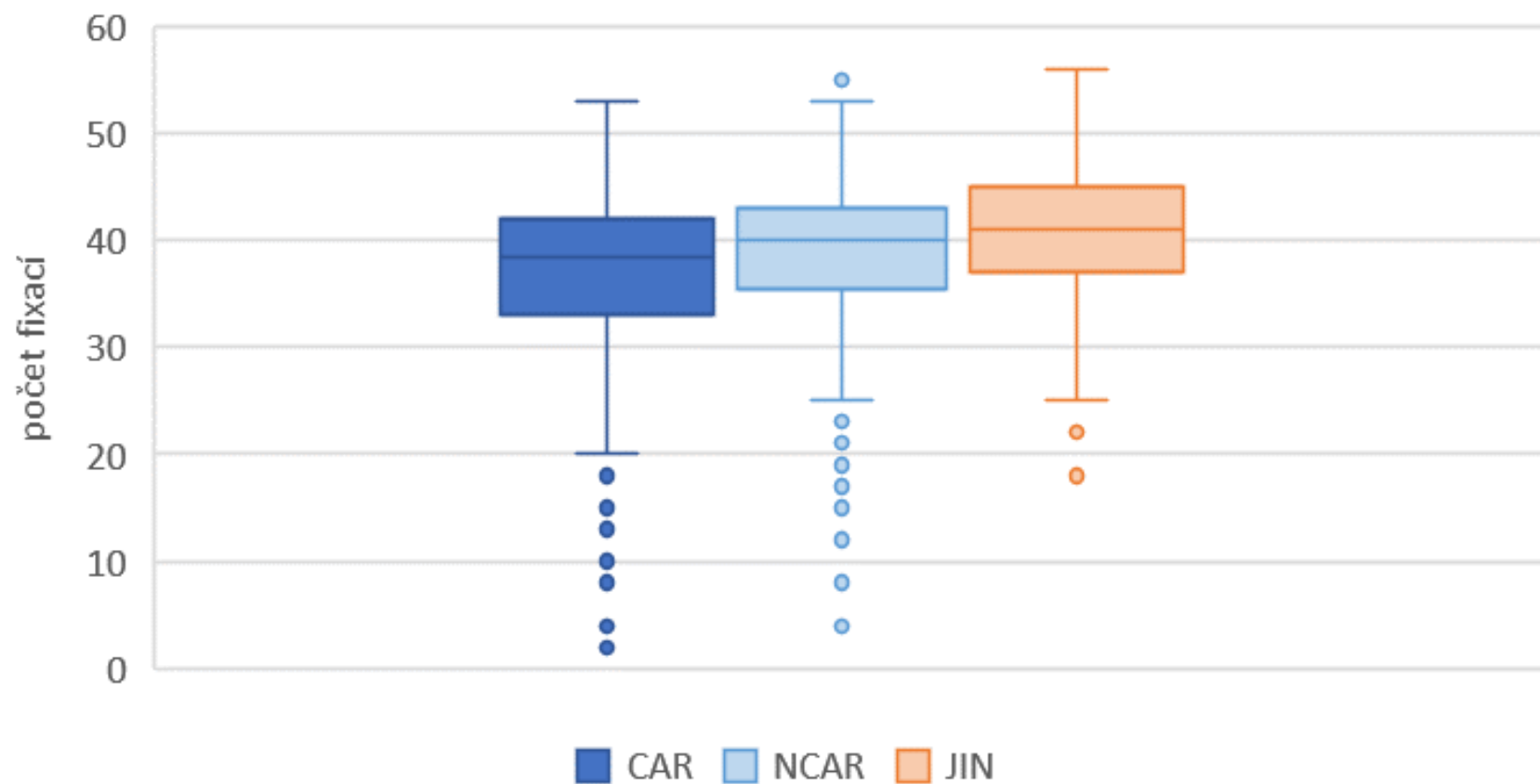


Boxplot

- krabicový graf
- *box-and-whisker plot*
- zkoumání více sad dat
- např. několik experimentů

