

# Vizualizace dat

---

ANALÝZA DAT

11. BŘEZNA 2021

# Vizualizace dat v médiích

---

# Typy proměnných

---

- nominální (nominal)
  - o dvou hodnotách lze říci pouze to, zda jsou stejné či různé
  - např. pohlaví, TV stanice, typ formy vysílání apod.

**kvalitativní data**

---
- ordinální (ordinal)
  - u hodnot je navíc možné určit jejich pořadí („co je více a co méně“)
  - např. úroveň dosaženého vzdělání, míra spokojenosti s obsahem pořadu apod.

**lze užít jako oba typy dat**

---
- intervalová (interval)
  - lze vypočítat, o kolik je jedna hodnota větší (menší) než druhá
  - např. míra růstu HDP, výše inflace apod.

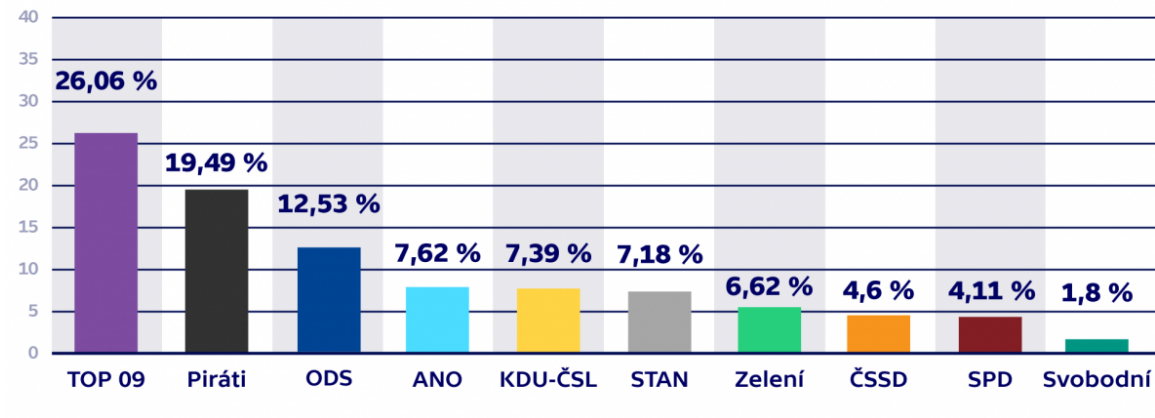
**kvantitativní data**

---
- poměrová (ratio)
  - lze vypočítat, kolikrát je jedna hodnota větší (menší) než druhá – existuje absolutní nula, pouze kladné hodnoty
  - např. počet hlasů získaných ve volbách, počet poslanců hlasujících pro radní ČRo apod.

# Zobrazení kvantitativních dat - sloupcový graf

- slouží k porovnání hodnot několika proměnných
- může být složený z několika součástí

## Výsledky voleb do PSP ČR 2017 – hlasování v zahraničí

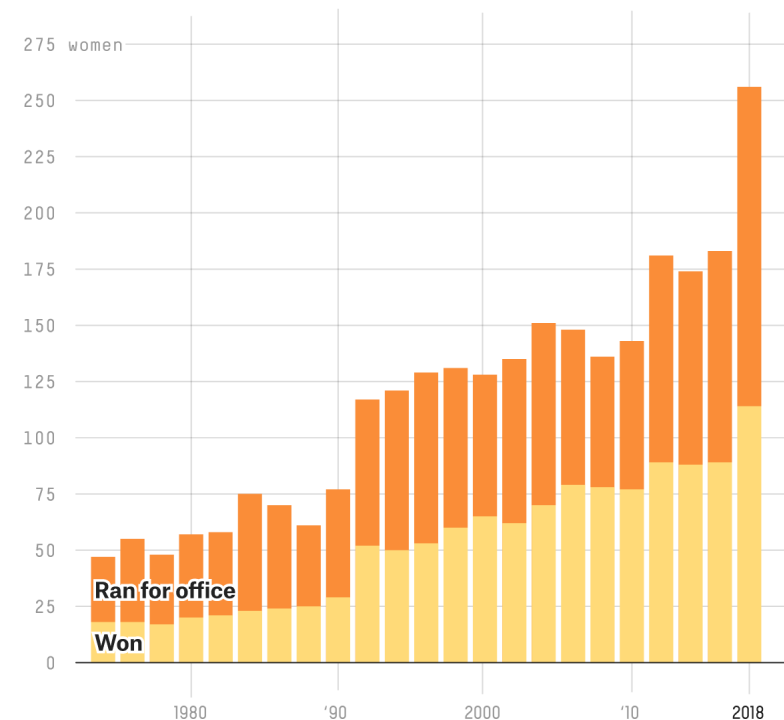


Zdroj: ČSÚ

ceskatelevize.ct24.cz

## A record year for female members of Congress

The number of women running for the U.S. House and Senate and the share who won, 1974-2018



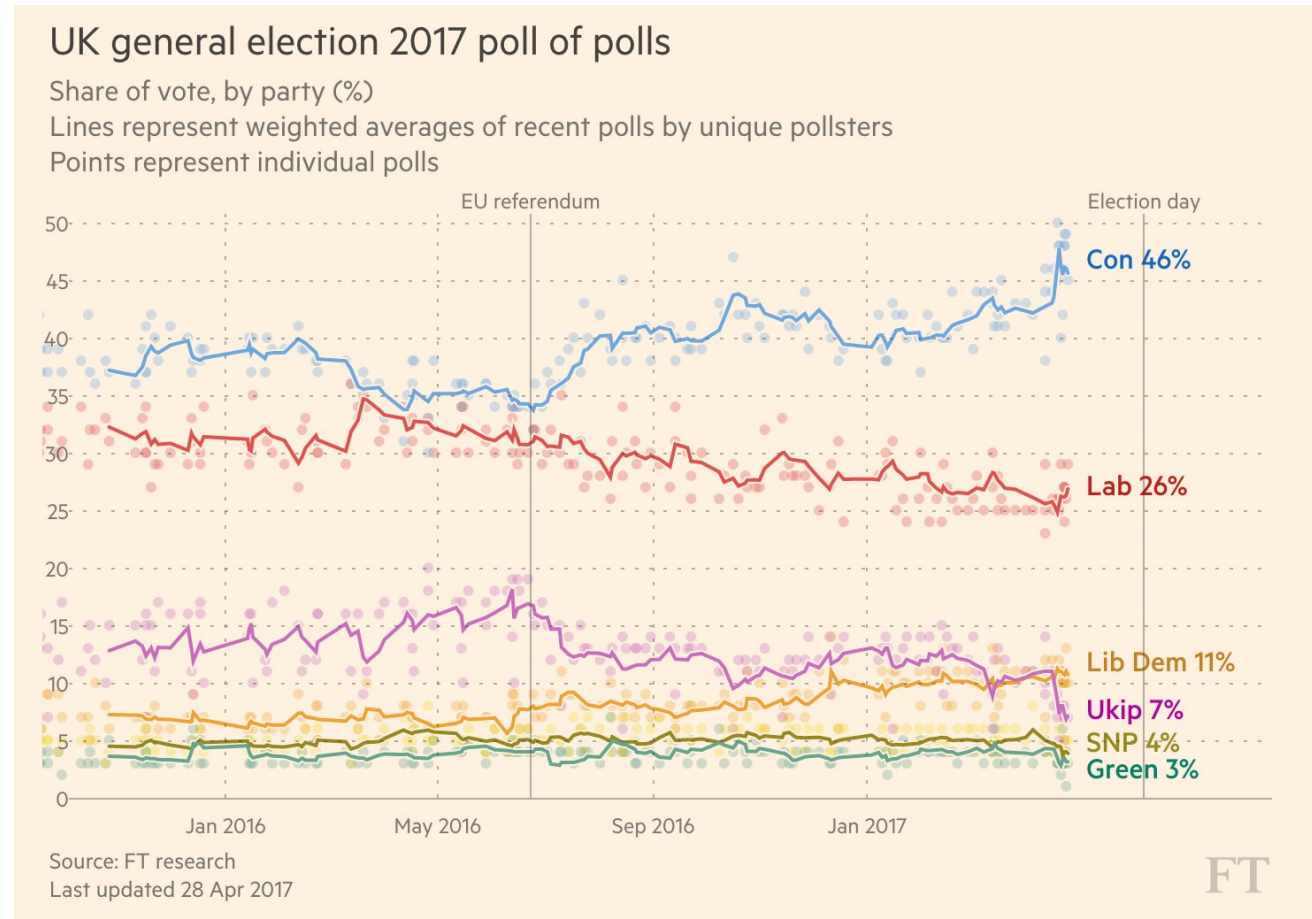
The data on women candidates who won in 2018 is still preliminary. Data is as of Nov. 13, 2018.

