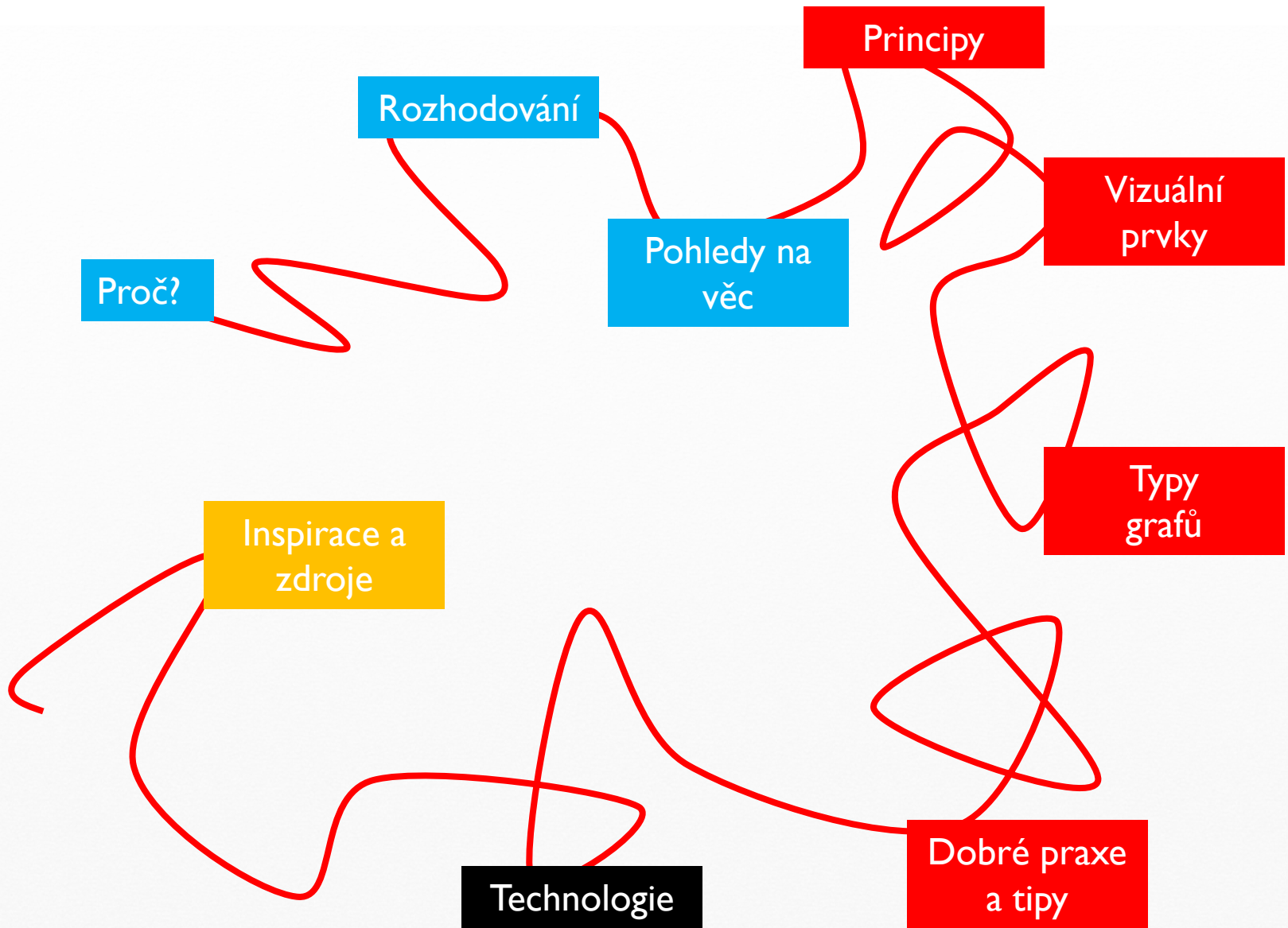


Evaluátorův průvodce vizualizací dat

Petr Bouchal

Evaluace ve světle nových příležitostí, 9. 6. 2015



Proč vizualizovat



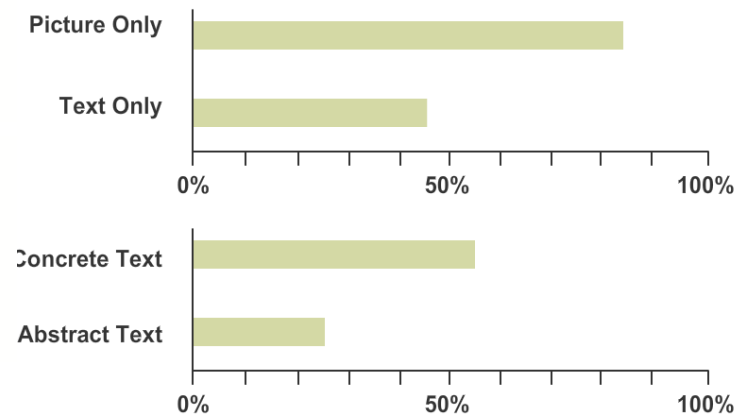
- Omezení média: vizualizace nabízí vyšší hustotu informací, možnost prozkoumat data
- Informační overload: vizualizace pomůže zdůraznit to podstatné v přebytku informací

Proč vizualizovat

a 85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898

b 85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898

pictures vs. text



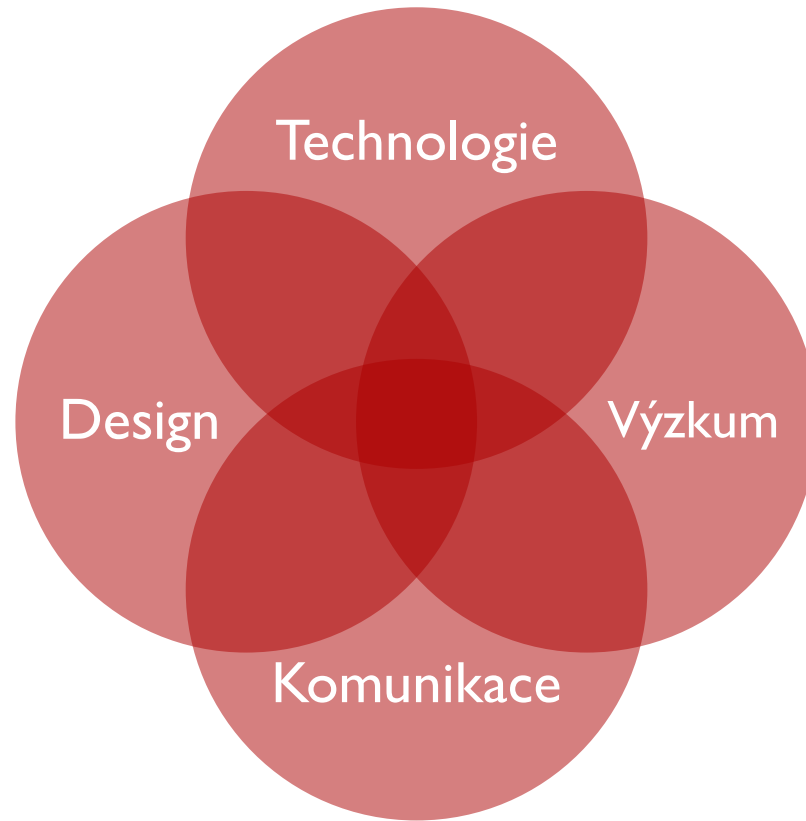
- Čtenář informace rychleji zachytí
- Lidský mozek si vizuální informace lépe pamatuje

Zdroj: Colin Ware. *Information Visualization: perception for design*. 2004
John Medina, *Brain Rules*. 2015 brainrules.net/vision

Udělejte zásadní rozhodnutí

- Proč?!
- Chci přesvědčit? Ukázat informaci? Vtáhnout diváka do dat?
- Co chci ukázat? (Data nemluví sama za sebe!)
- Co chci, aby si divák zapamatoval?
- Záleží na přesnosti?
- Kolik má divák času?
- Jaký je kontext a kolik z něj musí divák znát?

Zapojte různé pohledy a dovednosti



Zdroj: Jeff Knezovich, Visualising data: both a science and an art, On think tanks
onthinktanks.org/2015/05/01/visualising-data-both-a-science-and-an-art/

Držte se zásad dobré vizualizace dat

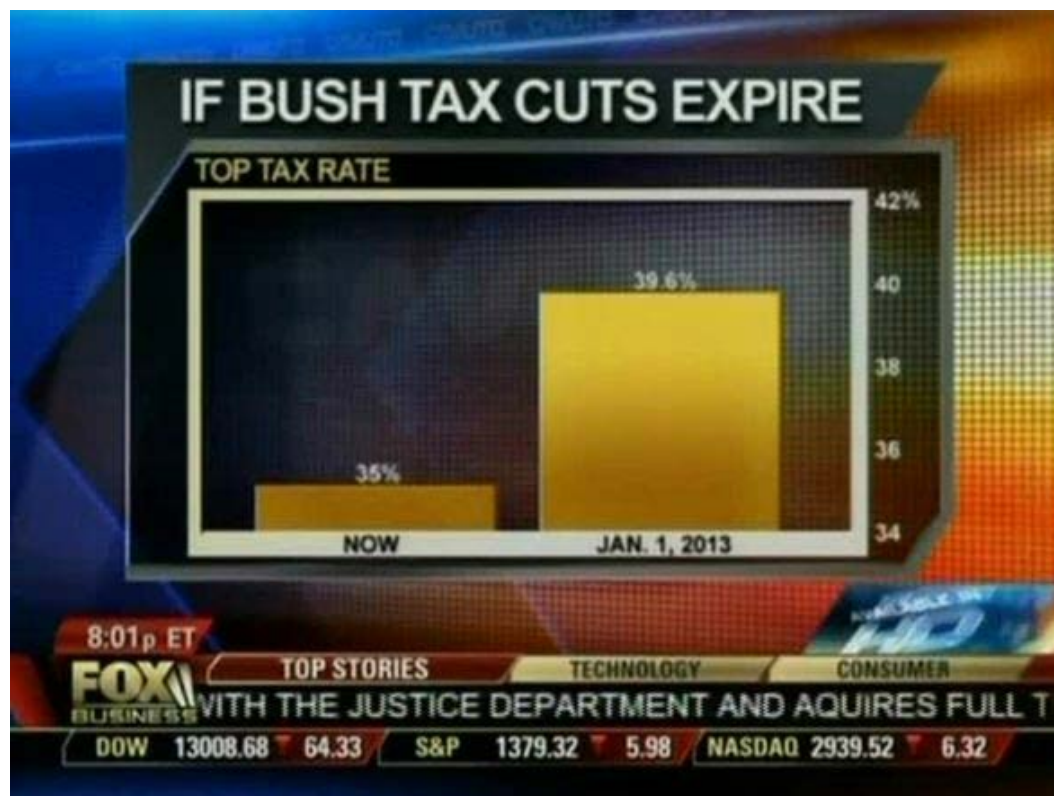
Integrita

Důraz na data

Čitelnost

Informativnost

Iterace a zlepšování



Zdroj: Media Matters, Dishonest Fox Chart: Bush Tax Cut Edition, 2012 mediamatters.org/blog/2012/07/31/dishonest-fox-chart-bush-tax-cut-edition/189046

Držte se zásad dobré vizualizace dat

Integrita

Důraz na data

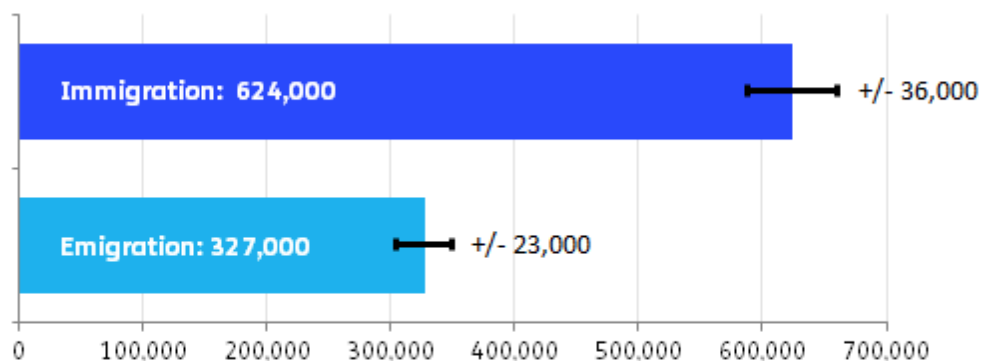
Čitelnost

Informativnost

Iterace a zlepšování

Uncertain migration

Inflow and outflow of all citizenships, year ending September 2014



Source: ONS migration statistics quarterly report, February 2015

Full Fact

Zdroj: Full Fact, Why don't we know how many migrants are entering and leaving the UK? fullfact.org/immigration/count_people_in_and_out_uk-37635

