

Hvordan og hvorfor modellerer og visualiserer man data

En CoTinker læringsaktivitet omkring datavisualisering



I CoTinker projektet på Institut for Datalogi ved Aarhus Universitet har vi udviklet en række læringsaktiviteter indenfor computational thinking målrettet faget informatik til gymnasieniveau. Den følgende læringsaktivitet omkring data visualisering fokuserer på at lære eleverne om, hvordan og hvorfor man modellerer og visualiserer data. Forløbet er tiltænkt til både informatik og matematik, men kan med fordel også anvendes i samfundsfag.

Læringsaktiviteten er bygget med CoTinker¹ platformen: En webbaseret platform til at udvikle læringsaktiviteter indenfor computational thinking med fokus på samarbejde og tinkering. Platformen gør det let for elever at samarbejde i grupper omkring kode, tekst og media via en browser på samme måde, som det kendes fra Google Docs. CoTinker kræver ingen installation af software på forhånd.

Følgende giver en kort introduktion til datavisualisering, en beskrivelse af læringsaktiviteten, en ordbog for fagbegreber i læringsaktivitet, forslag til en lektionsplan, samt tips, tricks og forslag til ekstramateriale, som kan være interessant at dykke ned i.

*Thomas Thomsen, Line Have Musaeus, Clemens Klokmosse Nylandsted
Institut for Datalogi, Aarhus Universitet, 2025*

¹ <https://cotinker.au.dk/frontpage/>

Indhold

Introduktion til datavisualisering.....	4
Hvad er data og hvad er datavisualisering?	4
Det visuelle alfabet	4
Hvordan fungerer datavisualisering?	5
Når visualiseringer går galt	6
Hvad kan vi bruge datavisualisering til?	6
Beskrivelse af CoTinker læringsaktiviteten om datavisualisering	7
Læringsmål	7
Beskrivelse af læringsaktiviteten	8
Ordbog over fagbegreber i CoTinker læringsaktiviteten	10
Forslag til lektionsplan	12
Ekstramateriale.....	14
Eksempler på manipulerende datavisualiseringer	16
Forskydning af akser	16
Omvendte akser	17
Dårlig brug af farver	19
Cherry picking af data	20

Introduktion til datavisualisering

Hvad er data og hvad er datavisualisering?

Data er information, der kan måles, tælles eller kategoriseres. Det kan være temperaturer, befolkningstal eller karakterfordeling på et gymnasie. I sig selv er data ofte svære at overskue. Lange rækker af tal fortæller sjældent en klar historie. Her kommer datavisualisering ind i billedet.

Datavisualisering er at gøre information synlig. Man kan tænke det som et sprog med sit eget alfabet, sin egen grammatik og sine egne regler. Ligesom vi bruger ord til at udtrykke tanker, bruger vi **visuelle elementer** som farver, former og positioner til at udtrykke data. Når det gøres rigtigt, hjælper visualisering os med at se mønstre, forstå sammenhænge og opdage historier i tallene.

Visualisering fungerer som et tegnsystem, hvor forskellige visuelle elementer har forskellige betydninger. En position på en akse betyder noget. En farve betyder noget. En størrelse betyder noget. Sammen skaber disse elementer en visualisering, der kan formidle information langt mere effektivt end tabeller med rå tal. Man kan sige, at data er til computeren og visualiseringer er til mennesker.

Det visuelle alfabet

Når vi skaber datavisualiseringer, arbejder vi med nogle grundlæggende visuelle egenskaber. Hver af disse kommunikerer på sin egen måde, og nogle er mere effektive end andre. **Position** er den mest præcise visuelle egenskab. Når vi placerer datapunkter langs en akse, som i et søjlediagram eller et punktdiagram, kan vores øjne sammenligne værdierne med stor nøjagtighed. Mennesker kan læse position langt mere præcist end andre visuelle egenskaber og bruges derfor ofte.