fivethirtyeight.com



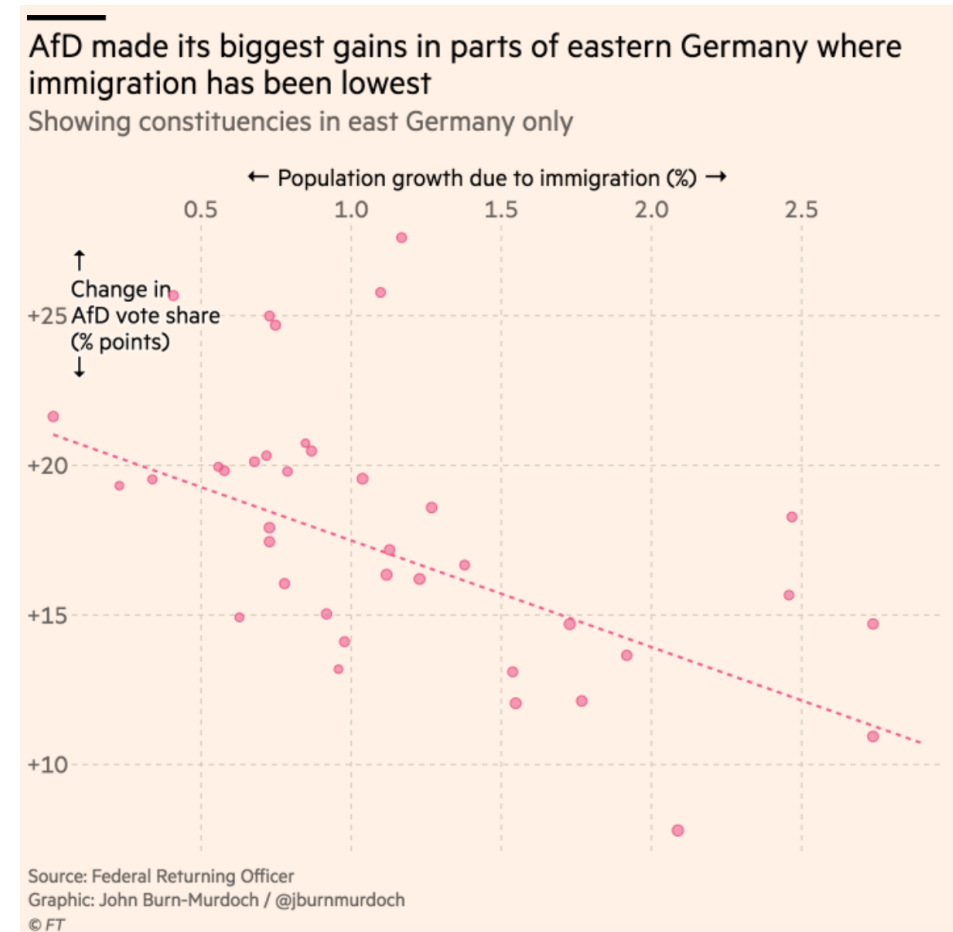
# Zobrazení kvantitativních dat - liniový graf

- zobrazení časových trendů
- většinou propojení bodového grafu a linie
- podstatou je spojitost jednotlivých pozorování



# Zobrazení kvantitativních dat - bodový graf (*scatter plot*)

- přímé zobrazení pozorování ve 2D prostoru (klidně i 3D)
- nejužitečnější ze všech statistických zobrazení
- na horizontální ose x je zobrazena nezávisle proměnná X, na vertikální ose y je zobrazena závisle proměnná Y (není to ale bezpodmínečné pravidlo)
- je možné pozorovat vztahy mezi proměnnými

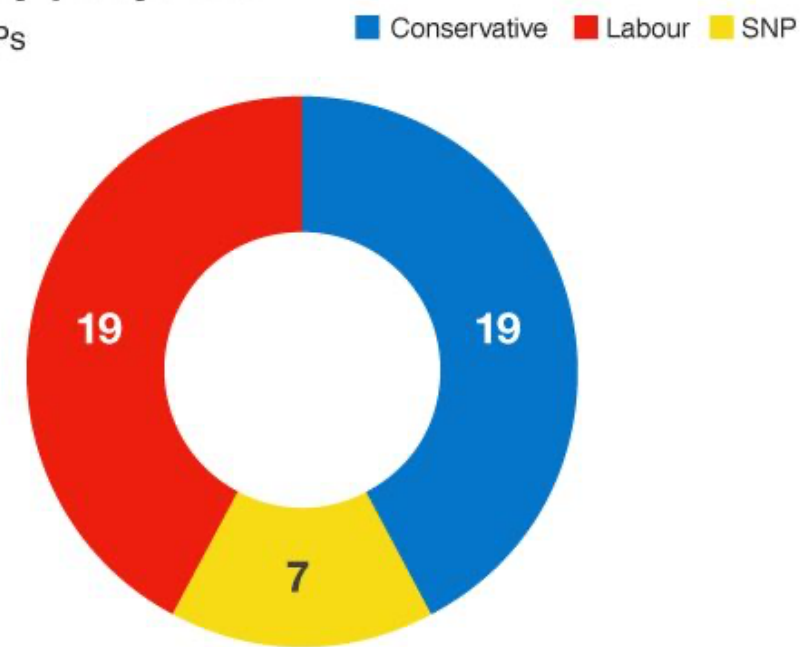


# Zobrazení kvantitativních dat - koláčový graf (*pie chart*)

- zobrazení dat s dopočtem do komplexního celku
- číselně lze zobrazit absolutní hodnoty, protože procentuální vztahy ukazuje graf
- různé druhy zobrazení