Držte se zásad dobré vizualizace dat

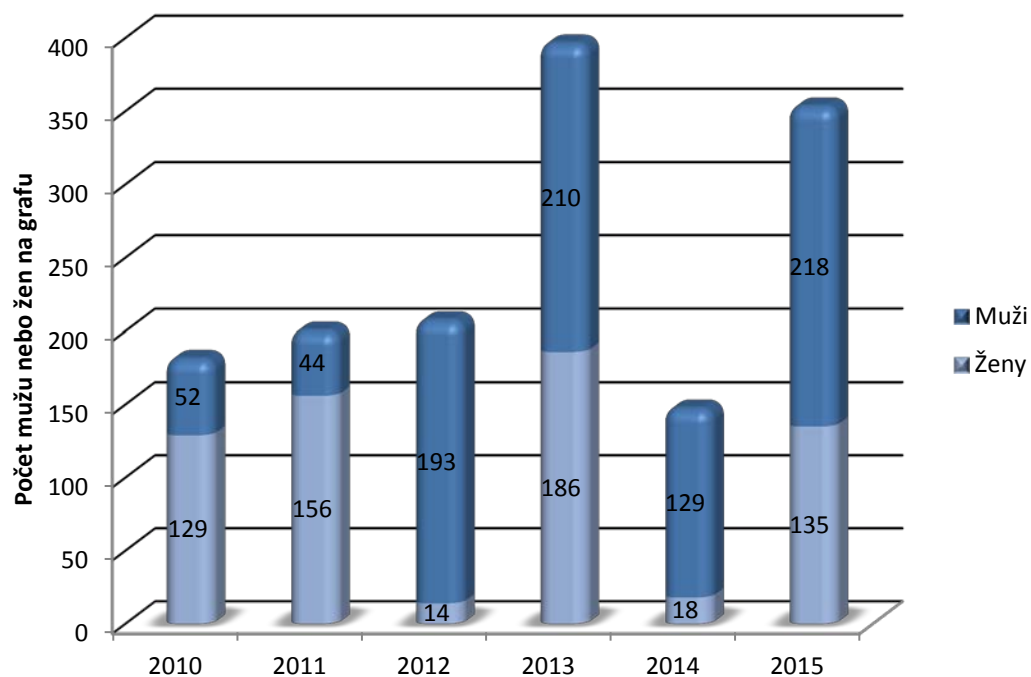
Integrita

Důraz na data

Informativnost

Čitelnost

Iterace a zlepšování



Zdroj: generátor náhodných čísel MS Excel, červen 2015

Držte se zásad dobré vizualizace dat

Integrita

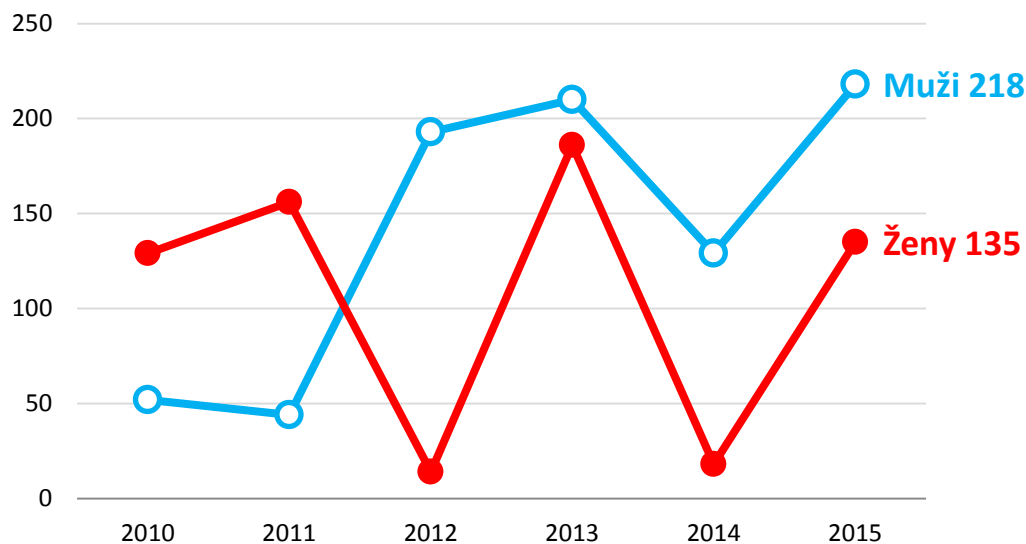
Důraz na data

Informativnost

Čitelnost

Iterace a zlepšování

Od roku 2012 bylo na grafu více mužů
počty mužů a žen na grafu, 2010-15



Zdroj: generátor náhodných čísel MS Excel, červen 2015

Poměr data : inkoust

Držte se zásad dobré vizualizace dat

Integrita

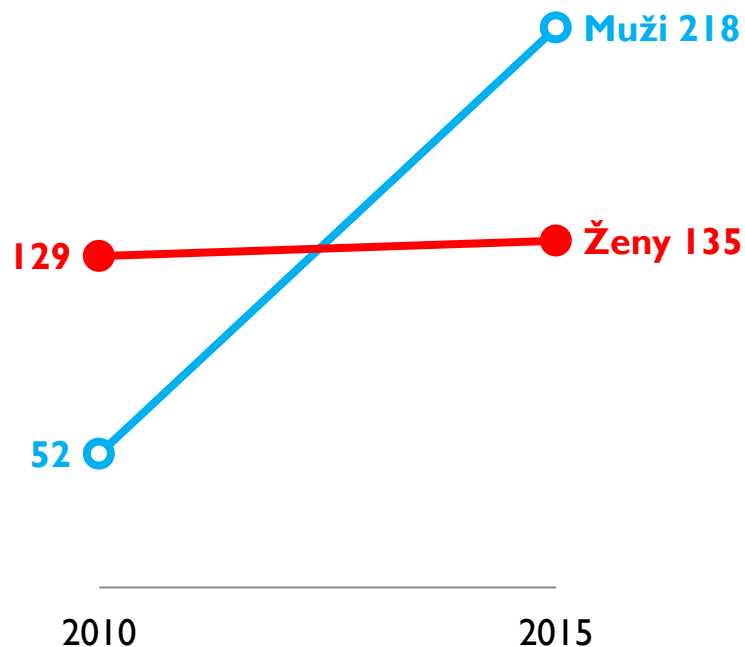
Důraz na data

Informativnost

Čitelnost

Iterace a zlepšování

Muži od roku 2010 předstihli ženy
počty mužů a žen na grafu, 2010-15



Zdroj: generátor náhodných čísel MS Excel,
červen 2015

Držte se zásad dobré vizualizace dat

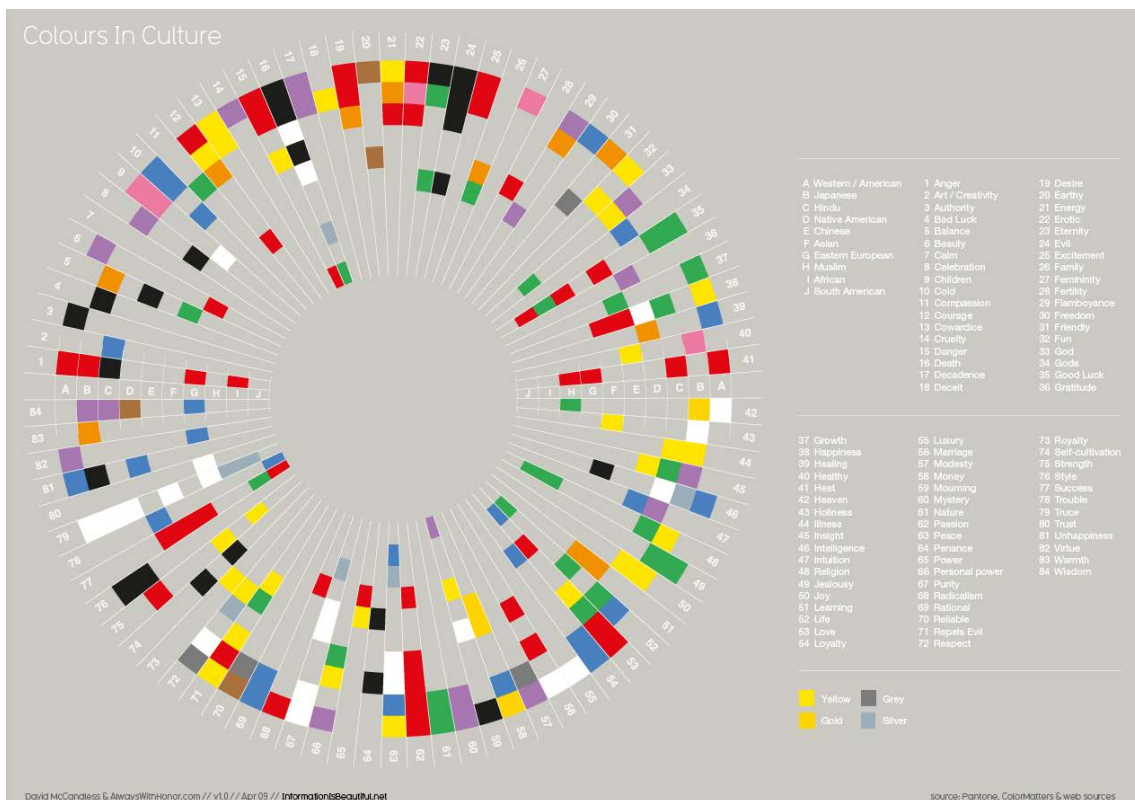
Integrita

Důraz na data

Informativnost

Čitelnost

Iterace a zlepšování



Zdroj: David McCandless, Information is Beautiful, Colours in Culture
informationisbeautiful.net/visualizations/colours-in-cultures/

Držte se zásad dobré vizualizace dat

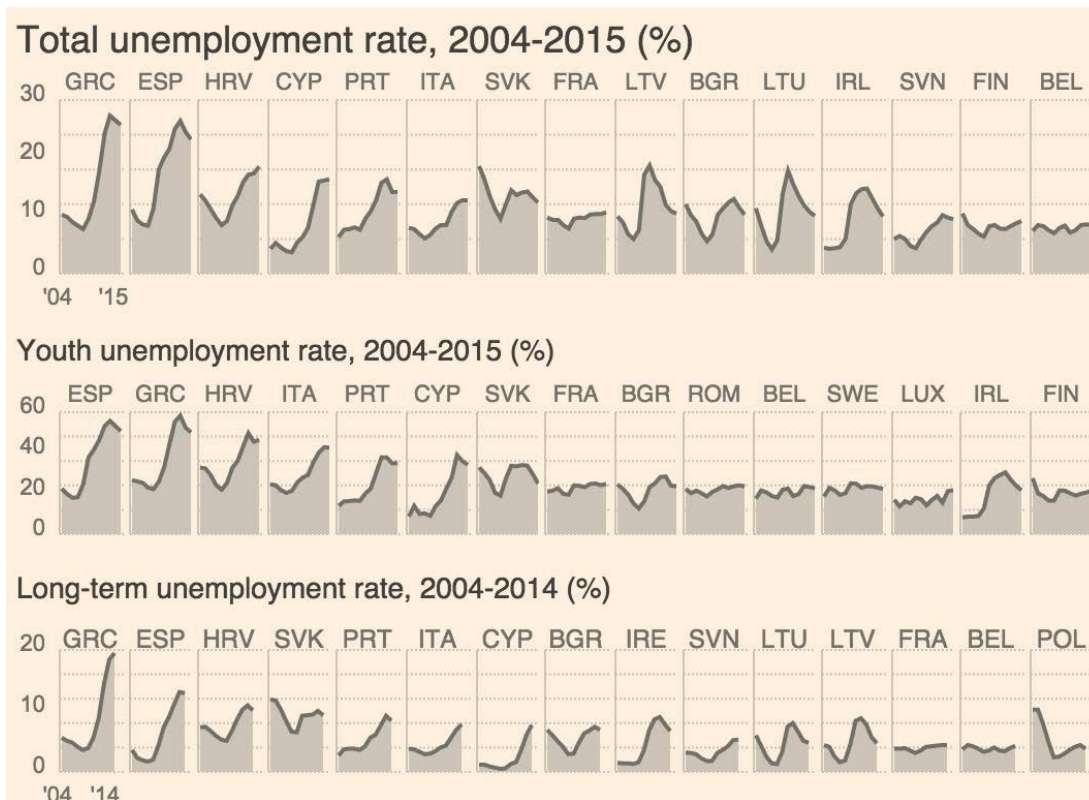
Integrita

Důraz na data

Informativnost

Čitelnost

Iterace a zlepšování



Datová hustota

Zdroj: Financial Times, Data Blog - EU Unemployment Tracker, 2015
blogs.ft.com/ftdata/2015/04/17/eu-unemployment-tracker/

Držte se zásad dobré vizualizace dat

"I SPEND A LOT OF TIME ON THIS TASK.
I SHOULD WRITE A PROGRAM AUTOMATING IT!"

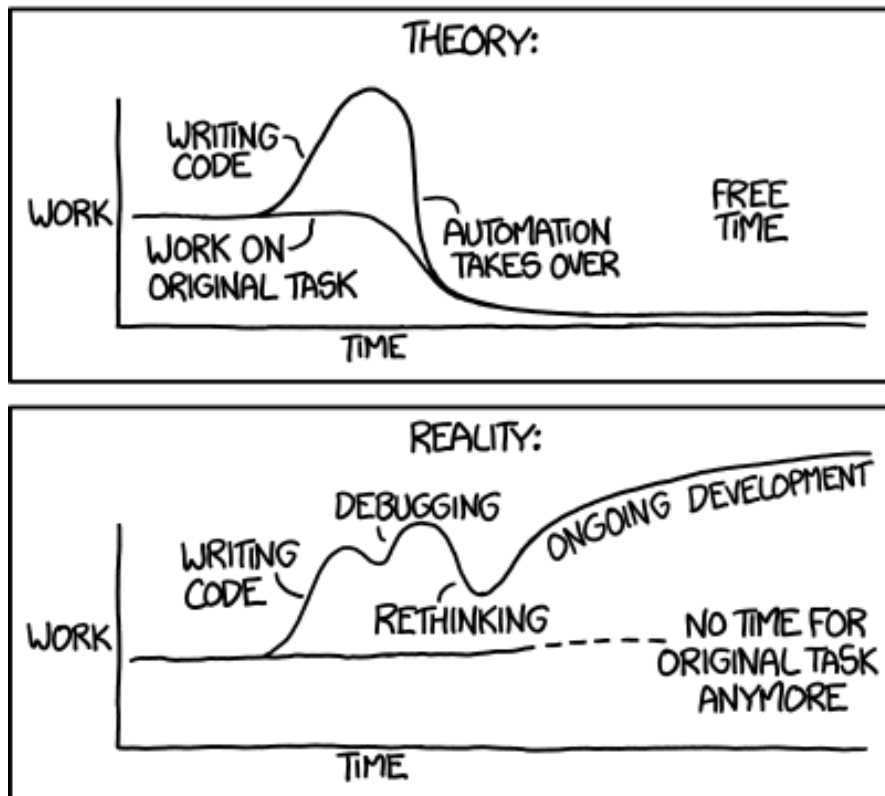
Integrita

Důraz na data

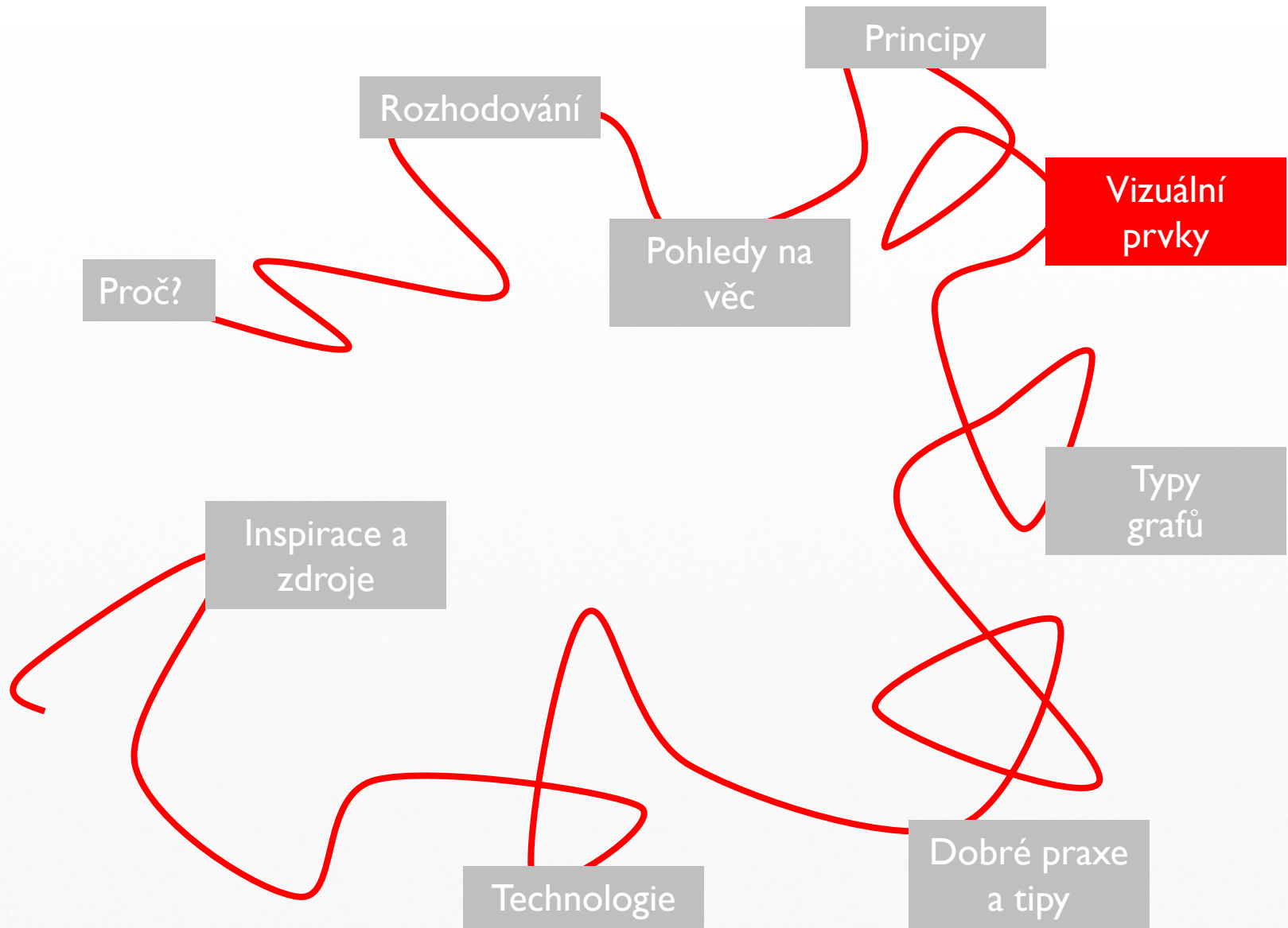
Informativnost

Čitelnost

Iterace a zlepšování



Zdroj: xkcd, xkcd.com/1319/

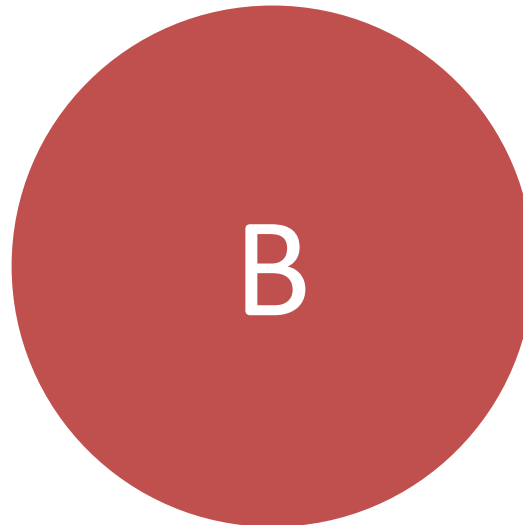


Vizualizaci vystavějte z vhodných prvků

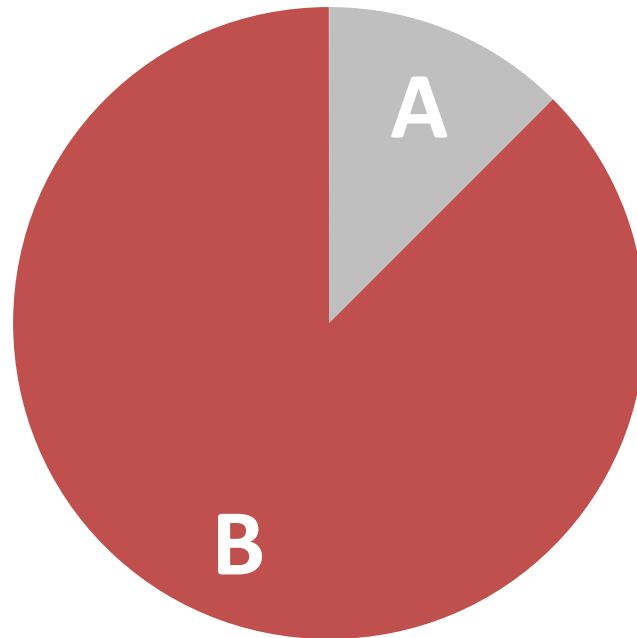
- **Vizuální prvky**, které nejlépe znázorní daná data
- **Typ grafu** je potřeba vybrat podle účelu vizualizace
- **Doplňkové prvky**, které pomáhají čitelnosti a neruší
- **Kontext** dokumentu či sdělení, v němž vizualizace funguje