Længde og **højde** fungerer næsten lige så godt, hvilket søjlediagrammer og linjediagrammer bruger. Vi kan let se forskelle i søjlehøjder eller følge en linjes bevægelse over tid. Her skal man dog være opmærksom på aksernes skala. En akse der ikke starter ved nul, kan overdrive små forskelle. **Størrelse** og **areal** er langt sværere at vurdere præcist. Selvom vi intuitivt forstår, at en større cirkel repræsenterer en større værdi, er vi ikke særlig gode til at sammenligne arealer. En cirkel med dobbelt så stor radius har faktisk fire gange så stort areal (For $r=2$ er areal $= \pi 2^2 \approx 12,57$ og for $r=4$ er areal $= \pi 4^2 \approx 50,26$). Den fejl overses let.

Farve er et komplekst værktøj der bruges forskelligt afhængigt af datatypen. Til **kategoriske data**, hvor vi skal skelne mellem forskellige typer, fungerer forskellige farvetoner godt. Rød for æbler, grøn for pærer, gul for bananer. Til **numeriske data**, hvor vi viser mængder, bruges gradienter. En gradient er to eller flere farver, der glider jævnt over i hinanden. Det kan være en overgang fra lys til mørk, eller det kan være en overgang mellem to farver. Begge viser et forløb fra lidt til meget. Til **ordinale data**, hvor værdierne har en rækkefølge, bruges farvetrin der både viser denne orden og adskiller hver kategori tydelig. Farverne flyder ikke sammen som i en gradient. Det kan f.eks. være trin fra lys grøn til mellemgrøn til mørkegrøn for at markere kategorier.

Former, mønstre og **vinkler** spiller også en rolle. Forskellige figurtyper kan repræsentere forskellige kategorier. Cirkler for en type, trekanten for en anden. Vinkler bruges i cirkeldiagrammer, selvom de er sværere at læse præcist end længde eller position.

Hvordan fungerer datavisualisering?

Datavisualisering udnytter måden vores hjerne bearbejder **visuel information**. Meget af arbejdet sker før vi overhovedet tænker bevidst over det. Visse mønstre springer øjeblikkeligt i øjnene. Andre kræver mere bevidst analyse.

Der findes et hierarki for hvordan præcist vi kan aflæse forskellige visuelle elementer. Position langs fælles akse er mest præcis. Derefter kommer position langs forskellige akser, så længde, vinkel og hældning, derefter areal, og sidst størrelse og farvetone. De vigtigste variabler bør kodes med de mest præcise visuelle egenskaber.

Der er tre centrale kompetencer i datavisualisering. For det første **fortolkning**: Evnen til at læse og afkode visuelle elementer, forstå hvad der vises, og identificere mønstre. For det andet **konstruktion**: Evnen til selv at skabe visualiseringer gennem bevidste valg af diagramtyper, og koblinger mellem data og visuelle egenskaber. For det tredje **kommunikation**: Evnen til at bruge visualiseringer til at argumentere, fortælle historier og dele indsigter med andre.

Visualiseringer kan bruges til forskellige ting. Visualiseringer kan bruges, når vi selv undersøger data for at finde mønstre, vi ikke kendte i forvejen eller når vi stiller spørgsmål til data. Et punktdiagram kan afsløre sammenhænge mellem to variabler. Et linjediagram kan vise udvikling over tid. Et søjlediagram kan sammenligne værdier på tværs af kategorier. Visualiseringen gør det muligt at se både generelle tendenser og interessante afvigelser.

Visualiseringer kan også bruges, når vi skal kommunikere mønstre eller fund til andre. En visualisering kan formidle information hurtigere og tydeligere end tekst eller tabeller. Den kan fremhæve det vigtige og nedtone det uvigtige for at fortælle en historie, som modtageren kan forstå. Det kræver ofte forskellige designvalg. En visualisering til egen udforskning kan være kompleks og detaljeret. En visualisering til kommunikation skal være fokuseret og letforståelig.

Gode visualiseringer lever op til flere kriterier. De skal være sandfærdige. De viser data ærligt uden at forvrænge eller manipulere. De skal være funktionelle. De tjener det formål de er skabt til. De skal helst være flotte. Æstetik hjælper med at fastholde opmærksomhed og gøre information tilgængelig. De skal være indsigtsfulde. De afslører noget vi ikke så før. Ideelt set skal de være oplysende. De ændrer hvordan vi forstår verden.