**LGBT MPs by party 2017**

45 out of 650 MPs

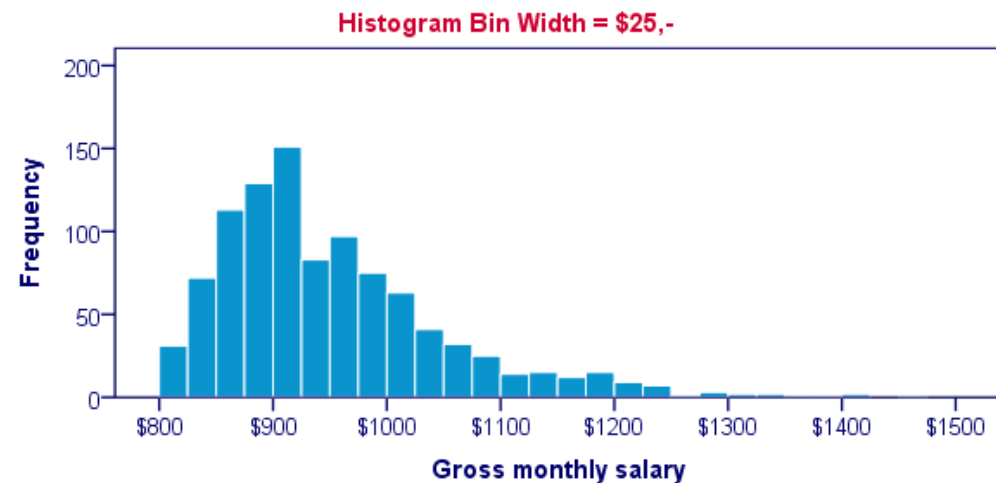
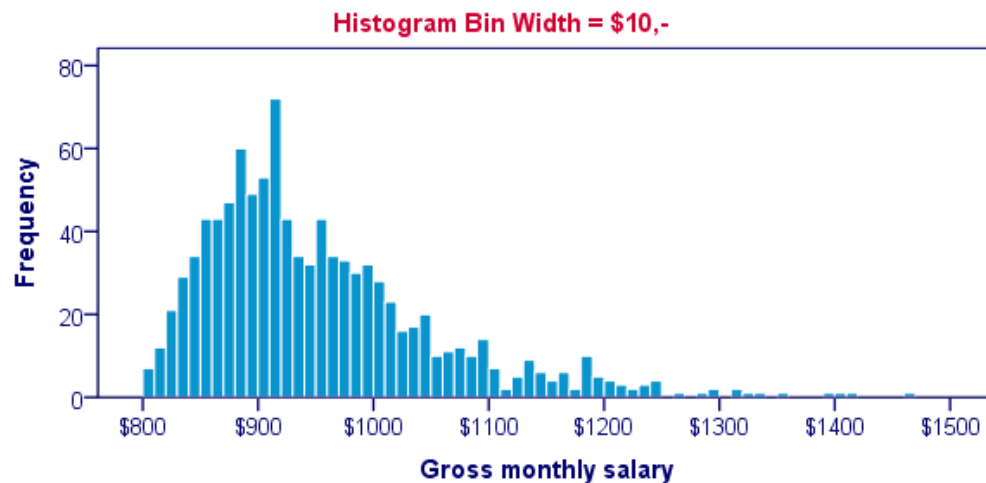


Source: John Peart, Andrew Reynolds UNC

BBC

# Zobrazení kvantitativních dat - histogram

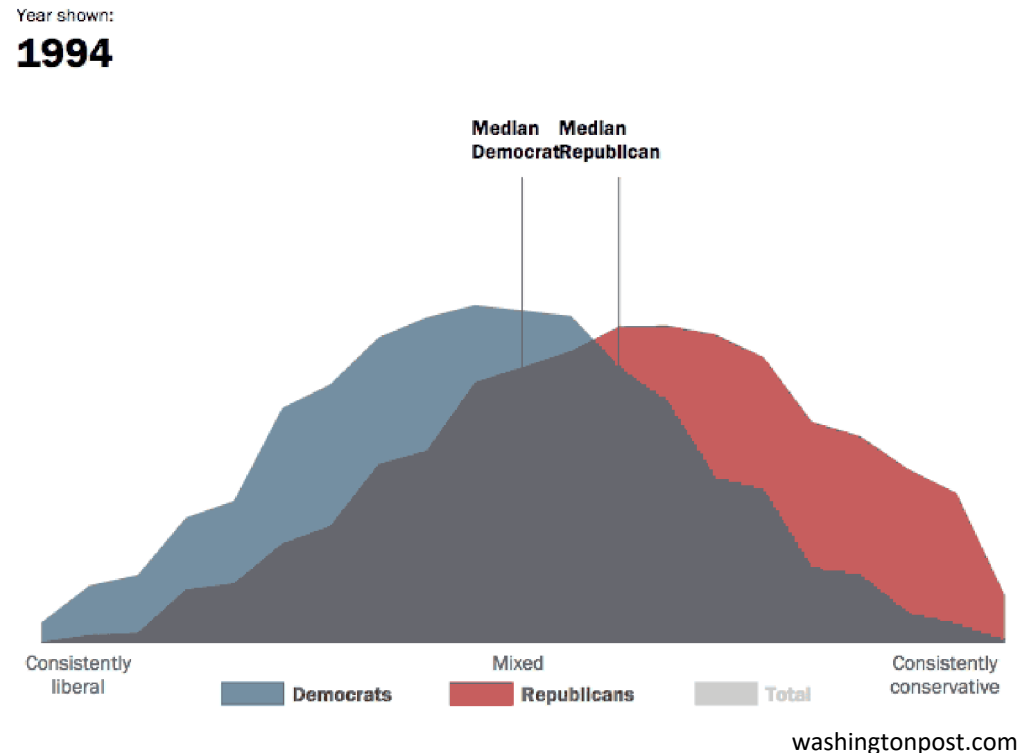
- ukazuje distribuci hodnot jedné proměnné (rozdíl oproti sloupcovému grafu)
- výška každého bloku ukazuje, kolik pozorování spadá do dané kategorie
- jednotlivé kategorie mohou být různě detailní (zobrazení vybíráme podle situace)



[spss-tutorials.com](http://spss-tutorials.com)

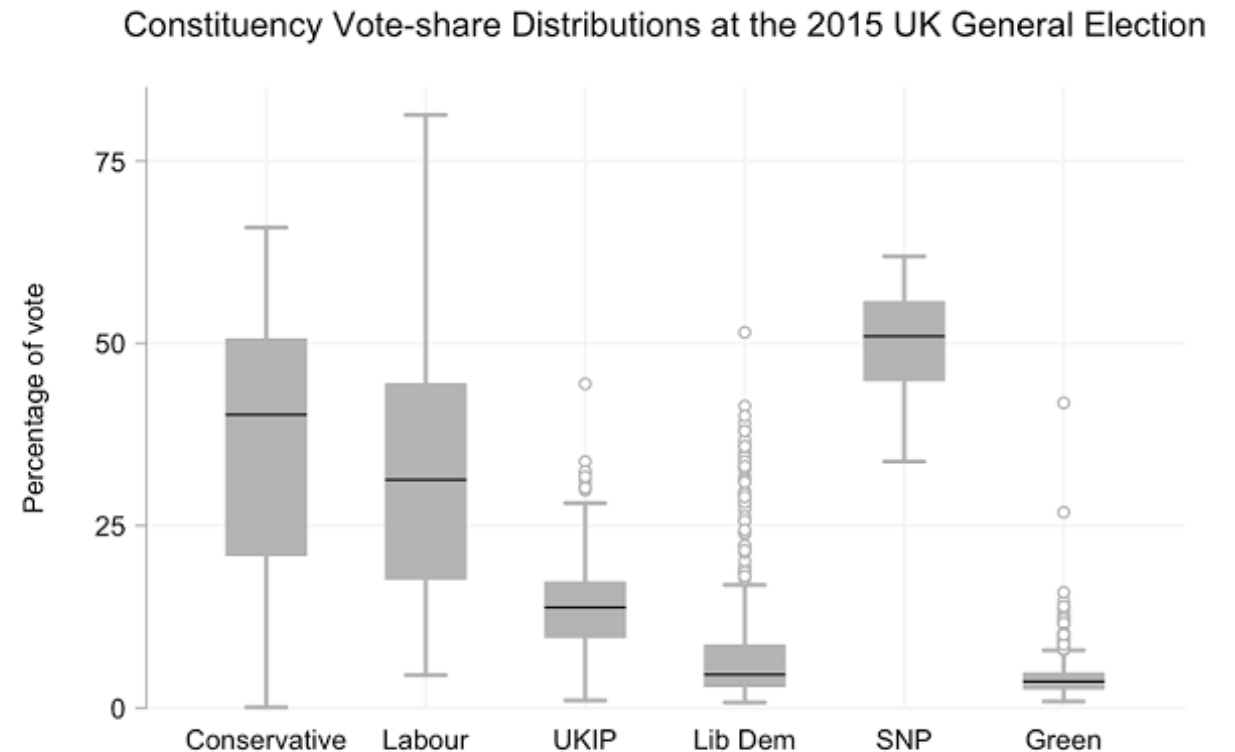
# Zobrazení kvantitativních dat - hustota rozložení (*density plot*)