Vizuální prvky ovlivňují sdílnost grafu

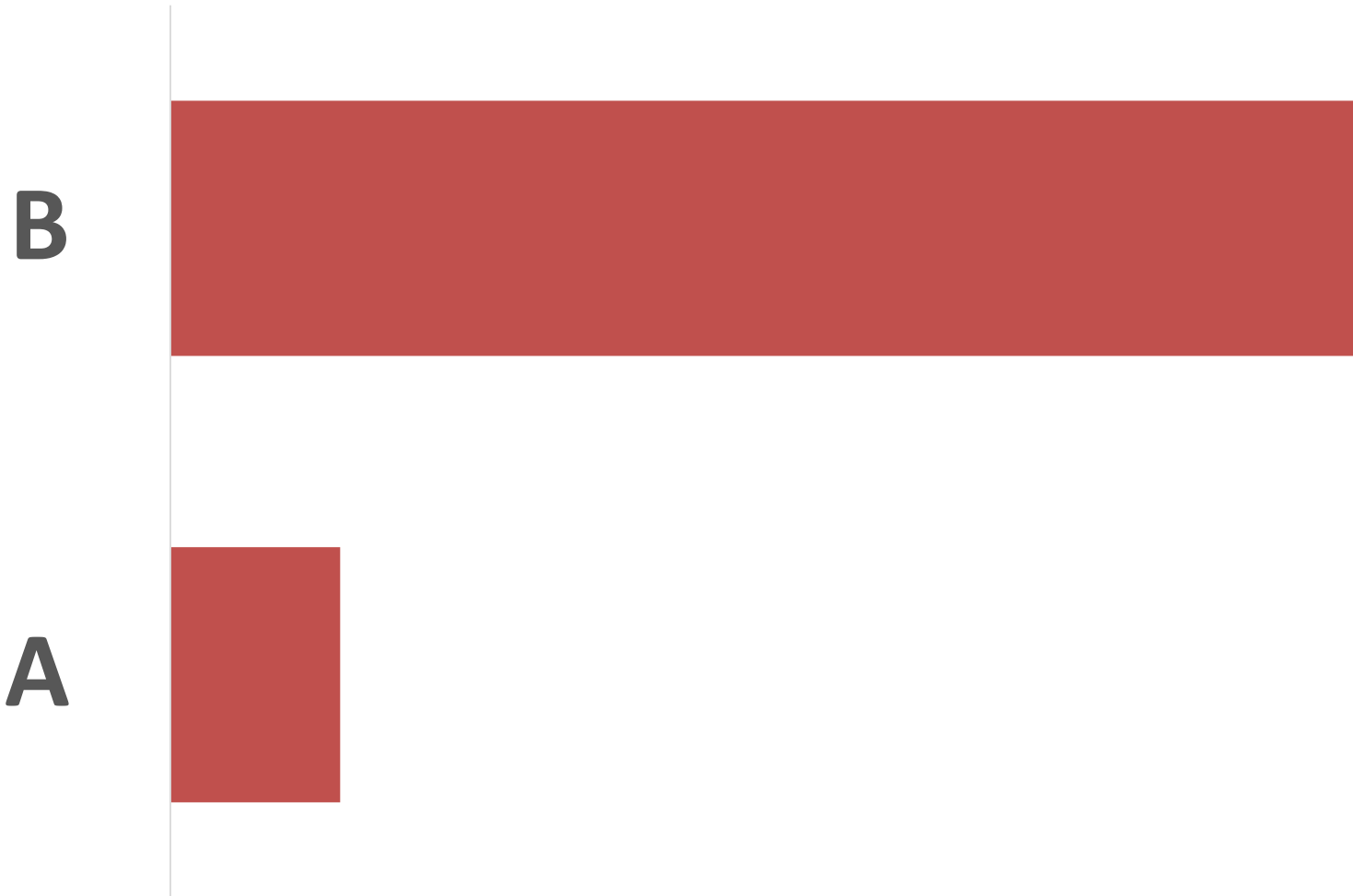
Kolikrát je B větší než A?



Kolikrát je B větší než A?



Kolikrát je B větší než A?



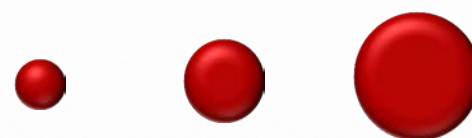
Zdroj experimentu: Jon Schwabish

Vizuální prvky se liší mj. přesností

Poloha



Objem



Délka



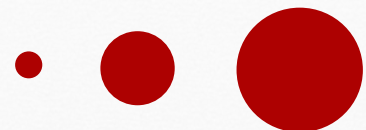
Svtost



Úhel



Plocha



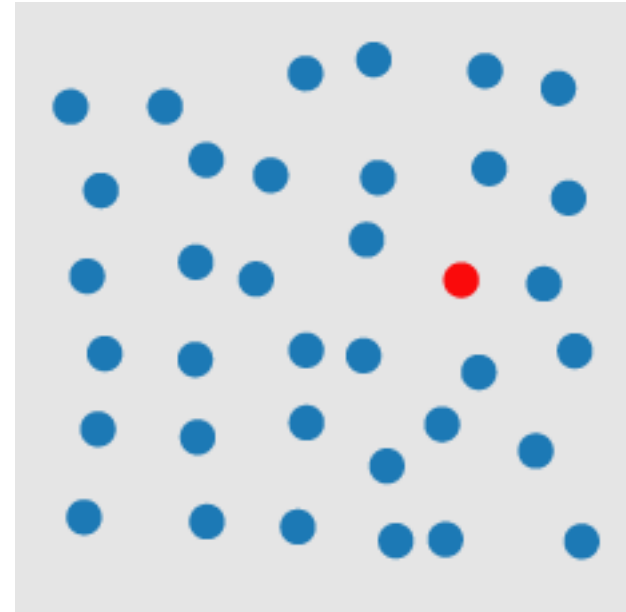
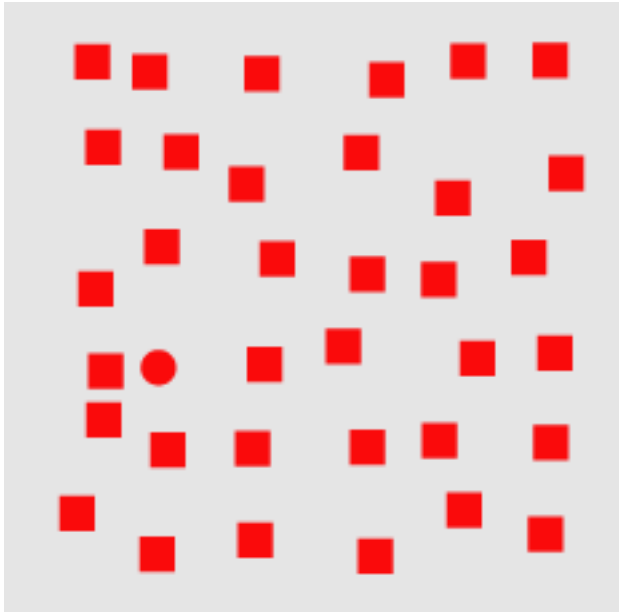
Tvar



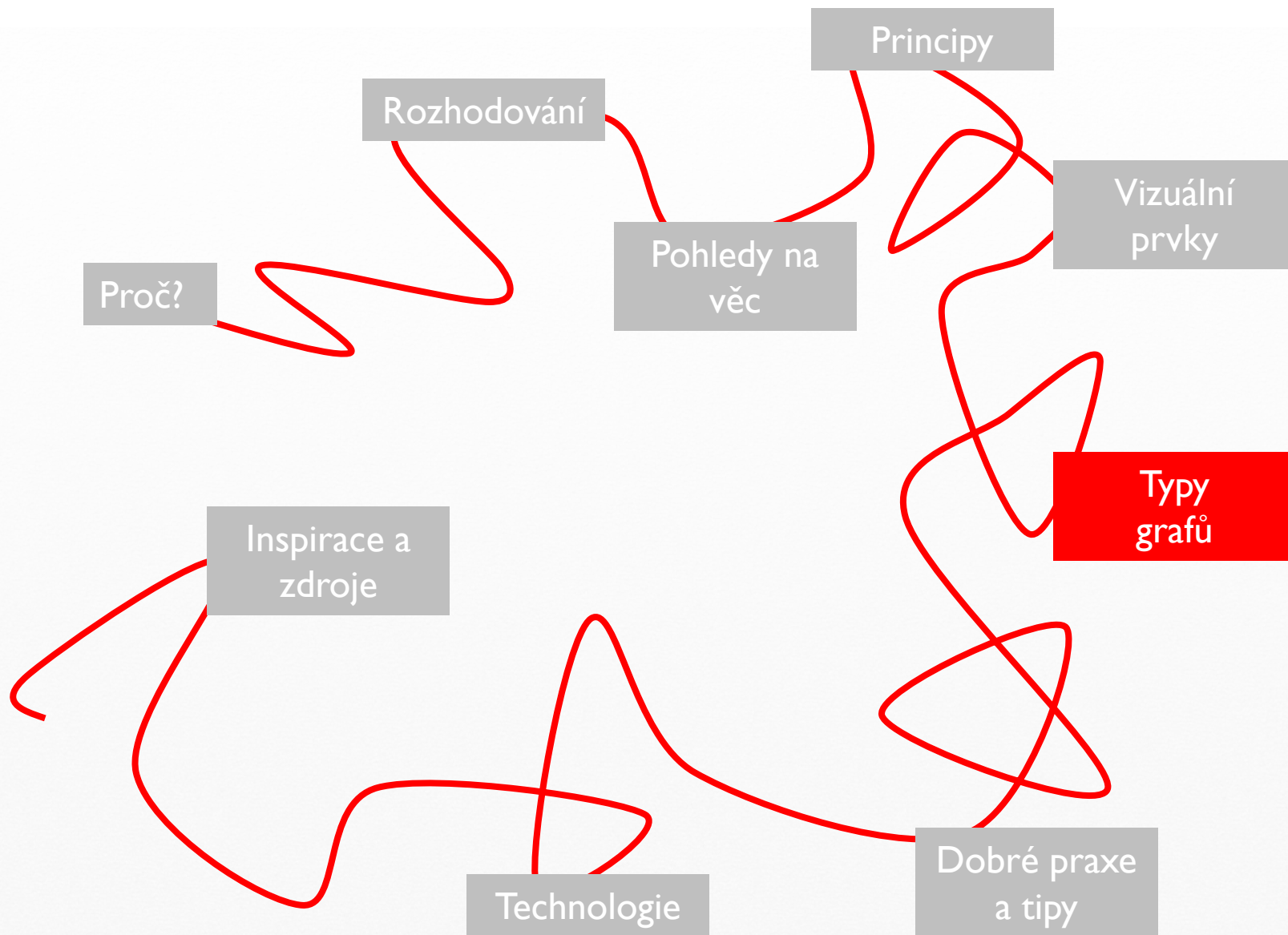
Přesnost vnímání

Zdroj: William S. Cleveland, Robert McGill, „Graphical Perception and Graphical Methods for Analyzing Scientific Data“, *Science* 229: 828-833. doi.org/10.1126/science.229.4716.828

...nebo výrazností



Zdroj: Christopher Healey, „Perception in Visualization“, Department of Computer Science, North Carolina State University, csc.ncsu.edu/faculty/healey/PP/index.html



Typ grafu zvolte podle účelu a typu dat

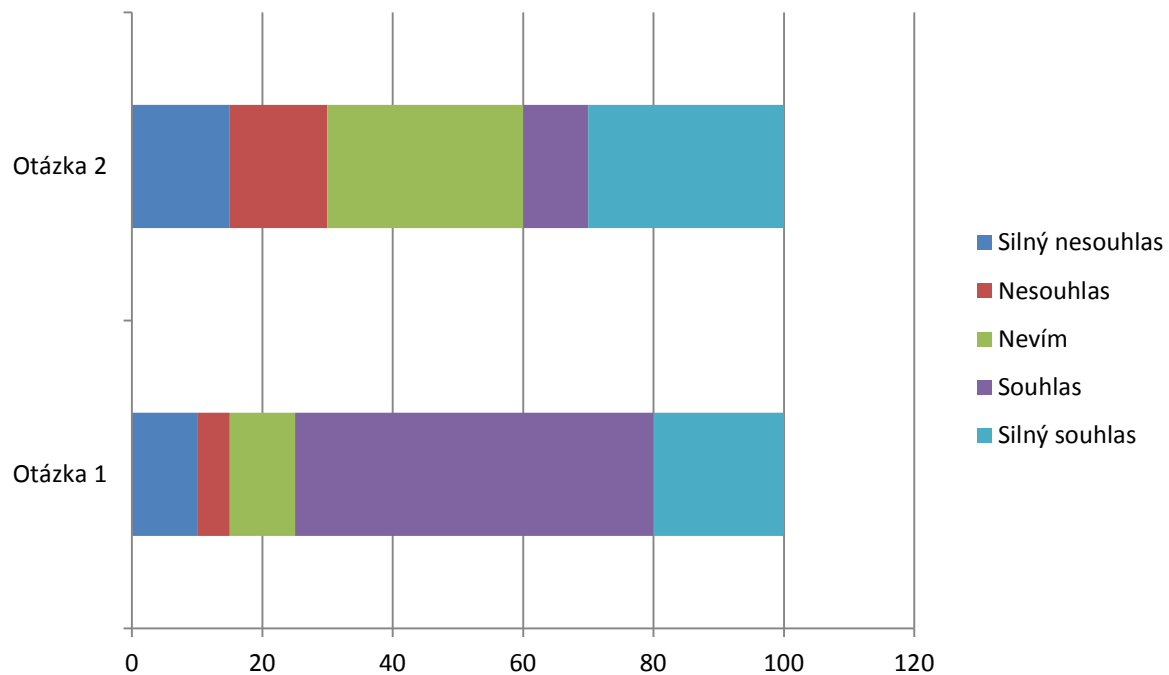
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných čísel MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