Når visualiseringer går galt

Visualiseringer er aldrig neutrale. Hvert **designvalg** former hvordan vi forstår data. Små ændringer kan skabe forskellige indtryk. Nogle gange sker det ved uheld. Andre gange er det bevidst **manipulation**. Almindelige fejl inkluderer **forskudte akser** der overdriver små forskelle. **Dobbelte akser** skaber forvirrende sammenligninger. Unødvendige **3D-effekter** forvrænger proportioner. Cirkeldiagrammer med for mange kategorier bliver uoverskuelige. Visse farvekombinationer er ulæselige for **farveblinde**. Mere problematisk er bevidst manipulation. **Cherry-picking** betyder, at man kun vælger de data, der understøtter ens historie. Det kan skabe et stærkt forvrænget billede.

Det gør **kritisk læsning** vigtigt. Når vi møder en visualisering, skal vi stille spørgsmål: Starter akserne ved nul? Hvad er udeladt fra visningen? Hvem har lavet visualiseringen, og hvilken historie vil de fortælle? Hvilke designelementer er der brugt og hvordan? Det handler ikke kun om at kunne skabe visualiseringer, men også om at kunne læse dem kritisk.

Hvad kan vi bruge datavisualisering til?

Datavisualisering er altså mere end pæne grafer. Det er et **systematisk sprog** med sit eget alfabet af visuelle elementer, sine egne regler for hvordan de kombineres, og sine egne principper for god praksis. At lære dette sprog kræver forståelse for både hvordan mennesker læser visuel information og hvordan designvalg former fortolkning. Det kræver også et kritisk blik for at kunne skelne ærlig formidling fra manipulation. Men med øvelse bliver datavisualisering et effektivt værktøj til både at forstå komplekse og udforske data for at opdage sammenhænge og mønstre, der er svære at se i rå tal.

Beskrivelse af CoTinker læringsaktiviteten om datavisualisering

Læringsaktiviteten består af 18 steps, som på forskellig vis indeholder noget tekst og teori, opgaver, dataset i tabelform og et blokbaseret værktøj til at lave datavisualiseringer. Eleverne arbejder sammen i grupper, hvor grupperne i eget tempo arbejder sig igennem de forskellige steps.

Som lærer har du adgang til et 'Teacher view' (i form af link til en hjemmeside), hvor du kan oprette gruppelinks til at dele med eleverne, se hvor langt eleverne er kommet ud af de 18 steps, og åbne et gruppelink for at dele elevernes arbejde oppe på en projektor eller for at hjælpe eleverne med at løse en opgave.

Læringsaktiviteten bruger et datasæt fra Spotify, som indeholder 50000 sange med informationer omkring kunstner, titel, album, genre, udgivelsesår samt en lang række variabler omkring en sangs karakteristika. Disse variabler er dels kendt data omkring sange (varighed, tempo) samt en lang række variabler af en mere anderledes karakter, som Spotify selv har udregnet for hver sang (dansevenlighed, energi og positivitet).

Læringsaktiviteten bruger en delmængde af det samlede datasæt for de forskellige slides, hvor sangene er udvalgt efter bestemte parametre. Fokus har været på at datasættet passer til opgaven, som eleverne skal løse. Delmængderne har altid 50 sange, der skifter undervejs.

Læringsaktiviteten kan godt bredes ud over flere dage. Gruppernes arbejde bliver hele tiden gemt, og de skal blot åbne deres gruppelinks igen for at fortsætte arbejdet. Du skal lærer også bare gemme linket til dit teacher view, hvor du også kan distribuere gruppelinks'ene igen, hvis det er nødvendigt.

Læringsmål

Forløbet har fokus på tre overordnede læringsmål:

- **Fortolkning:** Eleverne udvikler evnen til at læse, analysere og forstå eksisterende visuelle repræsentationer af data.
- **Konstruktion:** Eleverne lærer at konstruere visualiseringer ud fra data, hvor de skal træffe beslutning diagramtype, variabler der skal vises, visuelle egenskaber og koblingen af disse med variabler.
- **Kommunikation:** Eleverne forstår, hvordan visualiseringer bruges til at kommunikere budskaber, fortælle historier og understøtte argumenter.

Beskrivelse af læringsaktiviteten

I de første steps præsenteres eleverne for et overordnet 'Driving question' og de skal reflektere over, hvilke data man kan indsamle om sange. Formålet er at få eleverne gang i med at måde at tænke på musik som data.