- rozložení hodnot také můžeme zakreslit pomocí zprůměrování a zahlazení jednotlivých frekvencí z histogramu
- grafy hustoty rozložení jsou zvláště cenné, když se zajímáme o rozložení pravděpodobností výskytu
- na ose x jsou jednotlivé nabývané hodnoty
- osa y je obtížně interpretovatelná (záleží na konkrétní situaci) – zásadní je všimnout si tvaru, vrcholů apod.



# Zobrazení kvantitativních dat - krabicový diagram (*boxplot*)

- komplexní zobrazení distribuce hodnot proměnné
- krabice vyznačuje první, druhý (medián; černá čára) a třetí kvartil
- „vousy“ (šedé vertikální čáry s horizontálním ukončením) vyznačují krajní hodnoty mezi horním, respektive dolním kvartilem, a 1,5násobkem vzdálenosti mezi prvním a třetím kvartilem
- odlehlé hodnoty jsou od horního či dolního kvartálu vzdáleny více než 1,5násobku vzdálenosti mezi prvním a třetím kvartilem (vyznačeny prázdnými body)



infragistics.com

# Zobrazení kvantitativních dat - wordcloud

- zobrazení četnosti textových jednotek
- může jít o čistou frekvenci slov či slovních spojení nebo i výstupy složitějších analýz
- stanovuje se minimální četnost i vztah četností k velikosti zobrazení
- vyžaduje předběžnou analýzu slovních spojení či slov, která nebudou zobrazena
- výhodou je lehká interpretace pro laiky