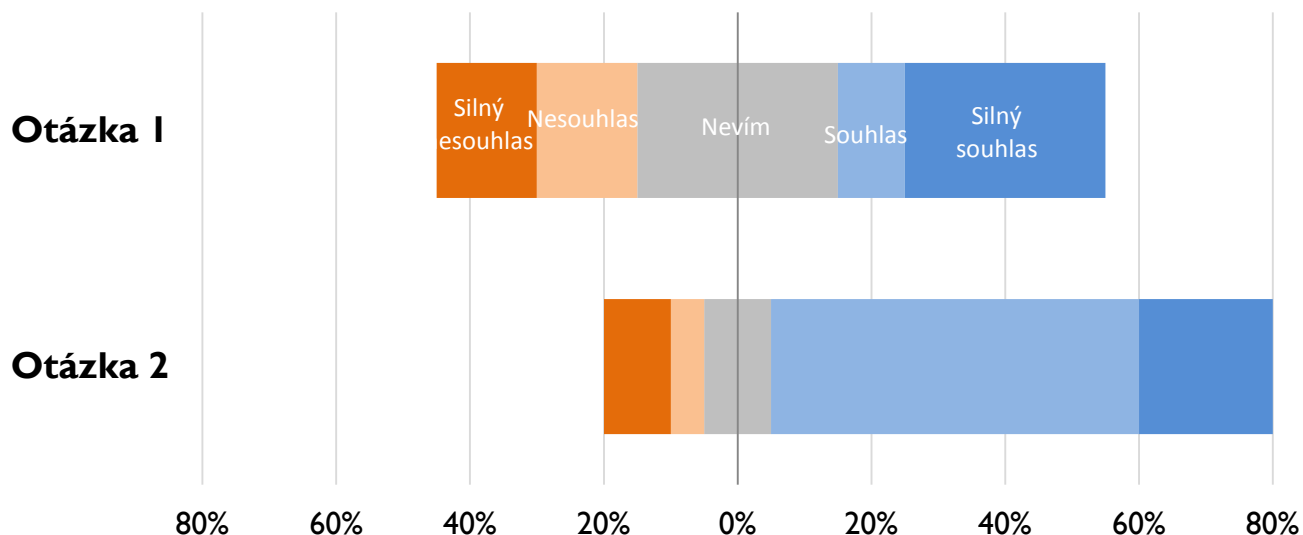
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných čísel MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

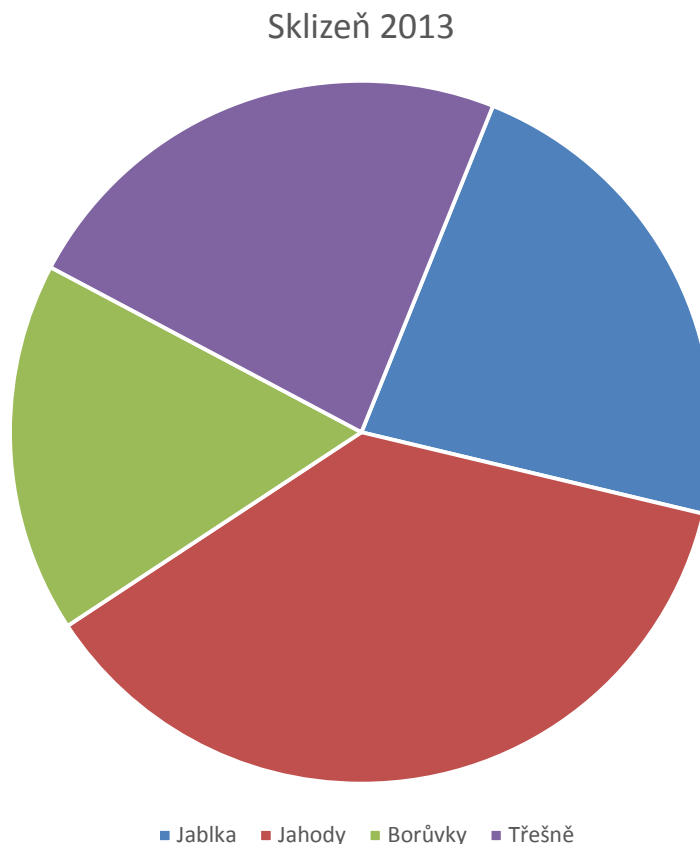
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Sklizeň 2013

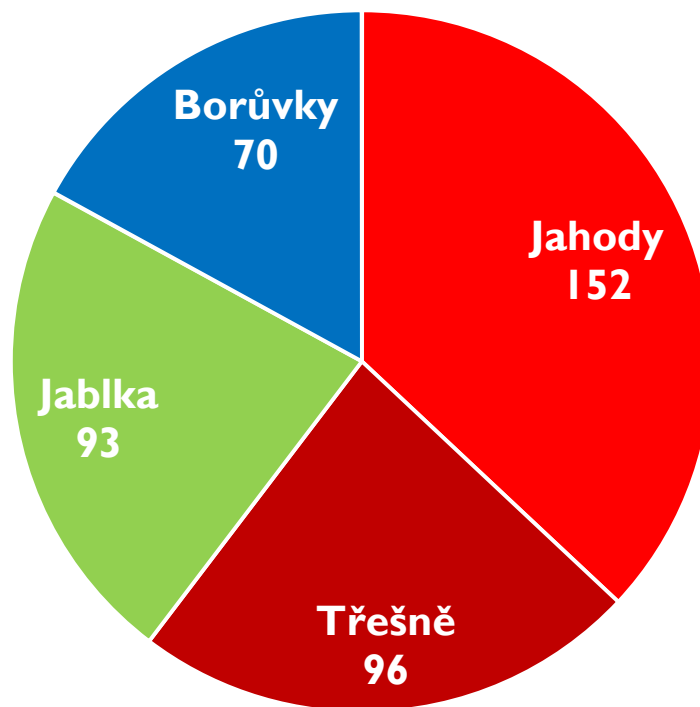
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Jahody v roce 2013 tvořily přes 1/3 sklizně

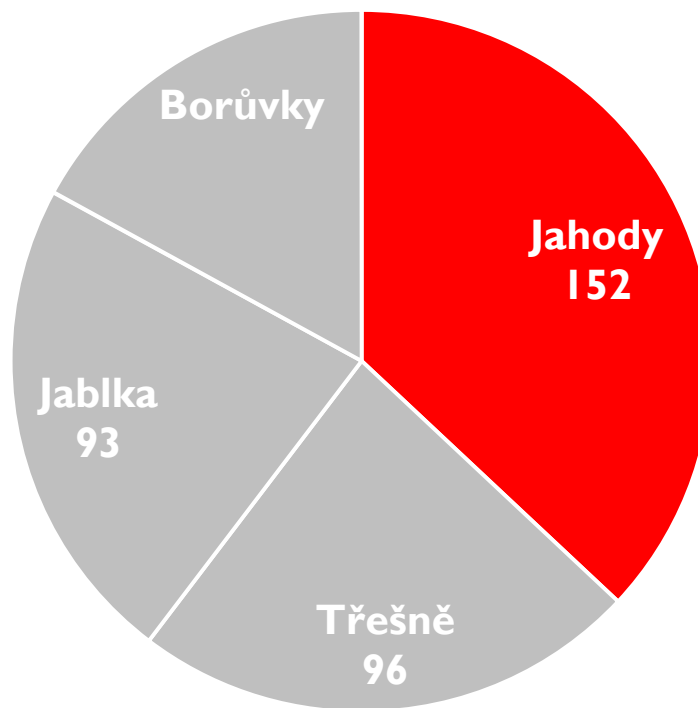
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

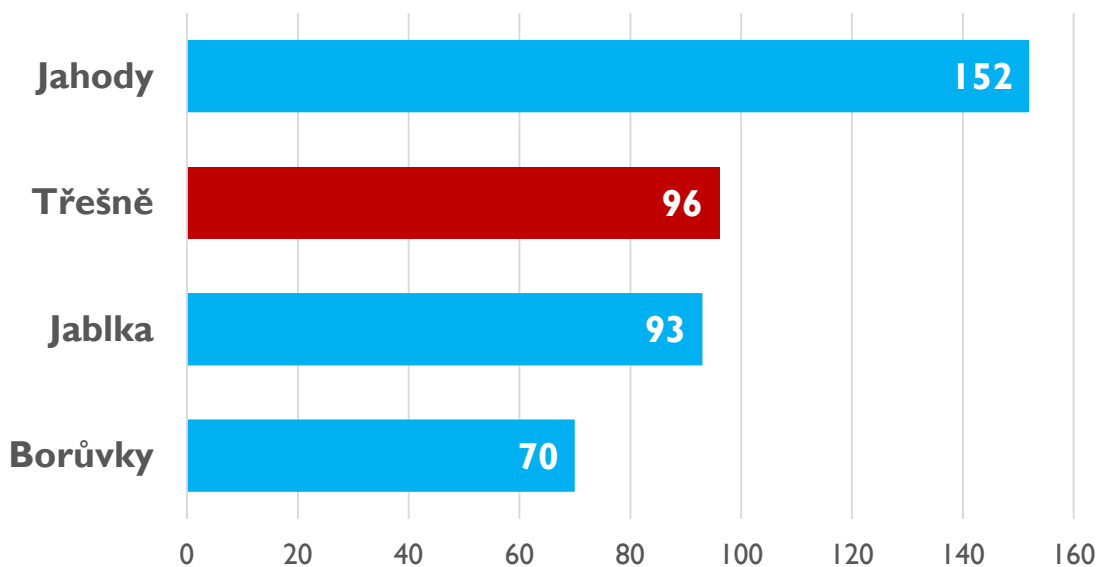
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

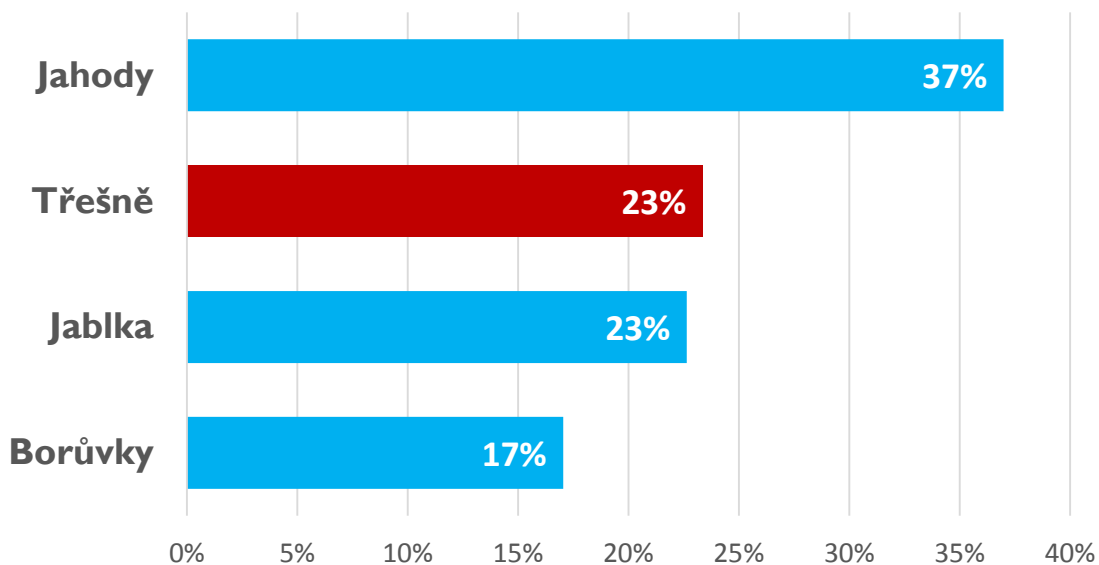
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

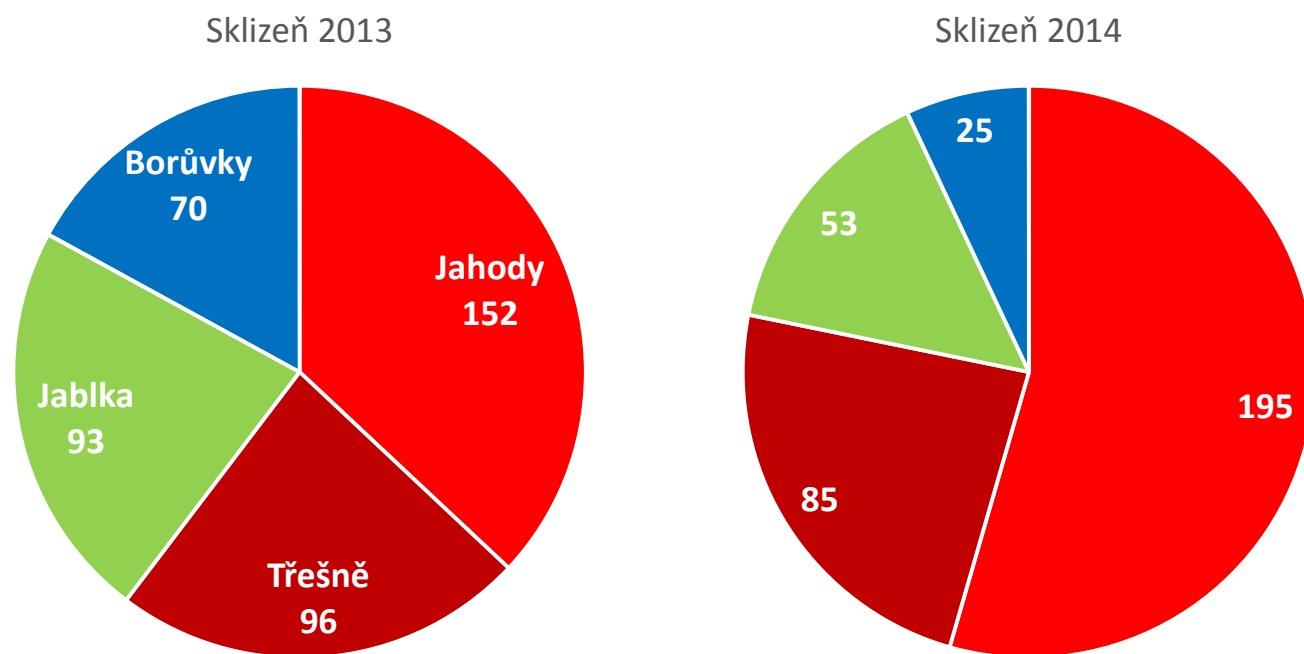
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

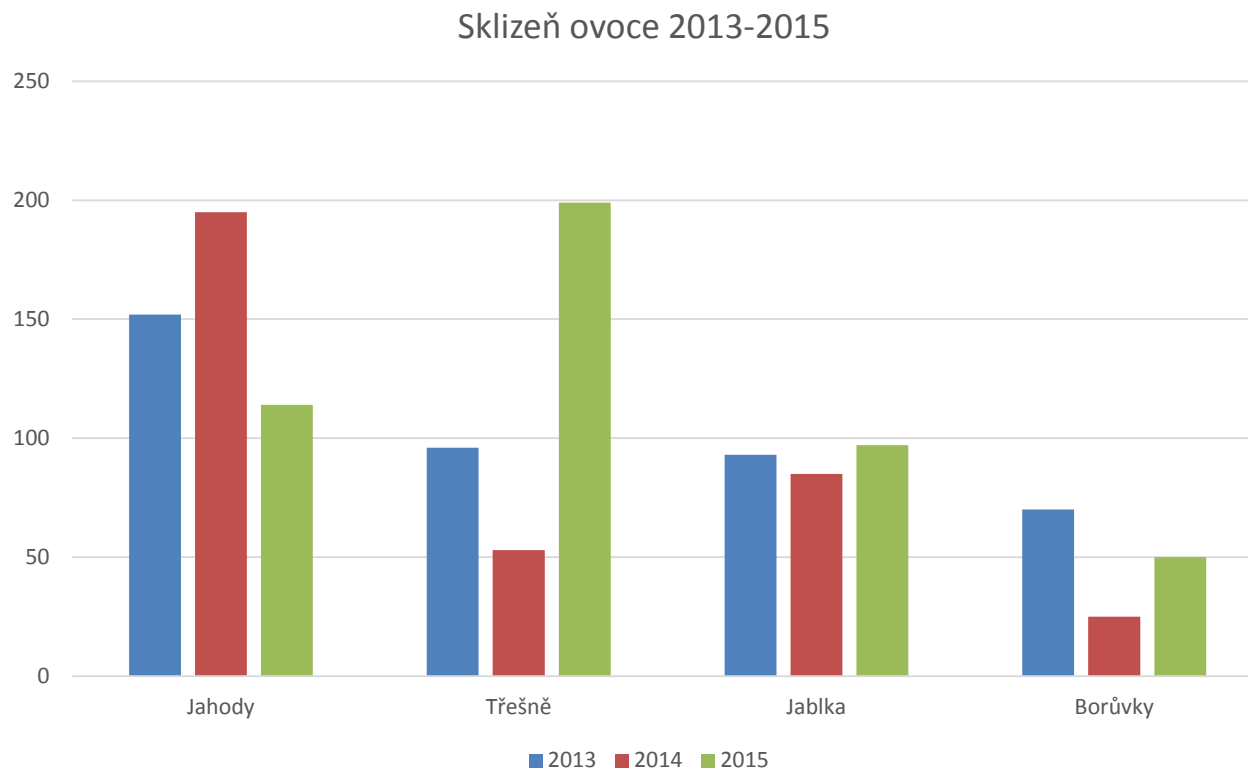
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