De næste steps introducerer centrale begreber omkring data, datatyper og strukturering. Eleverne skal diskutere mønstre i en tabel med sange for at opbygge en grundlæggende forståelse af, hvordan grundkomponenterne (data) i datavisualiseringer er bygget op, før de selv skal i gang med at lave visualiseringer i de efterfølgende steps. Disse indledende steps lægger grundlaget for **fortolkning** ved at etablere en forståelse af, hvad data er, og hvordan det er struktureret.

Derefter introduceres det blok-baserede værktøj med faste blokkonfigurationer. De tilgængelige blokke inkluderer figurtype (søjlediagram, punktdiagram, linjediagram), variabelvalg til x- og y-akser, farve samt gennemsigthed. I disse steps udvikler eleverne deres evne til **konstruktion** ved selv at bygge visualiseringer, samtidig med at de udvikler kompetencer i **fortolkning** ved at forstå, hvordan visuelle egenskaber kommunikerer information.

De følgende steps udvikler elevernes evne til at bygge visualiseringer med det blok-baserede værktøj (herfra ikke i en fast konfiguration), samtidig med at de skal lære at **læse** og **fortolke** visualiseringer ved at forstå, hvordan visuelle egenskaber kommunikerer information. Eleverne skal lave et søjlediagram, et linjediagram og et punktdiagram, hvor de skal prøve at koble visuelle egenskaber (som f.eks. farve) til variablerne fra datasættet. Eleverne skal også sammenligne og reflektere på tværs af alle tre diagramtyper og koblingen af forskellige visuelle egenskaber. Denne del styrker både **konstruktion** igennem skabelse af visualiseringer og **fortolkning** gennem sammenligning på tværs af visualiseringstyper.

De sidste steps beder eleverne om at skabe visualiseringer målrettet specifikke målgrupper og reflektere over designvalg med fokus på **kommunikation**. Eleverne udforsker manipulationsteknikker (akseskift, datafiltrering, dramatiske farver, vildledende titler, selektive beskrivelser) og reflekterer over deres effektivitet, hvordan de påvirker vores opfattelse, og etiske gråzoner, hvor tilpasning af visualiseringer kan være acceptabelt kontra eksplicit manipulativt. Dette berører hvordan manipulation påvirker **fortolkning** og hvad man som modtager, skal være opmærksom på.

Det allersidste step vender tilbage til læringsaktivitetens driving question og beder eleverne reflektere over dette for at binde læringen på tværs af **fortolkning**, **konstruktion** og **kommunikation**. Det sidste step inkluderer også mulighed for at eksportere alt arbejdet i PDF-format, hvilket inkluderer alle besvarelser, noter og visualiseringer.

Ordbog over fagbegreber i CoTinker læringsaktiviteten

Følgende er en oversigt og definition af fagbegreber inden for datavisualisering, som optræder i CoTinker læringsaktiviteten:

Data

Data er oplysninger, informationer eller kendsgerninger, som vi kan samle ved at observere og måle verdenen. Inden for datavisualisering er det oplysninger eller kendsgerninger, som kan behandles af computeren, som f.eks. tal.

Datapunkt

Et enkelt stykke data der har en værdi indenfor alle variabler i et datasæt. Oftest er det en række i en tabel for et datasæt.

Datasæt

Et datasæt er en organiseret samling af oplysninger, informationer eller kendsgerning. Et datasæt består en eller flere variabler med en eller flere datapunkter.

Datavisualisering

En datavisualisering er en visuel repræsentation af data, der bruger forskellige diagramtyper, farver og andre visuelle elementer til at vise data på en let og overskuelig måde.

Diagramtype

Forskellige diagrammer bruger forskellige figurer til at repræsentere data på hver deres måde, som f.eks. punkter, linjer eller søljer.

Kategoriske data

Data som kan deles op i tydelige grupper, hvor ingen gruppe er mere, mindre, før eller efter andre grupper. Dette kan f.eks. være genre, som rock, pop eller jazz.

Kobling

Kobling i datavisualisering er processen med at oversætte eller koble data til visuelle egenskaber for at gøre det lettere at kommunikere og fortolke datasættet.