## Povolání zvolených senátorských kandidátů (2008–2018)



Zdroj: ČT, ČSÚ

ct24.ceskatelive.cz



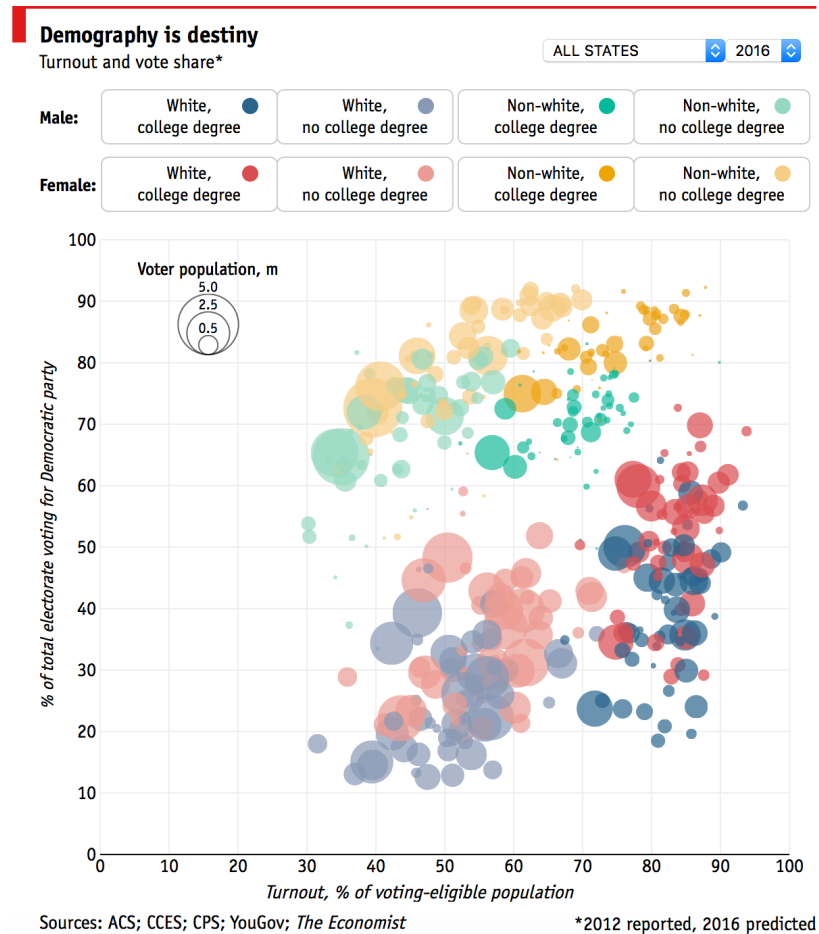
# Netradiční zobrazení



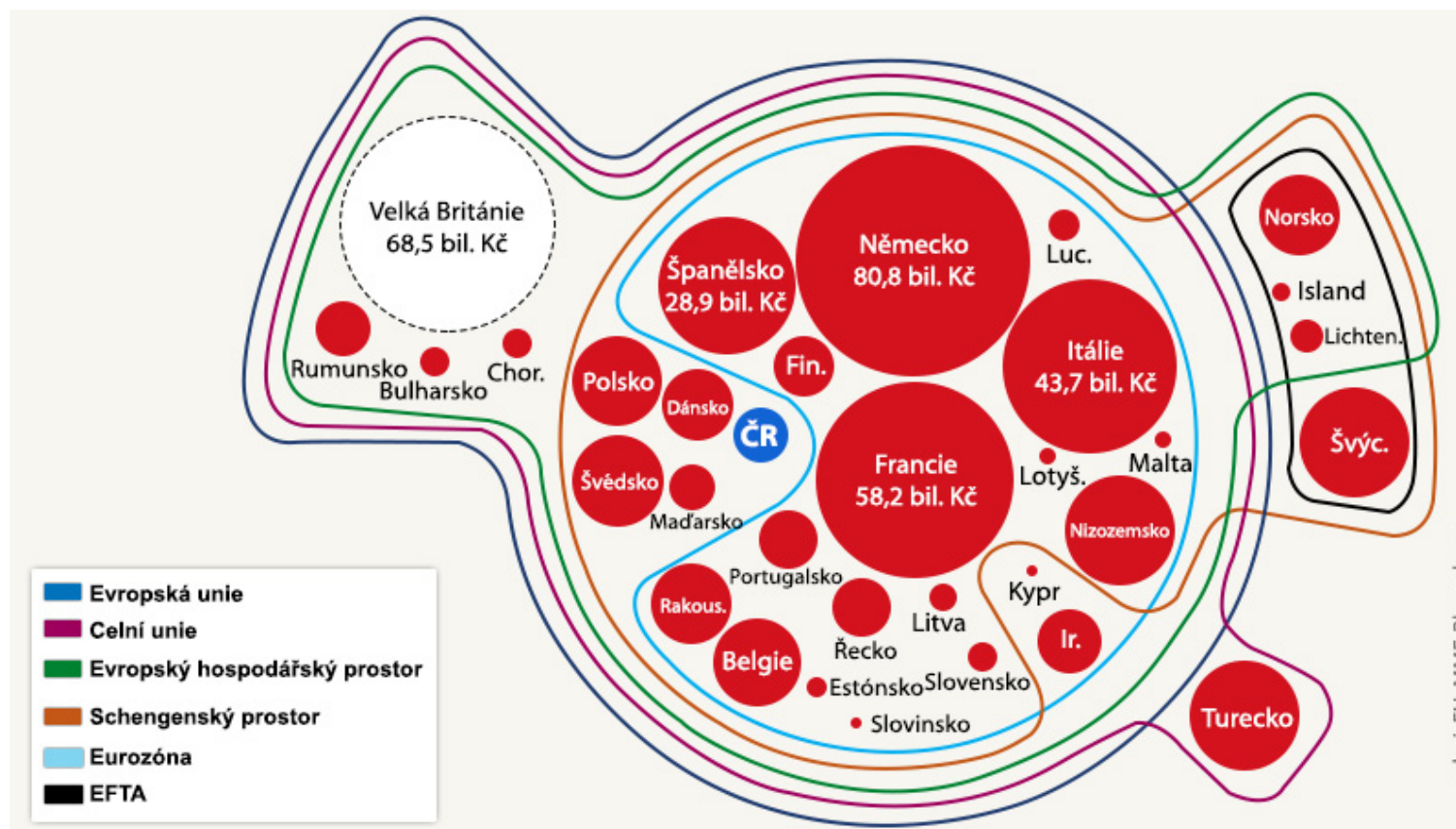
popsci.com



# Netradiční zobrazení

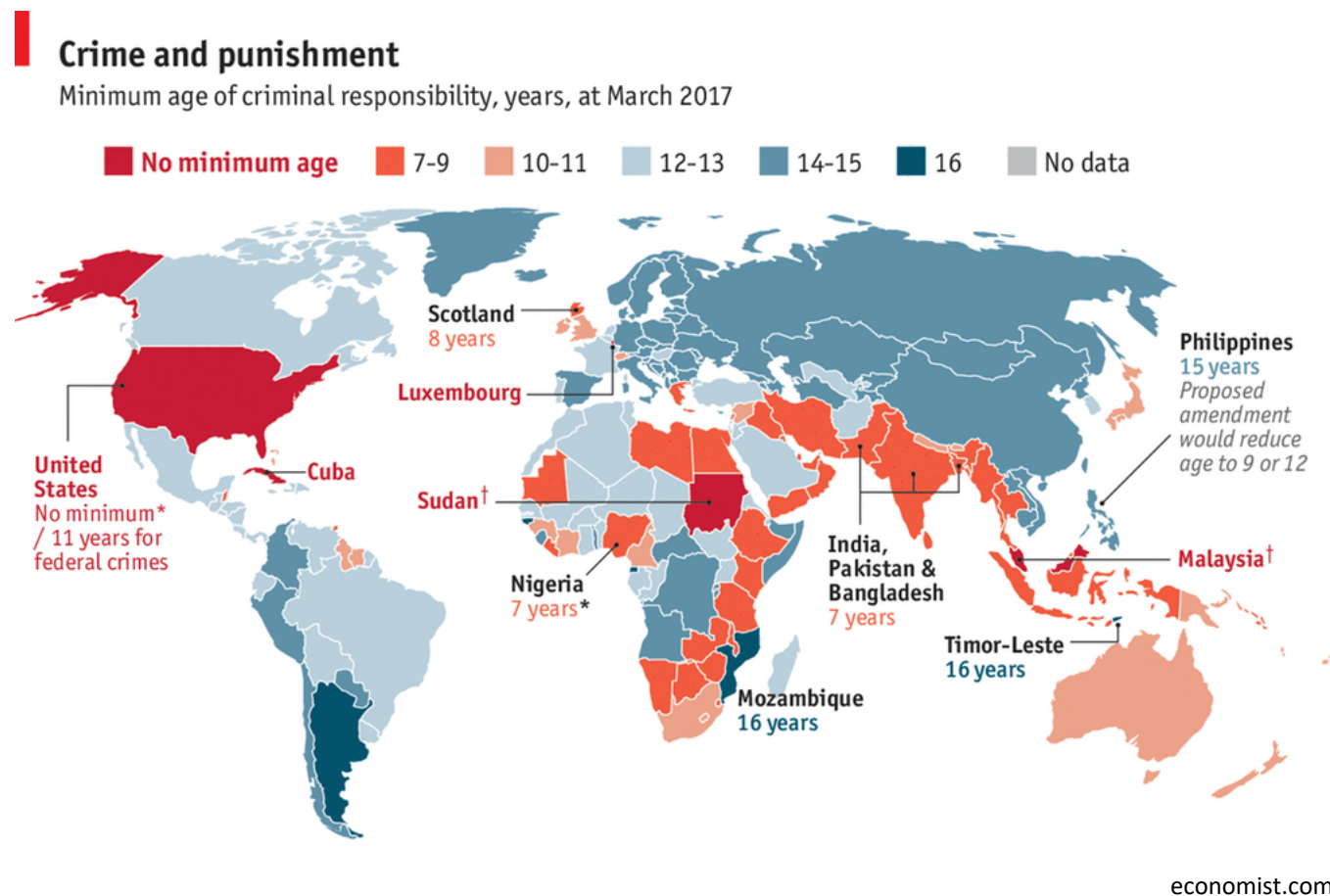


# Netradiční zobrazení



novinky.cz

# Netradiční zobrazení



# Netradiční zobrazení



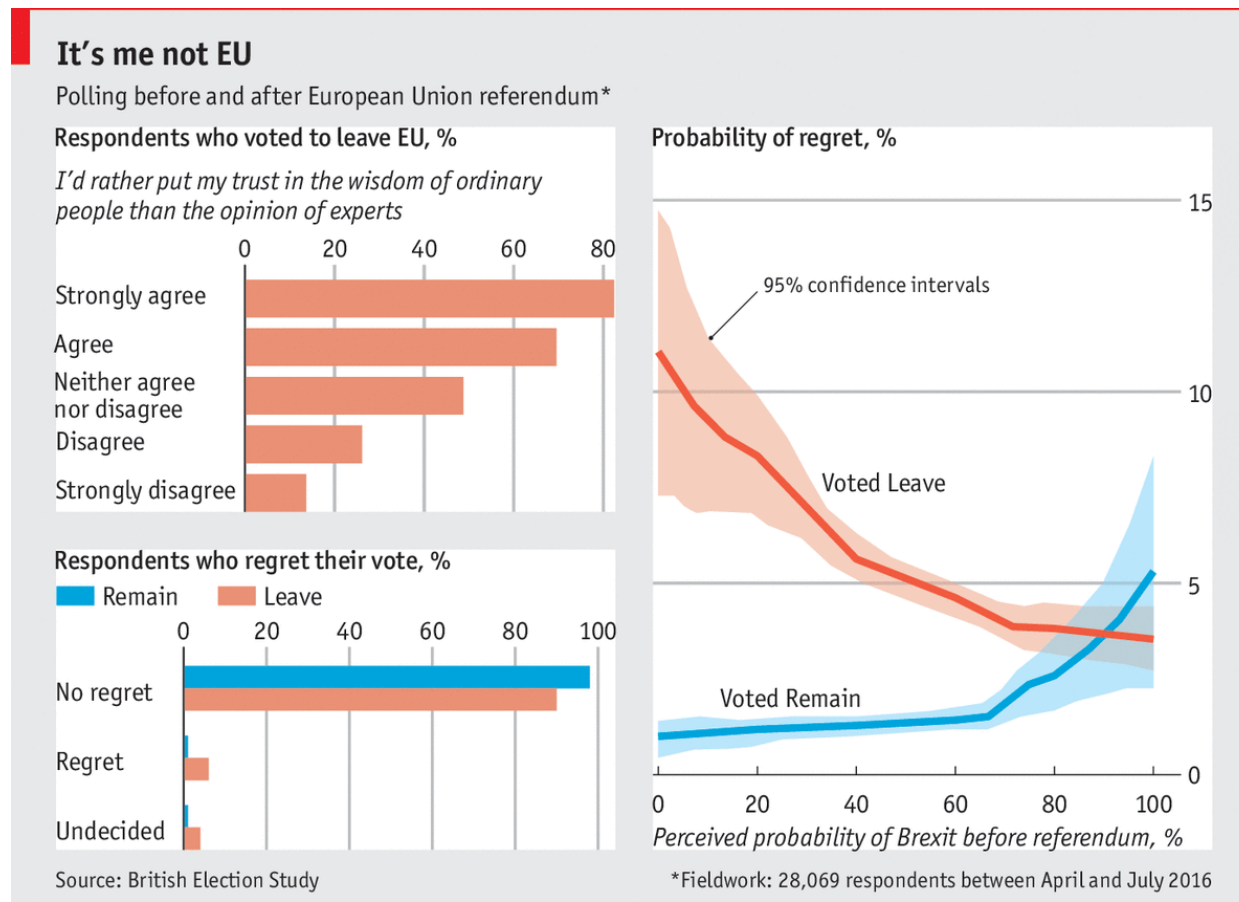
**Public Debt as Percentage of GDP**  
■ Less than 30% ■ 30% - 60% ■ 60% - 100% ■ 100% - Above

**How to read:** The more each citizen of a country owes, the closer to the center of the map the country is. Countries appear bigger as debt per capita to pay down countries' debt is