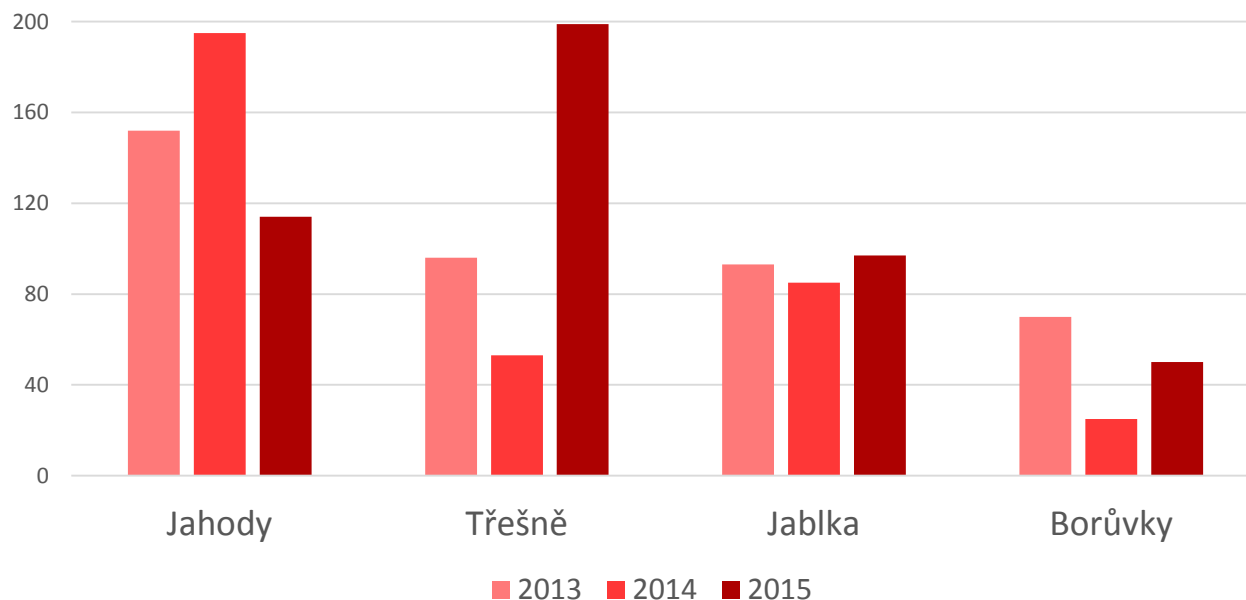
Nejvíce vzrostla sklizeň třešní
Sklizeň ovoce 2013-2015 (kilogramy)

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

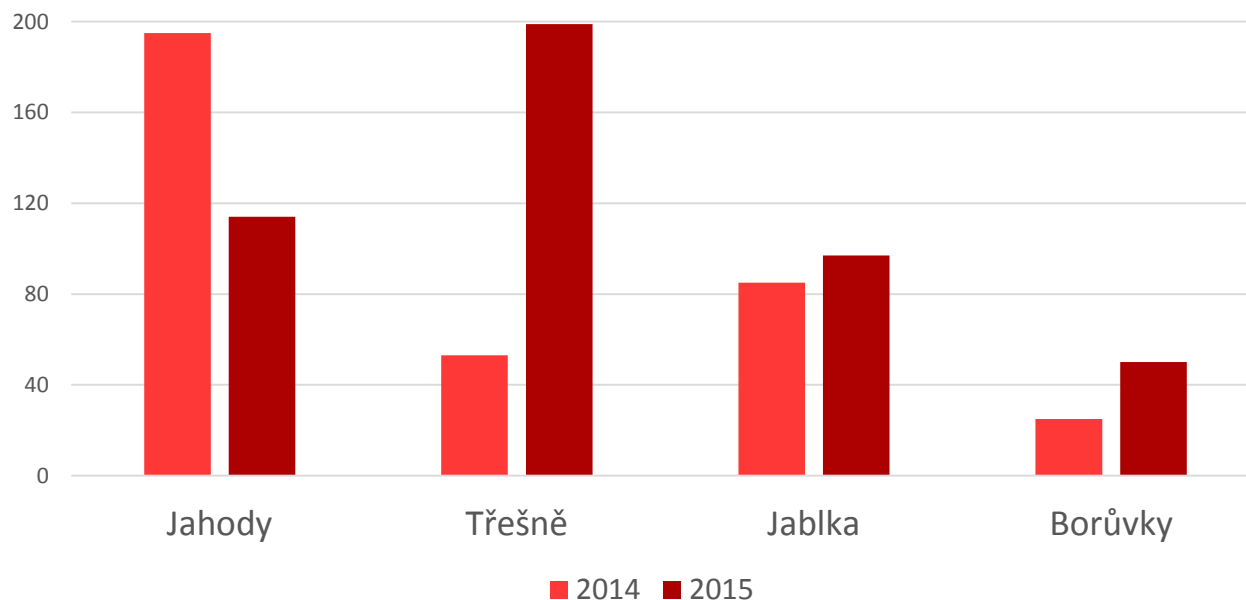
Nejvíce vzrostla sklizeň třešní
Sklizeň ovoce 2013-2015 (kilogramy)

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

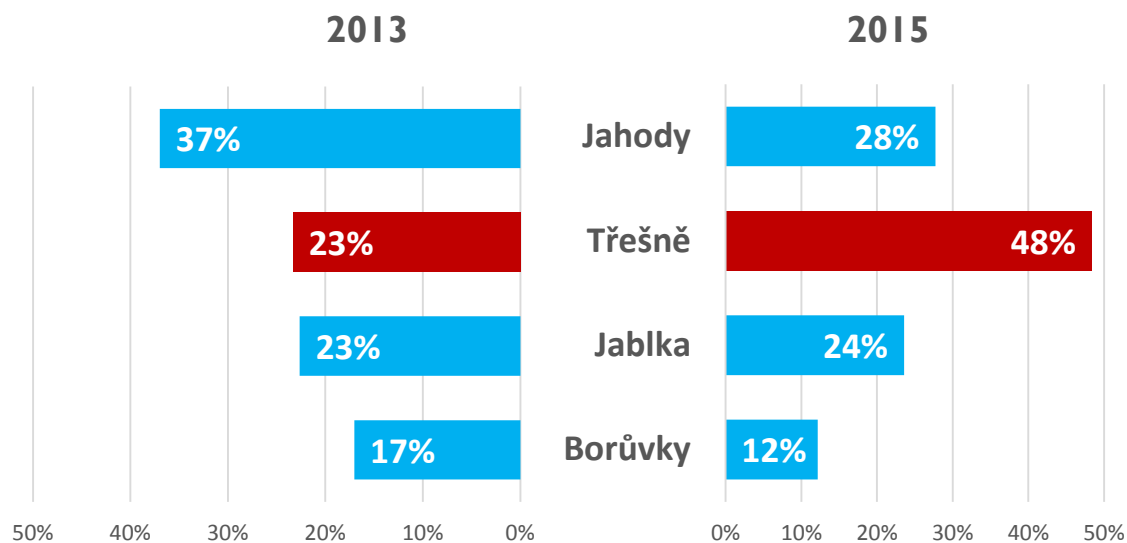
Vývoj v čase

Souvislost

Srovnání

Kombinace

Podíl třešní na sklizni se zdvojnásobil
Sklizeň ovoce 2013-2015 (kilogramy)



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

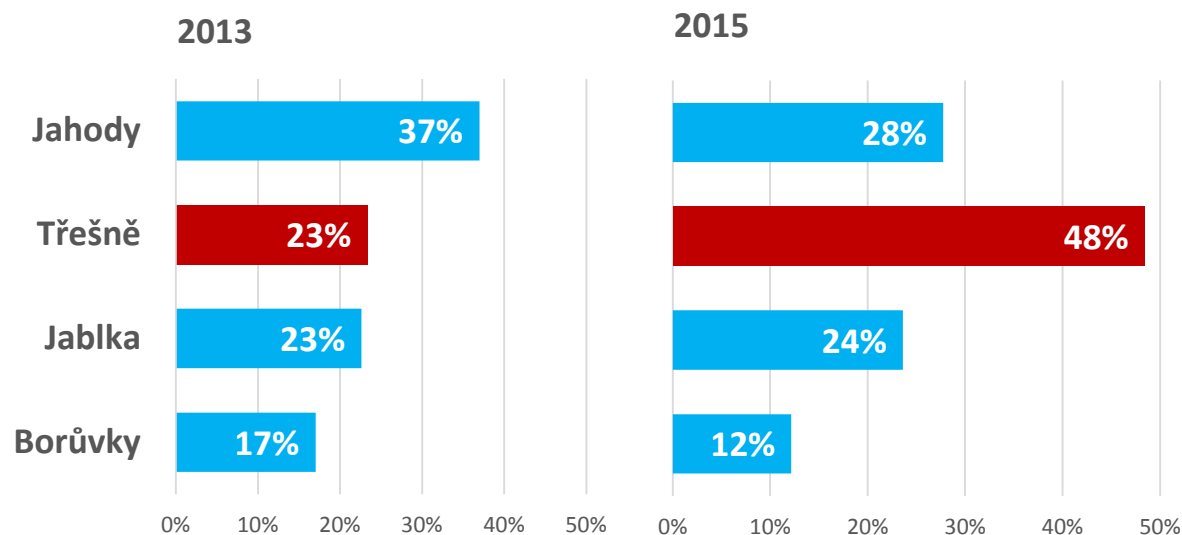
Vývoj v čase

Souvislost

Srovnání

Kombinace

Podíl třešní na sklizni se zdvojnásobil
Sklizeň ovoce 2013-2015 (kilogramy)



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

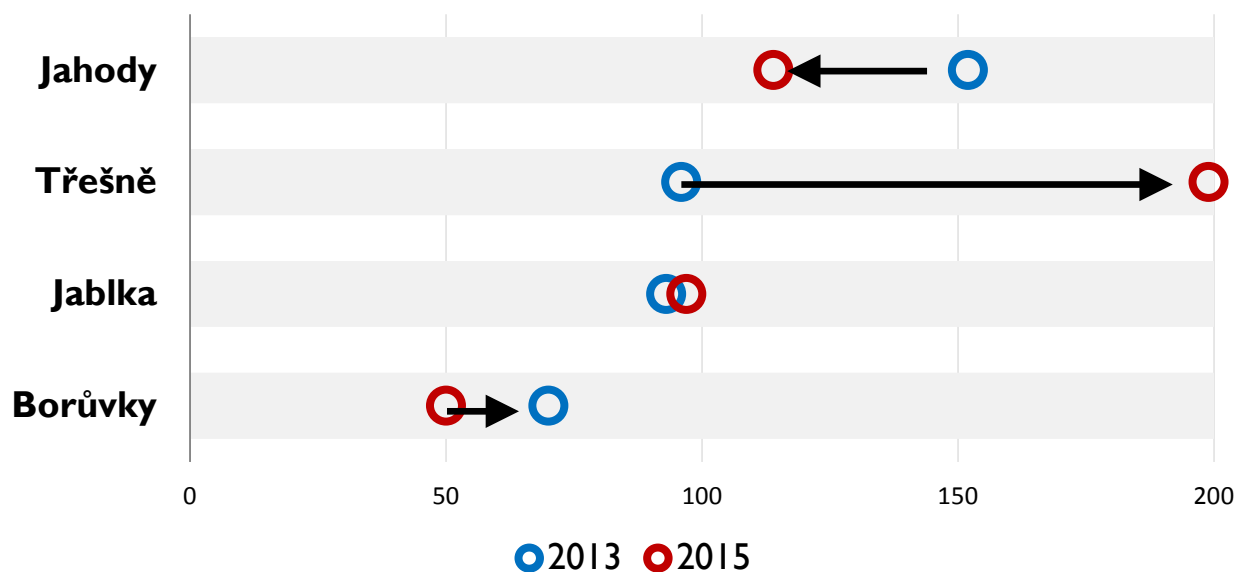
Vývoj v čase

Souvislost

Srovnání

Kombinace

Nejvíce od roku 2013 vzrostla sklizeň třešní
sklizené ovoce 2013-2015 (kilogramy)



Typ grafu zvolte podle účelu a typu dat

Sklizeň ovoce 2013-2015

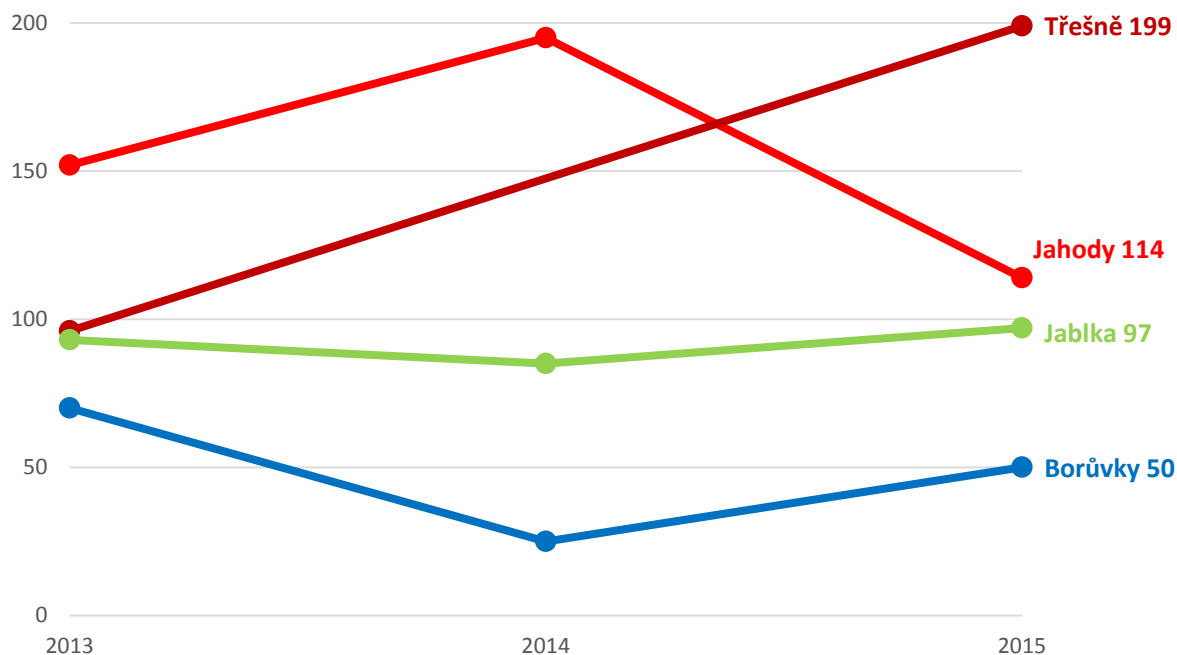
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných ovocí MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Nejméně se sklídilo borůvek, ale jejich úroda vzrostla
sklizené ovoce 2013-2015 (kilogramy)