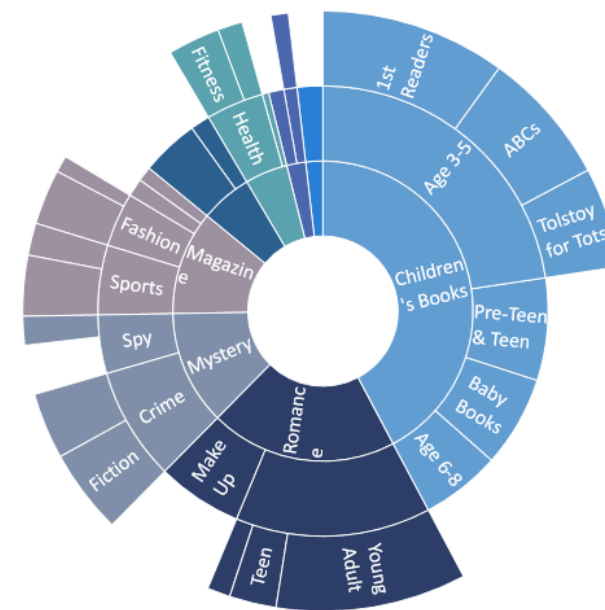
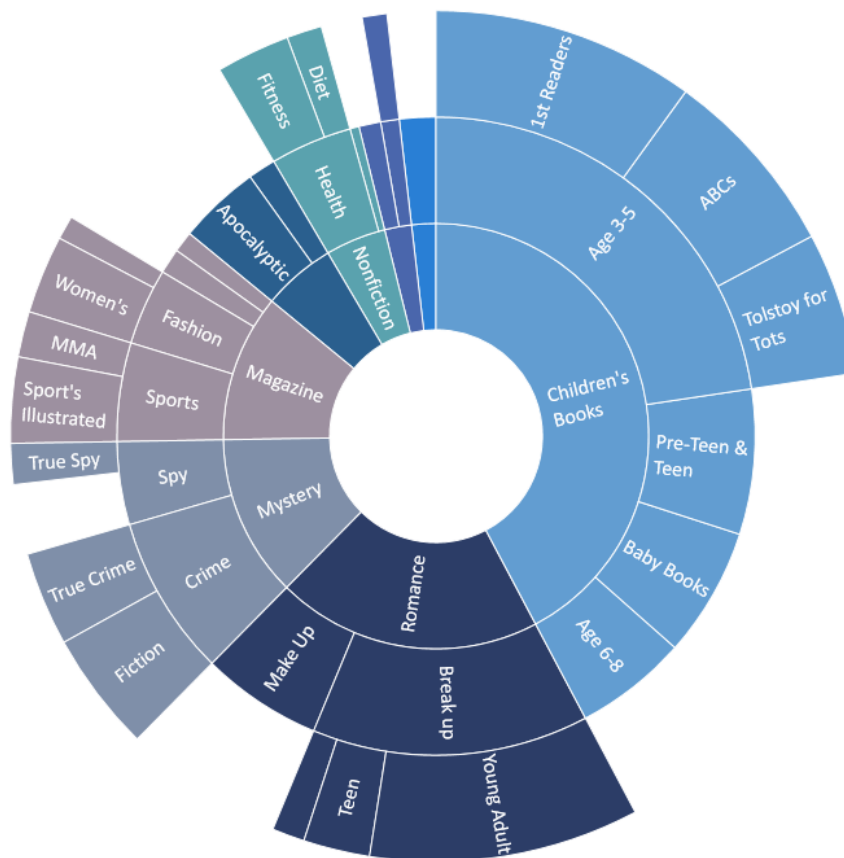


Mapy - Počet fixací



Hierarchická data

- treemapy
- sunburst
- dendrogram



Treemapy

- Ben Shneiderman
- hierarchická data
- např. rozpočet
- využití úložiště
- souborová struktura
- <https://windirstat.net/>



Kam dál?

- mapové formy - příště
- <https://datavizcatalogue.com/>