higher, e.g. Japan. Conversely, countries where debt per capita is lower appear smaller in size and farther from the center, e.g. Yemen. The colour of the country indicates the public debt as a percentage of GDP.

Sources: <https://howmuch.net/articles/per-capita-debt>  
International Monetary Fund, World Bank and CIA World Factbook

howmuch.net  
visualcapitalist.com

# Netradiční zobrazení



Economist.com

# Netradiční zobrazení

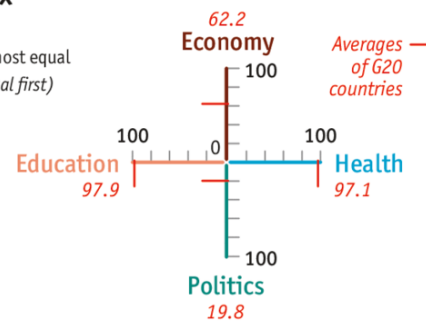


fivethirtyeight.com

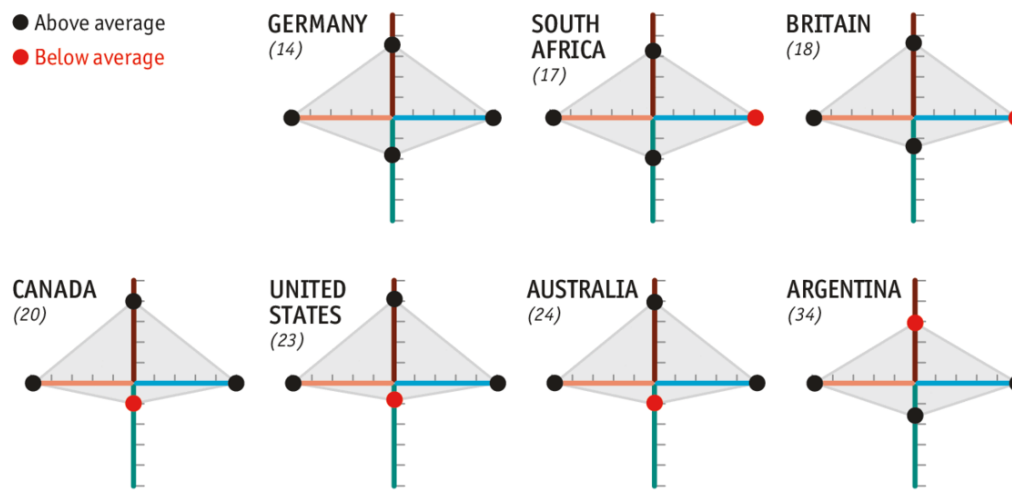
# Netradiční zobrazení

## Gender inequality index

G20\* countries, 2013  
Indexed across four measures, 100=most equal  
(Ranked by overall equality<sup>†</sup>, most equal first)