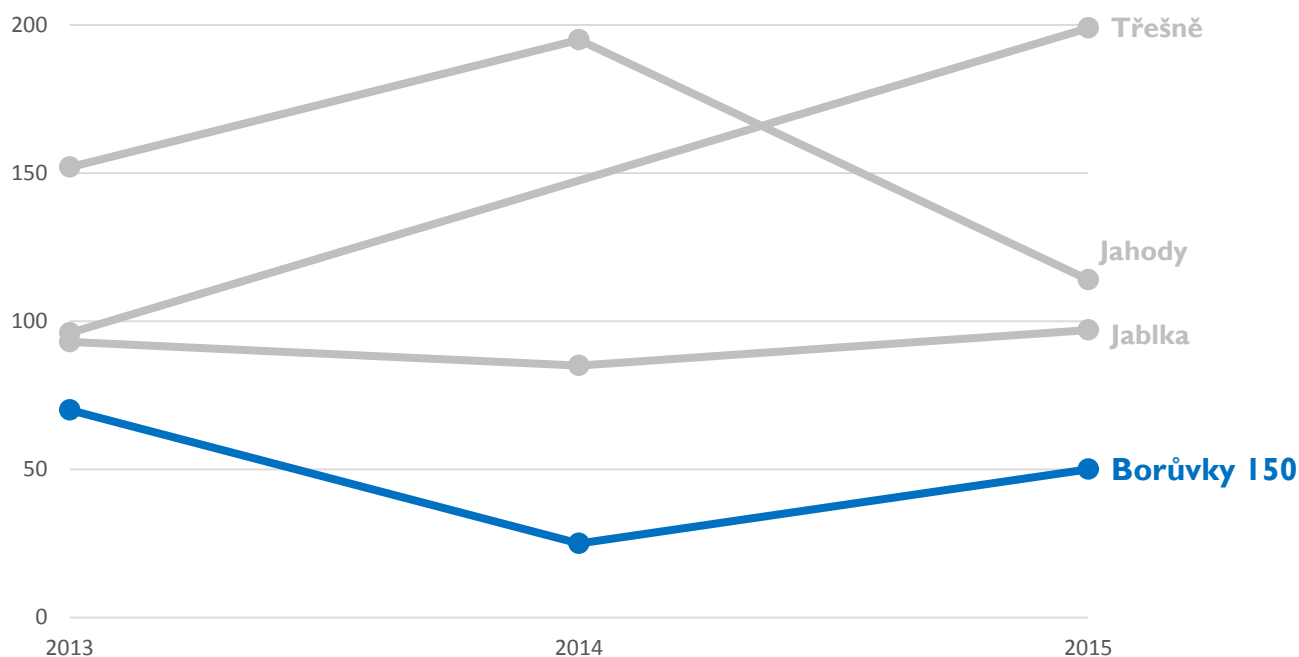
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných ovocí MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

Vývoj v čase

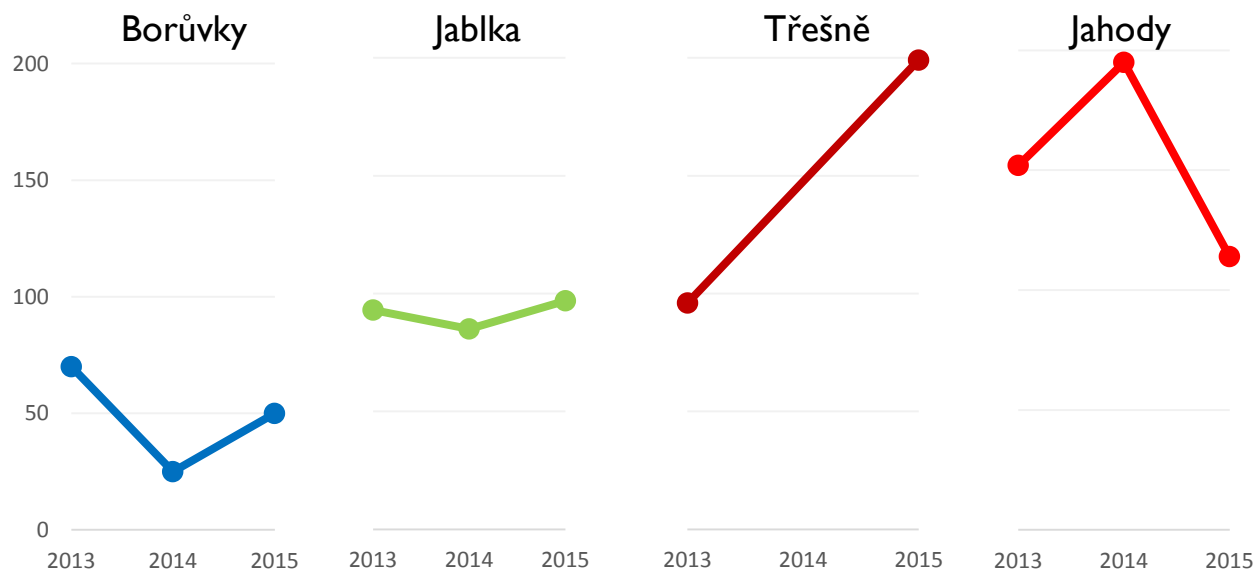
Souvislost

Srovnání

Kombinace

small multiples

Úroda různých druhů ovoce se liší a u většiny kolísá
sklizené ovoce 2013-2015



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

Vývoj v čase

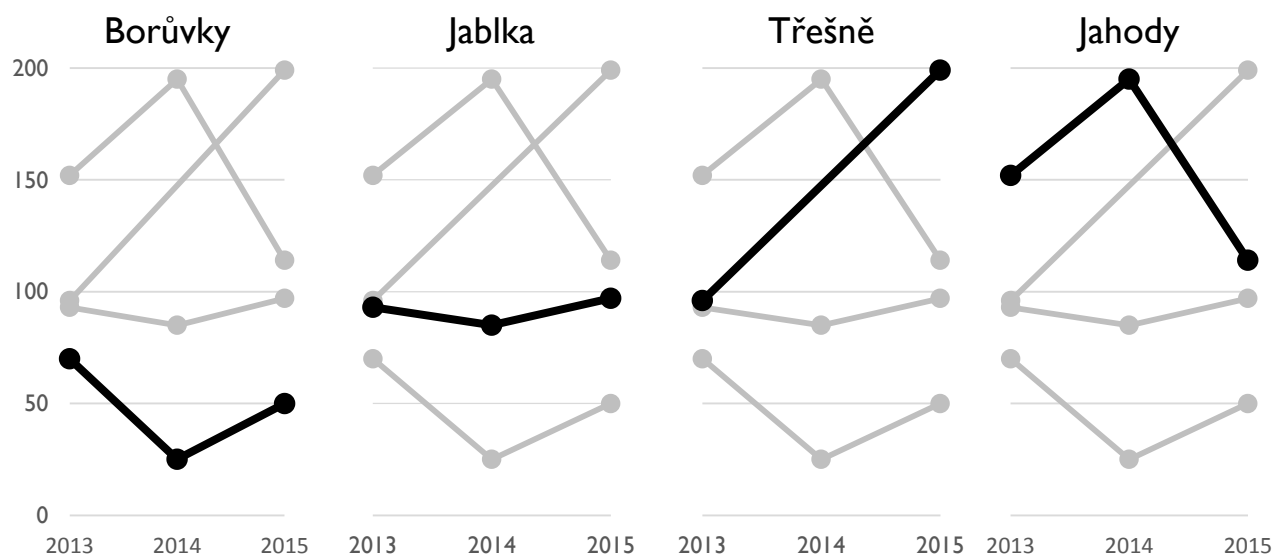
Souvislost

Srovnání

Kombinace

small multiples

Úroda různých druhů ovoce se liší a většinou kolísá
sklizené ovoce 2013-2015



Zdroj: generátor náhodného ovoce MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

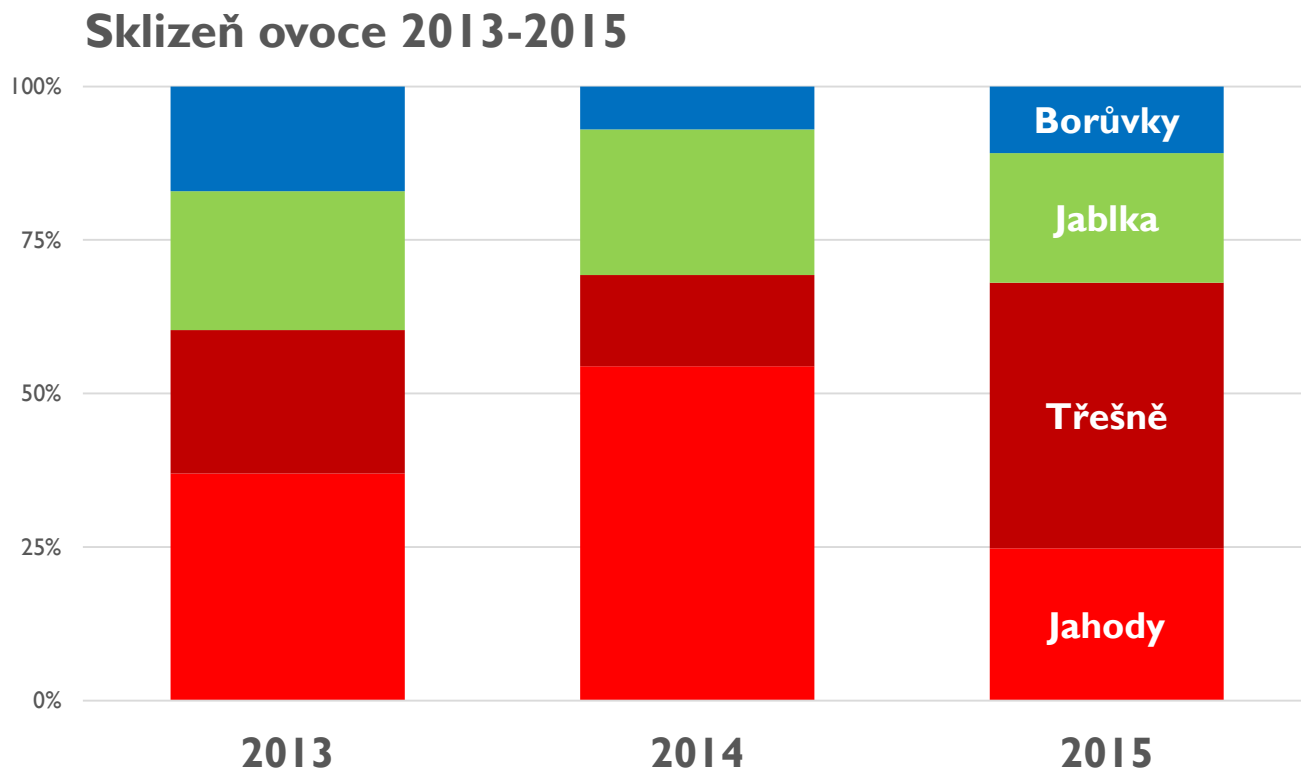
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných ovocí MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

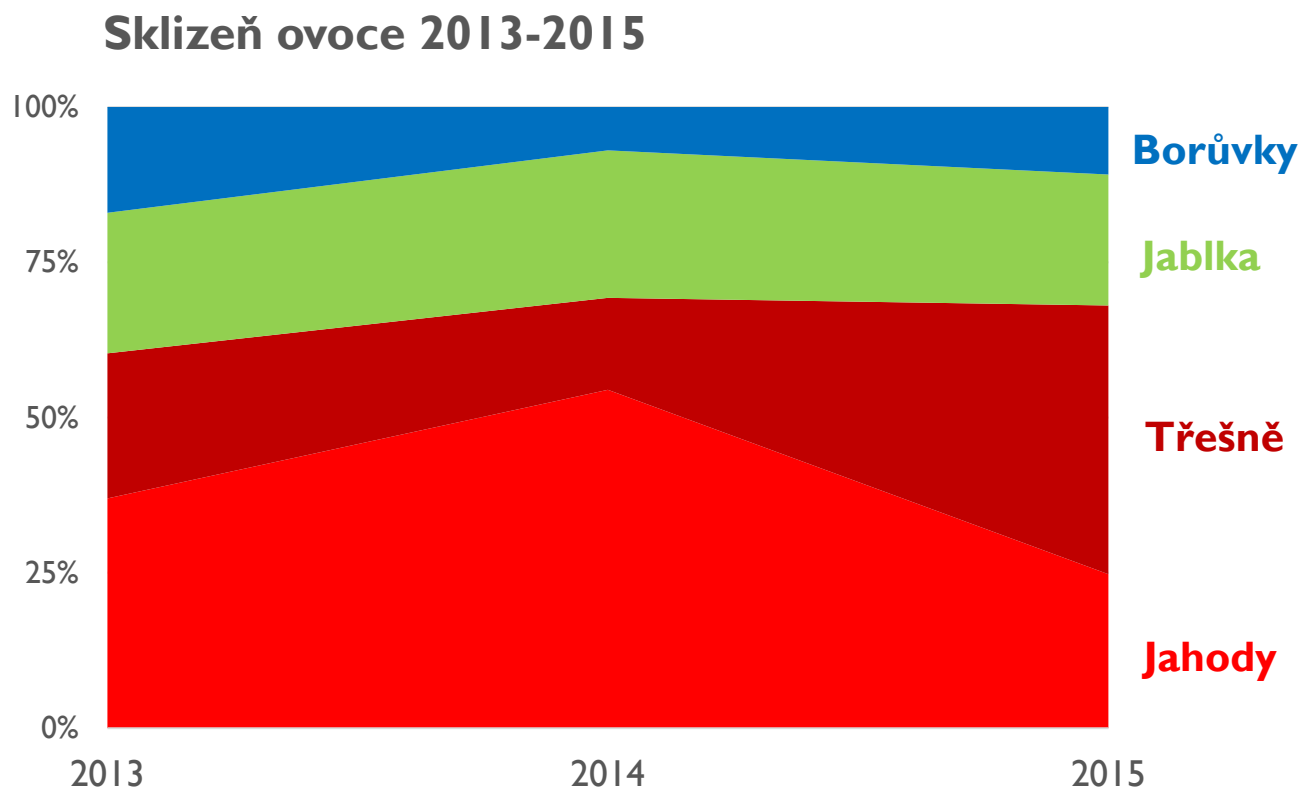
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných ovocí MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

Sparklines

Vývoj v čase

Souvislost

Srovnání

Kombinace

50 BEAUTIFUL EVIDENCE

SPARKLINES have obvious applications for financial and economic data—by tracking and comparing changes over time, by showing overall trend along with local detail. Embedded in a data table, this sparkline depicts an exchange rate (dollar cost of one euro) for every day for one year:

	2003.4.28	12 months	2004.4.28	low	high
Euro foreign exchange \$	1.1025		1.1907	1.0783	1.2858

Zdroj: Edward Tufte, *Beautiful Evidence*.