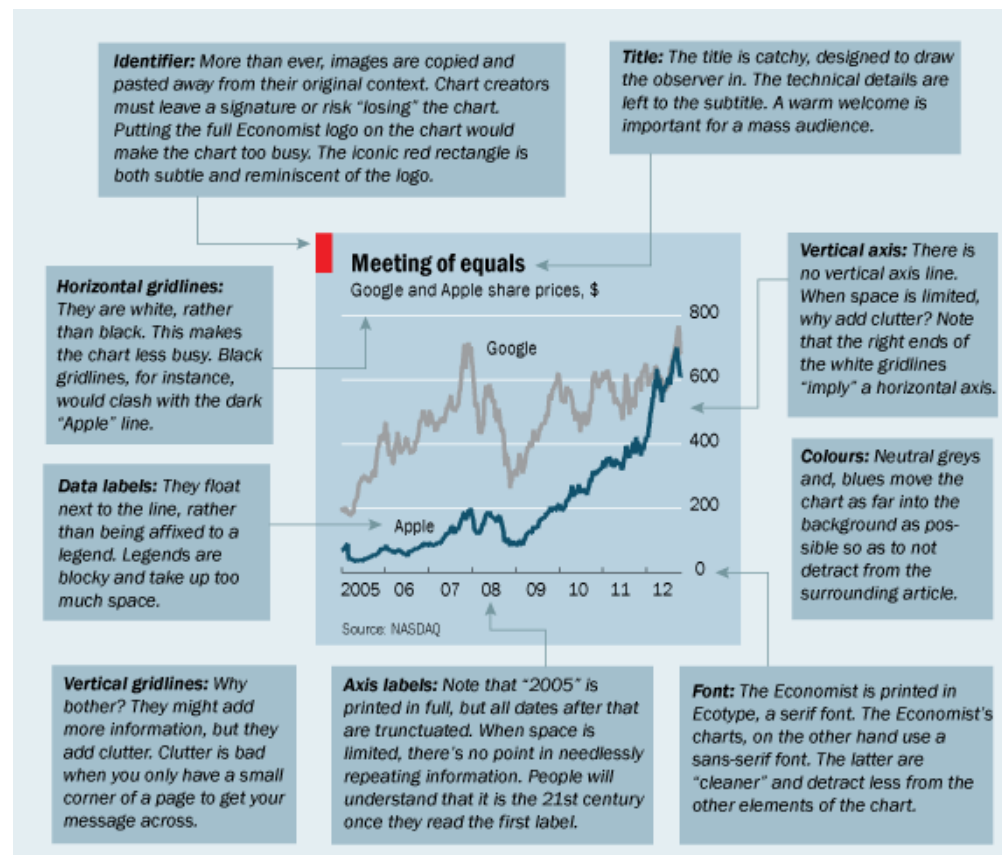
- Above average
- Below average



economist.com



# Na detailech záleží

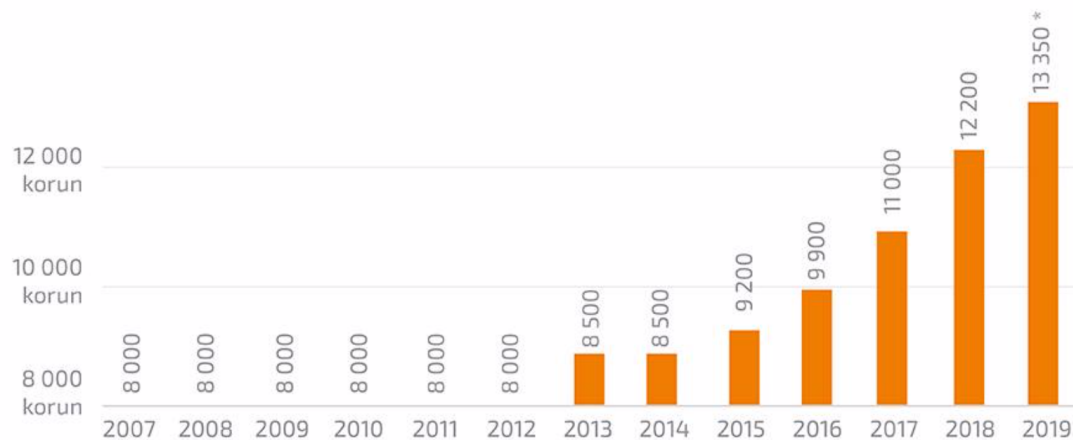


economist.com



# Grafické fauly

## Jak u nás rostla minimální mzda?



\* Od 1. 1. 2019 jsme vyjednali zvýšení minimální mzdy o 1 150 korun



ČSSD

Page Liked · 1 hr ·

Podařilo se nám vyjednat zvednutí minimální mzdy o 1 150 korun 👍 Proč jsme o její zvýšení tolik usilovali?

Zatímco ceny rok od roku stoupají, minimální mzda zůstávala dlouhé roky stejná. Nyní, když se naší zemi ekonomicky daří, máme nejlepší příležitost tuto nespravedlnost napravit a zajistit tak i lidem s nejnižšími příjmy důstojnou odměnu za jejich práci 📌

👍❤️👍 34

16 Comments 17 shares

👍 Like

💬 Comment

➦ Share

Most relevant ▾



**Josef Šedl** Zajímáte se vůbec o to, co se v této zemi nyní děje? Nebo se rozhodla ČSSD strčit hlavu do písku kvůli korytům?

Like · Reply · 1h

👍 9

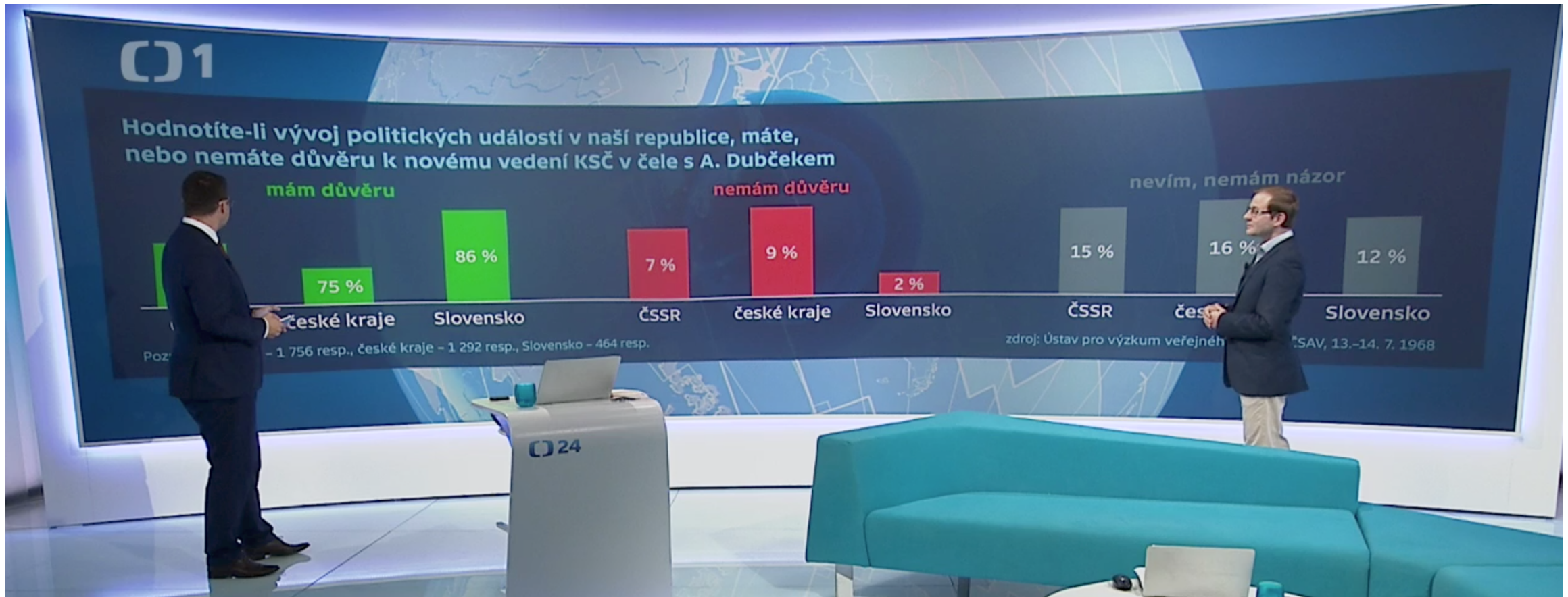


**Jan Hlaváček** Vyrábíte nebezpečnou past na nízkopříjmové zaměstnance, až dojde k ekonomické recesi.

Nejhorší je, že to píšu zbytečně, protože vy to dobře víte, nejspíše je to součástí vaší strategie.