Typ grafu zvolte podle účelu a typu dat

Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace

Skupina	Trend	2010	2011	2012	2013	2014	2015
Muži		52	44	193	210	129	218
Ženy		129	156	14	186	18	135
Muži do 15 let		37	134	6	73	102	189
Muži nad 15 let		32	19	182	56	83	91
Ženy do 15 let		7	128	35	150	14	5
Ženy nad 15 let		98	19	175	152	150	83

Skupina	Trend	2010	2011	2012	2013	2014	2015
Muži		52	44	193	210	129	218
Ženy		129	156	14	186	18	135
Muži do 15 let		200	176	194	110	43	131
Muži nad 15 let		29	125	186	113	17	149
Ženy do 15 let		158	144	159	15	49	147
Ženy nad 15 let		59	84	91	64	151	65

Zdroj: generátor náhodných statistik MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

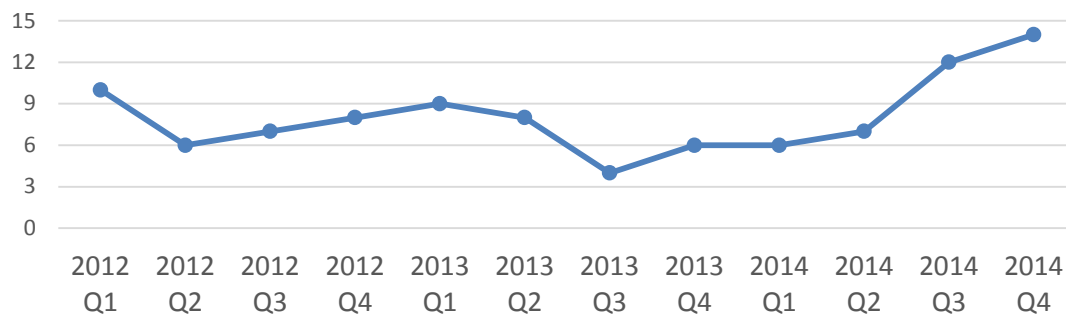
Vývoj v čase

Souvislost

Srovnání

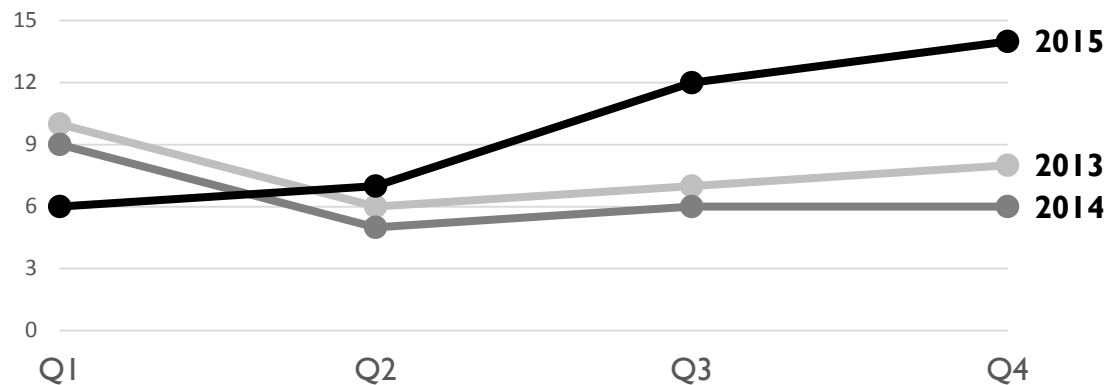
Kombinace

Počet schválených projektů 2012-2014



Trend v roce 2014 se lišil od předchozích let

Počet schválených projektů 2012-2014



Zdroj: generátor náhodných statistik MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

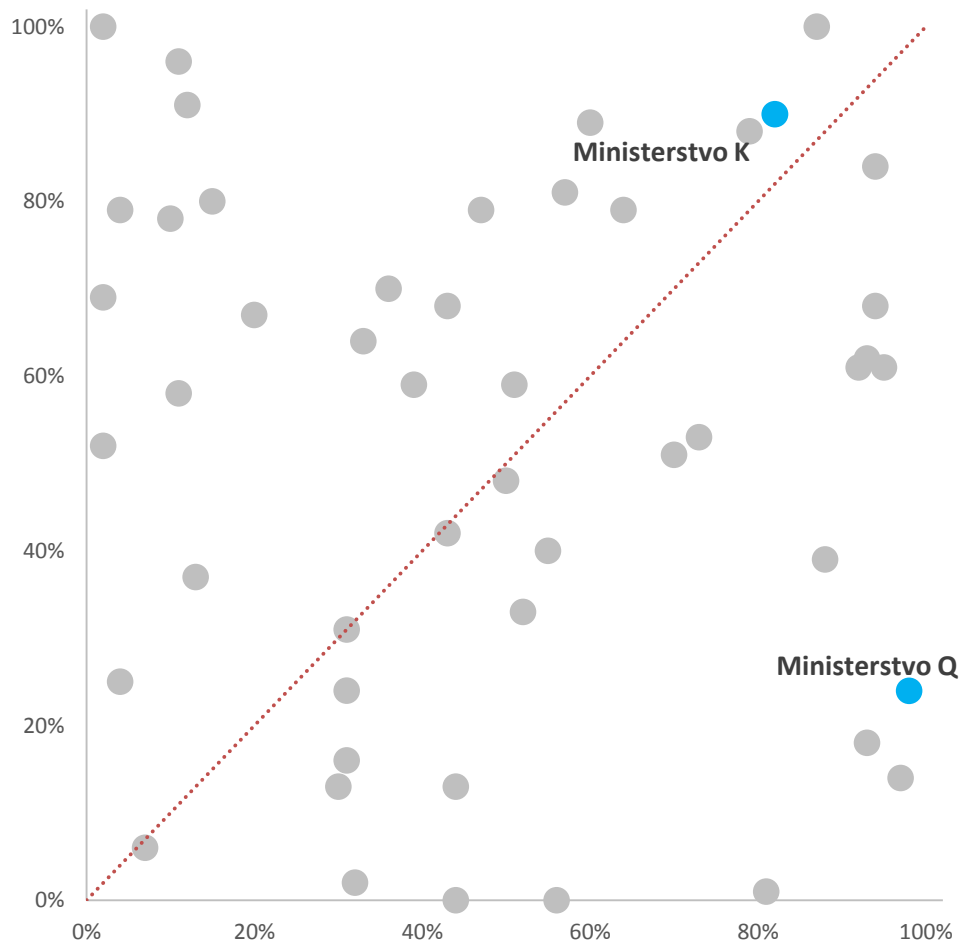
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Typ grafu zvolte podle účelu a typu dat

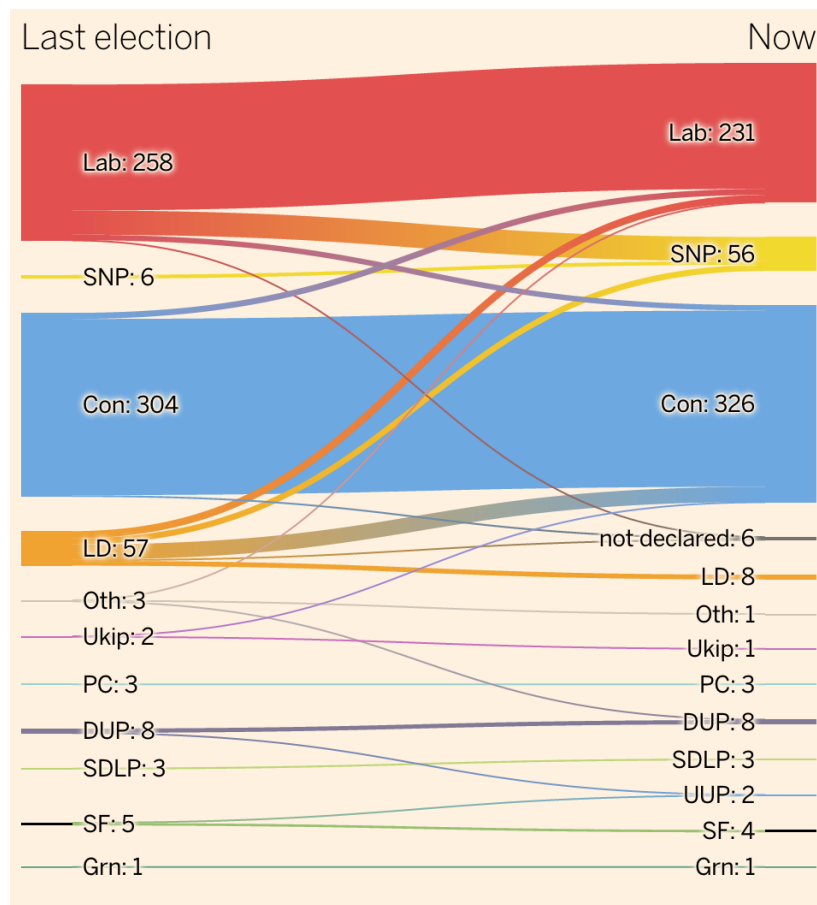
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: Financial Times online, elections.ft.com/uk/2015/results/

Typ grafu zvolte podle účelu a typu dat

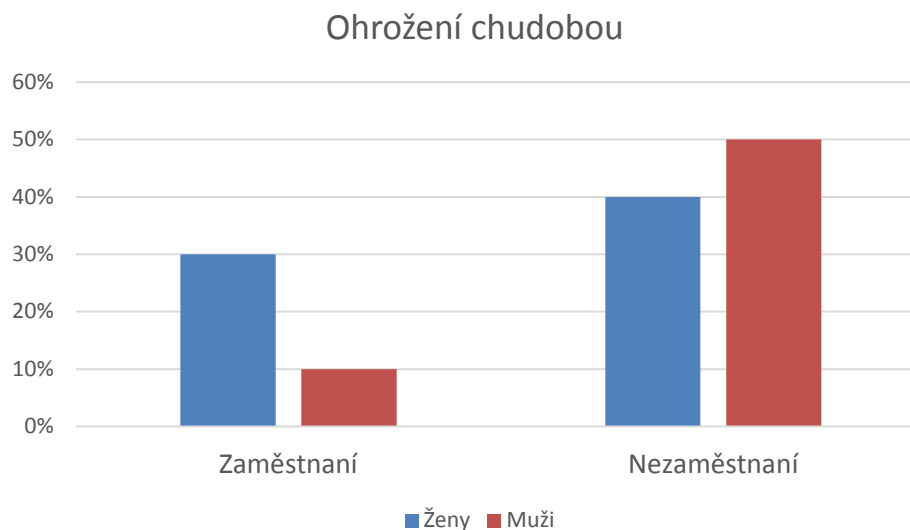
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných statistik MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

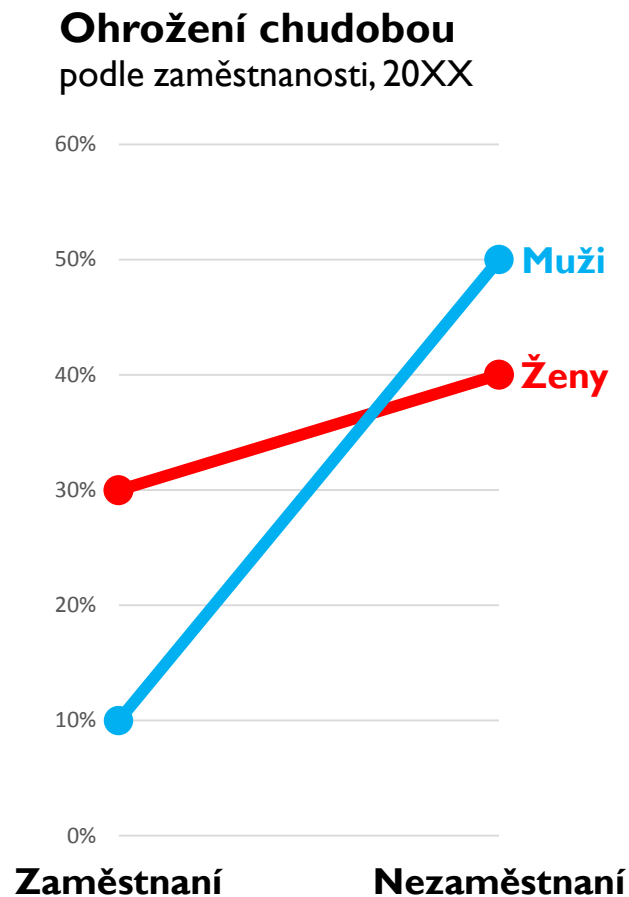
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: generátor náhodných statistik MS Excel, červen 2015

Typ grafu zvolte podle účelu a typu dat

Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace

Sustainability Assessment for Program E shows growth in 3 domains from start to end.



Zdroj: Stephanie Evergreen stephanieevergreen.com/radar-graphs/

Typ grafu zvolte podle účelu a typu dat

Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace

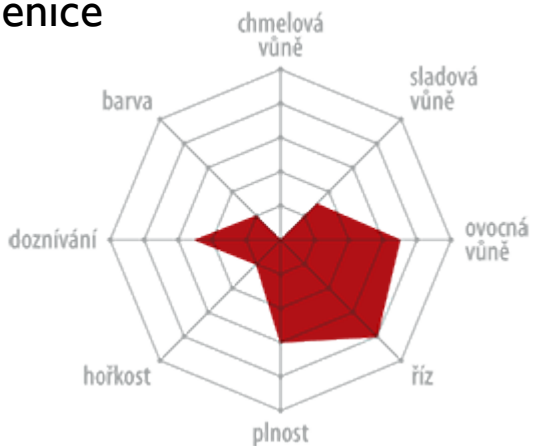
11°



12°



Pšenice



Pale Ale



Zdroj: Pivovar Hostivař pivovar-hostivar.cz/pivo

Typ grafu zvolte podle účelu a typu dat

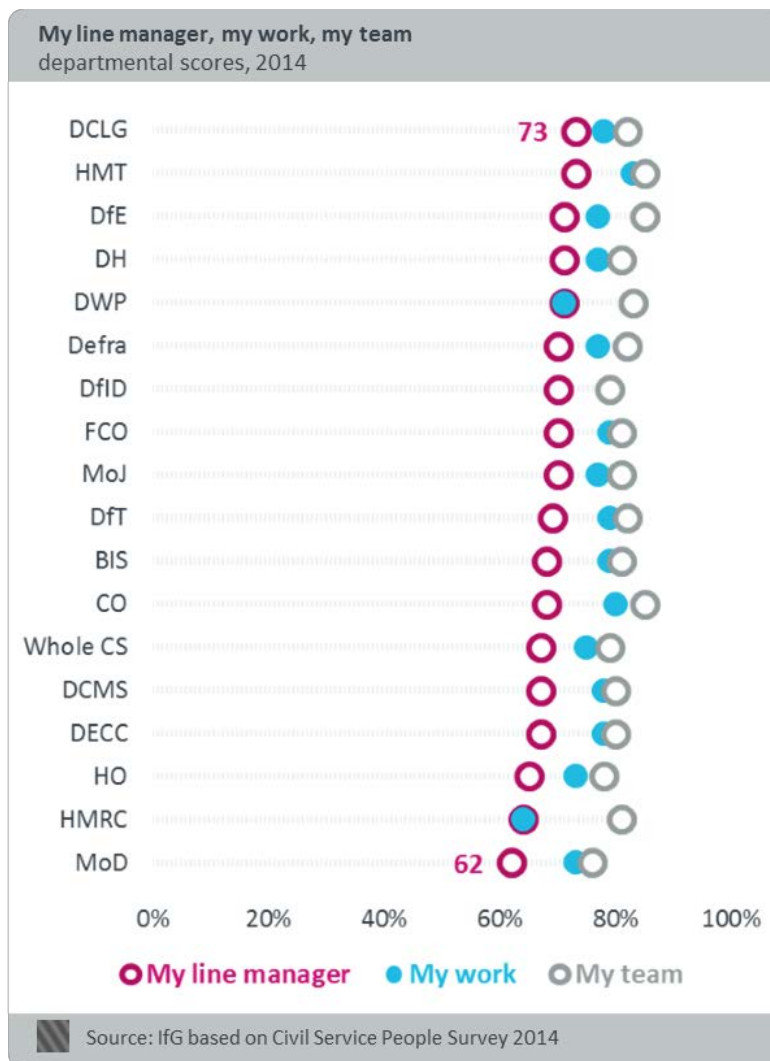
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Typ grafu zvolte podle účelu a typu dat

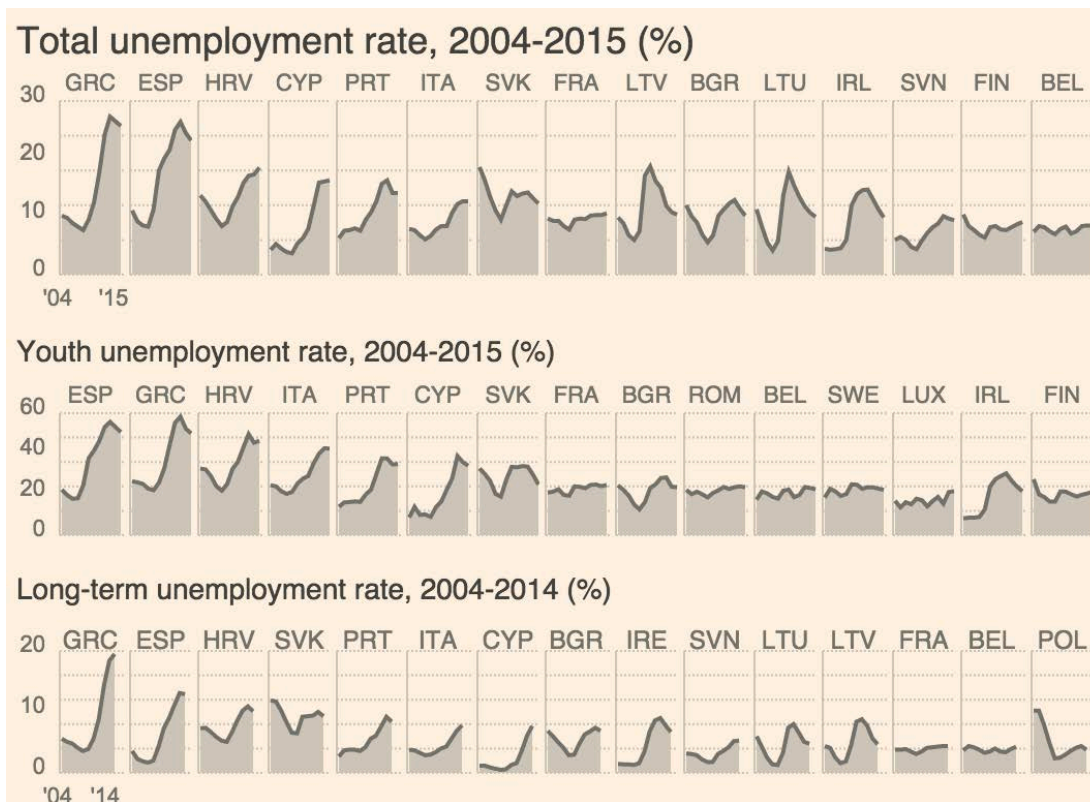
Složení

Vývoj v čase

Souvislost

Srovnání

Kombinace



Zdroj: Financial Times, Data Blog - EU Unemployment Tracker, 2015
blogs.ft.com/ftdata/2015/04/17/eu-unemployment-tracker/

Typ grafu zvolte podle účelu a typu dat

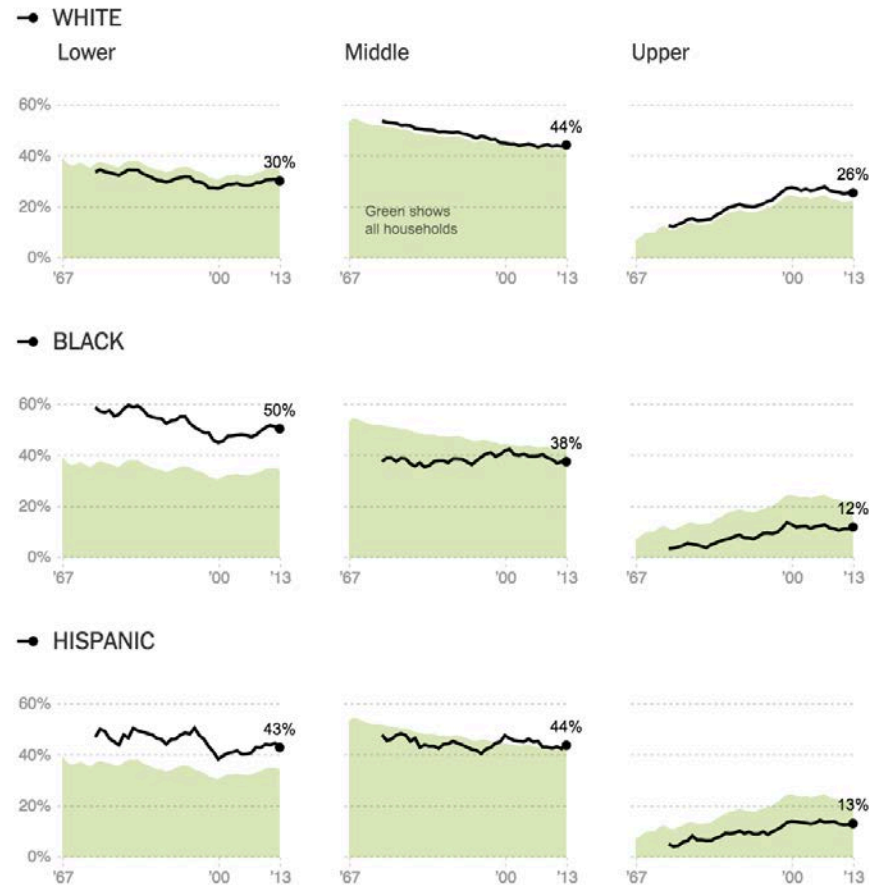
Složení

Vývoj v čase

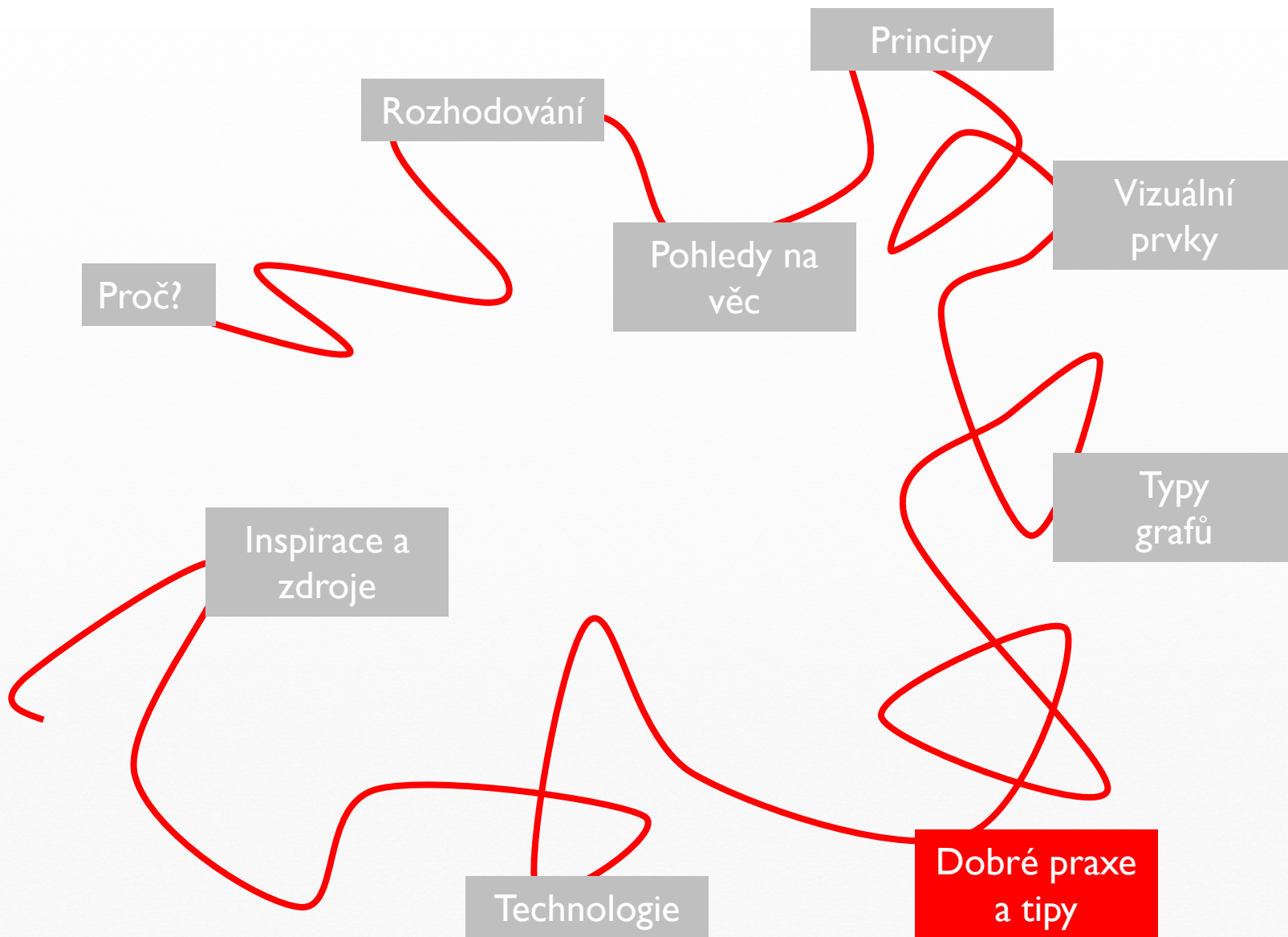
Souvislost

Srovnání

Kombinace



Zdroj: *New York Times*, The Upshot – The Shrinking Middle Class, 2015
nytimes.com/interactive/2015/01/25/upshot/shrinking-middle-class.html



Dobré praxe

- Výchozí prvky je často třeba změnit: barvy, rozsah os, popisky
- Nedatové prvky nesou význam: barvy, délky os, řazení kategorií
- Některé ale lze vizuálně upozadit
- 3D prvky jsou matoucí
- Osa Y by měla začínat na nule – pokud ne, upozorněte na to
- Popisky os a kategorií by měly být vodorovné a čitelné
- Legendy nejsou vždy potřeba a nejlepší je integrovat je s daty
- Pozor na barvoslepost (červená – zelená); lze otestovat online
- Půjde to vytisknout černobíle?

Dobré rady

- Omezte množství kategorií, pokud detaily nejsou zásadní
- Zvýrazněte, co je důležité
- Šetřete datovými popisky a desetinnými místy
- Ukažte, které body byly v datech (časové řady)
- Popis osy lze začlenit do nadpisu nebo podnadpisu
- Zařad'te kontextová data do pozadí (např. špagetový graf)
- Nadpis by měl vyjadřovat hlavní sdělení vizualizace
- Podnadpis může doplnit technické detaily
- Ne všechno musí být graf – zvažte tabulku nebo kombinaci (sparkline, heatmap)
- Ukažte to někomu, kdo nezná vaše data a vaše závěry

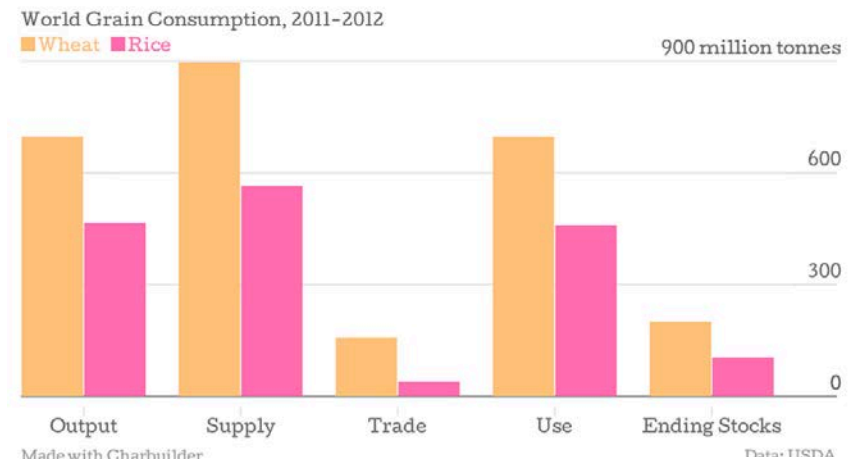
Technické tipy: když Excel nestačí



Datwrapper

Create charts and maps in just four steps. This tool reduces the time you need to create visualizations from hours to minutes.

[Login / Sign Up](#) [Try it now!](#)



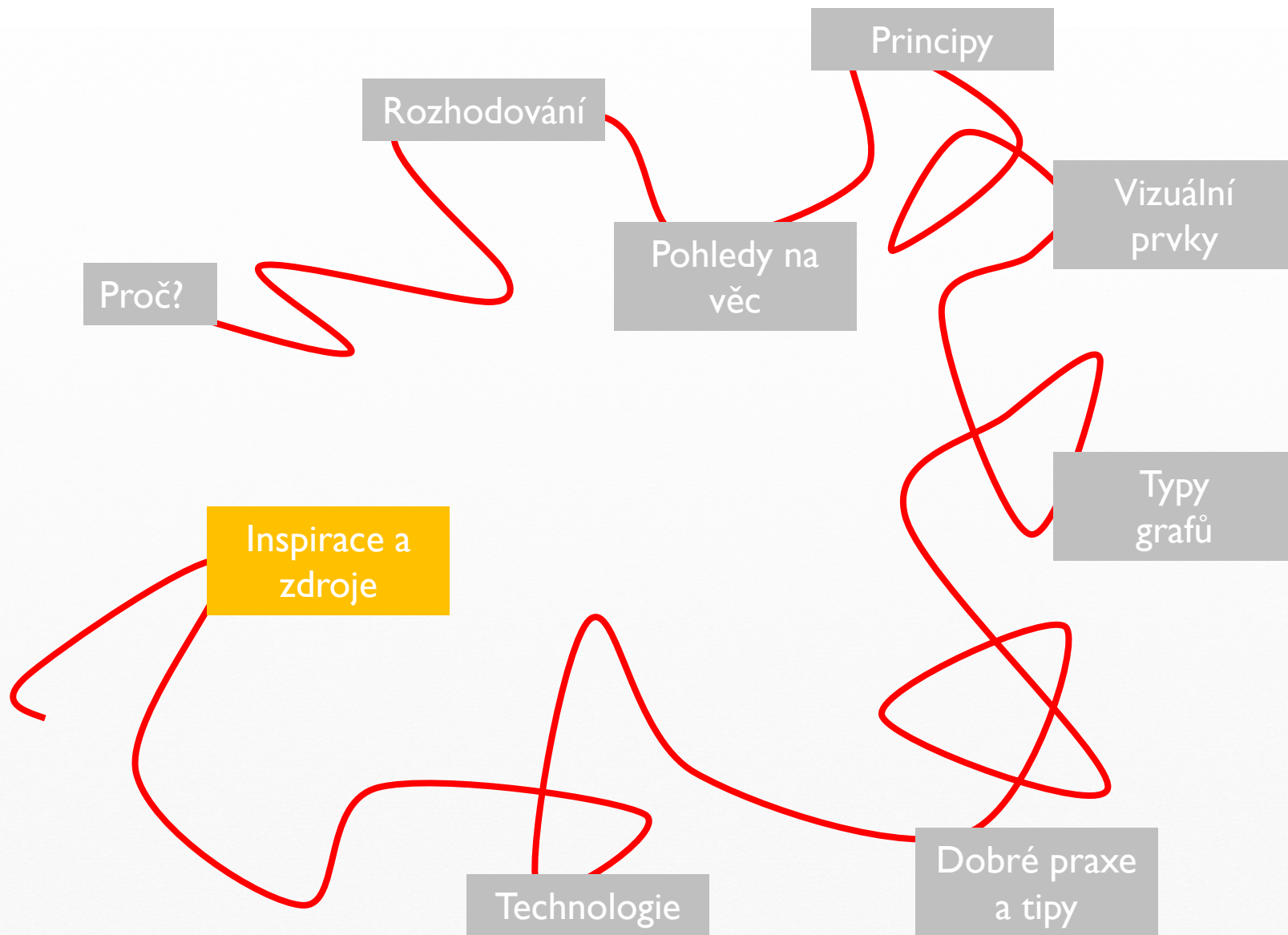
ChartBuilder

1. Input your data

names	apples	oranges
juicyness	5.5	23
color	10.2	10
flavor	6.1	13
travelability	3.8	7

Závěrem

- Neexistuje jediný správný způsob vizualizace
- Autor má zodpovědnost rozhodnout, co ukázat a jak
- Ale existují principy, kterých je třeba se držet.
- Stejně tak je třeba respektovat zákonitosti lidského vnímání.
- Vizualizace neexistují ve vakuu – jsou součástí prezentace či dokumentu, a záleží na podání a stavbě celku



Authority a inspirace

Edward Tufte: knihy + edwardtufte.com/tufte

Jacques Bertin *Semiology of Graphics*

Jean-luc Dumont, *Trees Maps and Theorems*

Stephen Few perceptualedge.com

Jonathan Schwabish policyviz.com

Ann K. Emery annkemery.com

Stephanie Evergreen stephanieevergreen.com

Andy Kirk visualisingdata.com

Robert Kosara kosara.net

Alberto Cairo thefunctionalart.com

The Economist economist.com/blogs/graphicdetail

The New York Times [@nytgraphics](https://twitter.com/nytgraphics) + Gregor Aisch vis4.net

Guardian Data Blog theguardian.com/datablog

Financial Times Data Blog blogs.ft.com/ftdata

Nathan Yau flowingdata.com

Hans Rosling gapminder.org

Office of National Statistics (Velké Británie) visual.ons.gov.uk

Český rozhlas DATA rozhlas.cz/zpravy/data

Tomáš Marek blog.vizualizacedat.cz

Děkuji za pozornost

Petr Bouchal
Evaluační jednotka NOK
petr.bouchal@mmr.cz
@petrbouchal