facebook.com

# Grafické fauly



Česká televize

# Grafické fauly



mediahub.cz

# Grafické fauly



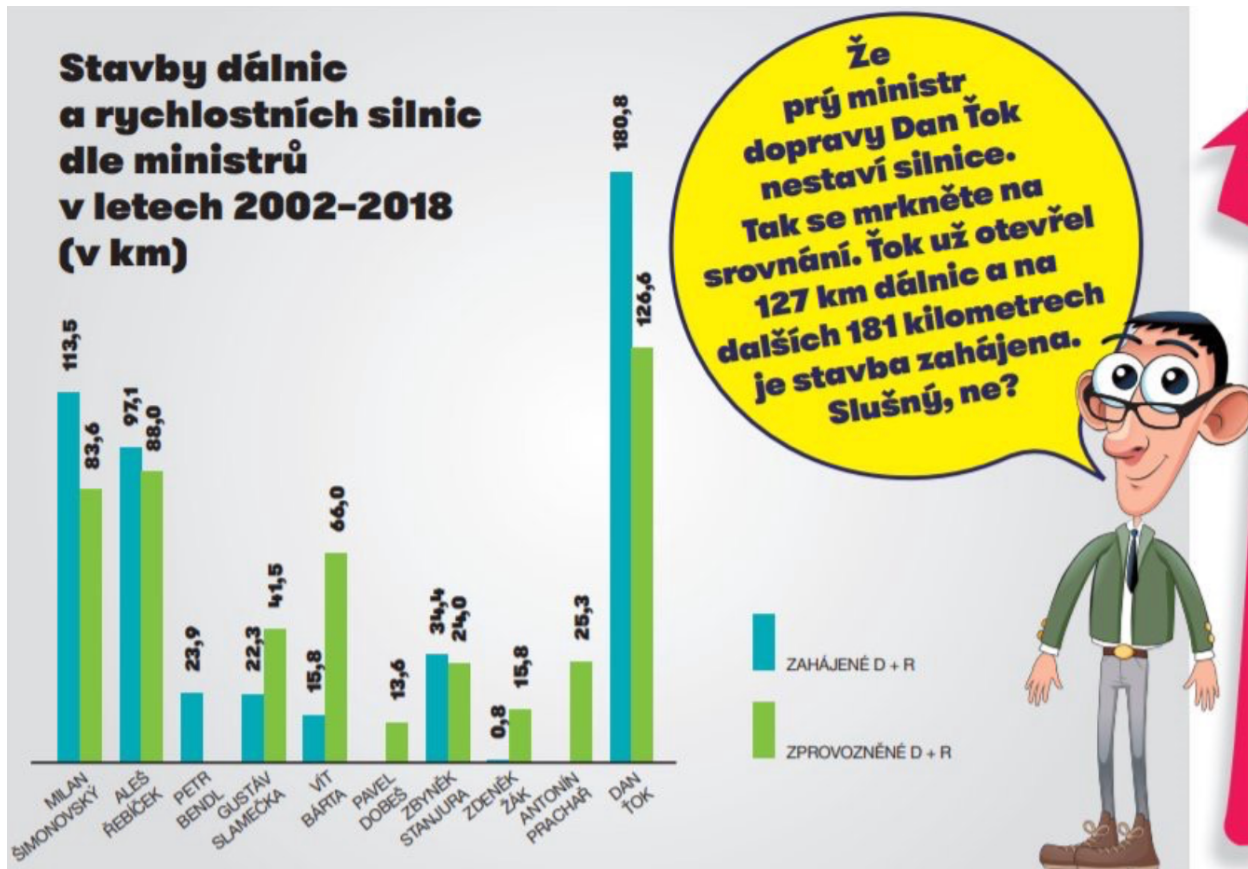
mediahub.cz



mediahub.cz

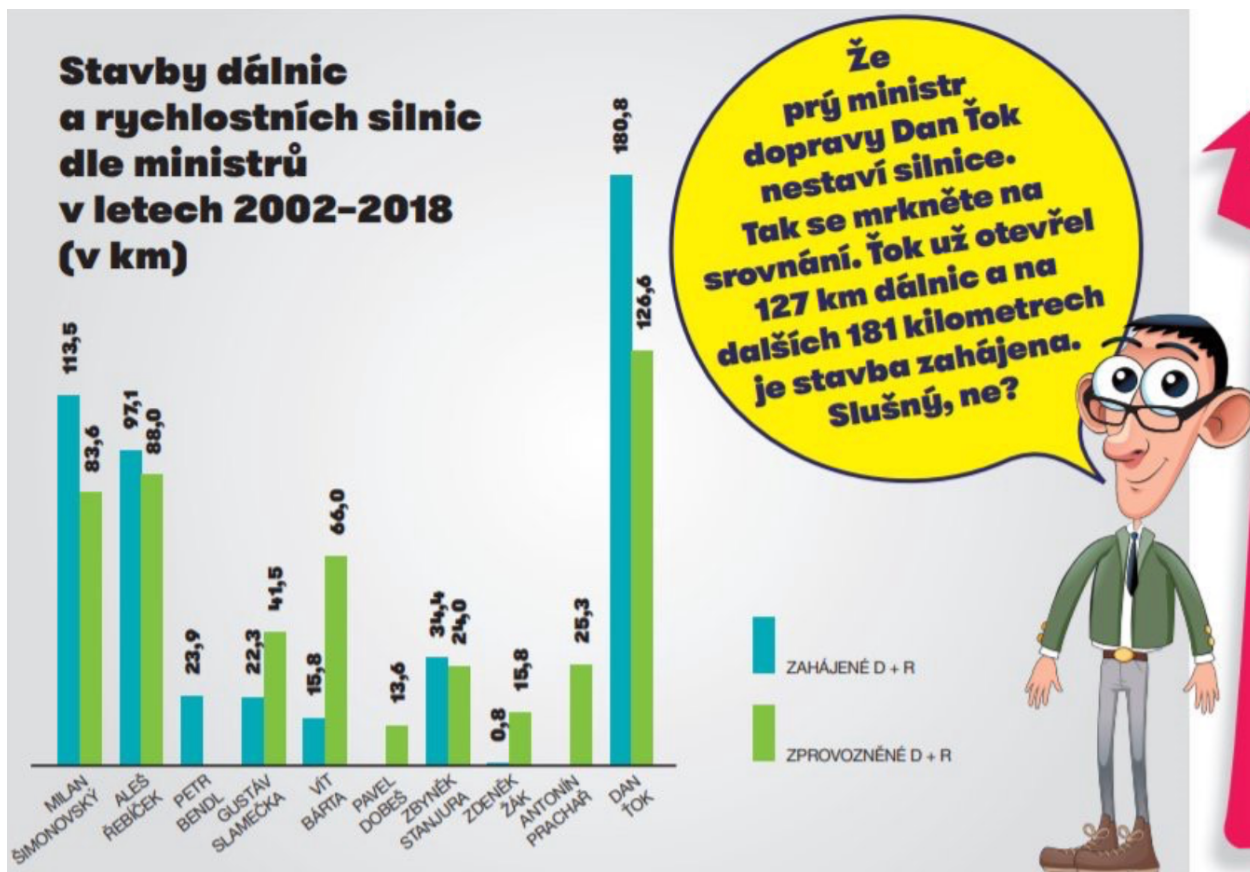


# Grafické fauly



medium.com

# Grafické fauly



medium.com

## Počet dokončených metrů za den

Pořadí	Jméno	metrů/den
1.	Vít Bárta	234,0
2.	Zbyněk Stanjura	114,3
3.	Aleš Řebíček	100,9
4.	Gustav Slamečka	96,3
5.	Zdeněk Žák	77,8
6.	Milan Šimonovský	55,3
7.	<b>Dan Ťok</b>	36,6
8.	Antonín Prachař	29,5
9.	Pavel Dobeš	26,1
10.	Petr Bendl	0,0
11.	Radek Šmerda	0,0

medium.com

# Grafické fauly

---

**Upraveno 7.1.2018 — šotek řádil, takže se musím omluvit za chybu v přepočtu otevřených metrů za den. Hloupou chybou jsem pohoršil Danu Ťokovi o dvě příčky, takže není nejhorší ani podle této metodiky. V textu nahrazeny oba grafy správnými čísly a věty, které chyba ovlivnila, jsou v hranatých závorkách []. Na ostatních datech a závěrech se nic nemění. Omlouvám se hluboce.**

medium.com

# Shrnutí

---

- existuje celá řada možností grafického zobrazení kvantitativních dat
- výběr vhodného způsobu zobrazení je zásadní součástí prezentace výsledků analýzy
- občas je vhodné sáhnout po tradičních nástrojích, ale nebát se inovace
- vždy myslete na to, pro jaké publikum data zobrazujete (a celou analýzu děláte)
- vyvarujte se jakýmkoliv grafickým faulům (podívejte se na <https://www.data-to-viz.com/caveats.html>)