Linjediagram

En diagramtype der viser data som en række punkter (datapunkterne) forbundet af linjer, der viser hvordan to variabler udvikler sig på samme tid – oftest en udvikling over tid.

Numeriske data

Data som består udelukkende af tal, som ikke er inddelt i grupper. Det kan f.eks. være tempo for en sang (ofte målt i BPM – beats pr. minute).

Ordinale data

Ordinale data er en type kategoriske data, der har en naturlig rangordning eller rækkefølge men hvor forskellene mellem kategorierne ikke er ens eller kan måles, som f.eks. alfabetet.

Punktdiagram

En diagramtype der bruger punkter til at vise sammenhængen mellem to variabler hvor den ene variabel vises på x-aksen og den anden variabel vises på y-aksen. Hvert punkt i diagrammet repræsenterer et datapunkt med de på gældende variabler.

Søjlediagram

En diagramtype der bruger søljer og længden/højden af disse til at vise mængder indenfor kategorier, hvor hver sølje repræsenterer en kategori.

Variabel

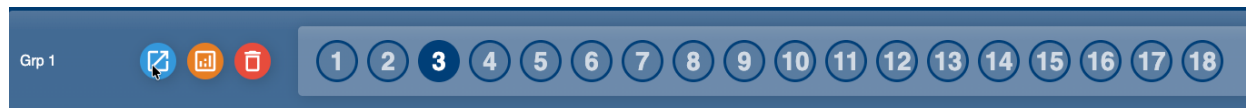
Data omkring en specifik ting (f.eks. genre eller årstal), der ofte vises som en kolonne i en tabel for et datasæt.

Visuel egenskab

I en datavisualisering kan vi bruge forskellige visuelle egenskaber (position, farve, størrelse osv.) til at variere, hvordan data bliver konstrueret, kommunikeret og fortolket.

Forslag til lektionsplan

Følgende er et forslag til en lektionsplan for læringsaktiviteten. Hvis der ikke laves diskussioner på klassen, som angivet i lektionsplanen, tager læringsaktiviteten 90 min. Hvis alle diskussioner på klassen inkluderes, tager læringsaktiviteten 120 min. Husk at du kan åbne en gruppes link for at vise deres arbejde på en projektor for hele klassen ved at klikke på den blå knap ud for en gruppes navn i teacher view.



Figur 1: Den blå knap som åbner en gruppes link på en ny fane.

Læringsaktiviteten kan breddes ud over flere dage. Gruppernes arbejde bliver hele tiden gemt, og de skal blot åbne deres gruppelinks igen for at fortsætte arbejdet. Du skal som lærer også bare gemme linket til dit teacher view, hvor du kan distribuere gruppernes link igen. På side 16 kan der ses eksempler på manipulerende datavisualiseringer, som kan bruges i undervisningen.

Diskussion 1 (Step 6):

Kontekst: Eleverne arbejder med data, datatyper og hvordan musik kan forstås som data

Diskussionsspørgsmål: "Hvad er egentlig data og hvornår bliver noget til data?"

Ekstra spørgsmål: "Hvilke ting ved en sang kunne I forestille jer at måle eller tælle?" og "Hvor kommer data fra?"

Diskussion 2 (Step 13):

Kontekst: Eleverne har lavet tre visualiseringstyper (søjlediagram, punktdiagram, linjediagram) og reflekteret over koblinger mellem variabler og visuelle egenskaber.

Diskussionsspørgsmål: "Hvornår er en diagramtype bedre end en anden? Og for hvem?"

Ekstra spørgsmål: "Var der noget, I kunne se i ét diagram, som var svært at opdage i et andet?"

Diskussion 3 (Step 17)

Kontekst: Eleverne har arbejdet med bevidst manipulerende visualiseringer og reflekteret over hvilke teknikker der virkede bedst.

Diskussionsspørgsmål: "Hvornår går en visualisering fra at være 'tilpasset' til at være 'manipulerende'?"

Ekstra spørgsmål: "Hvad skal I være opmærksom på, næste gang I ser en datavisualisering i nyhederne eller på sociale medier?"

MINUT	STEP	AKTIVITET
0-5	Ingen	Aktiviteten præsenteres, eleverne deles ud i grupper og links til grupper distribueres.
6-15	Step 1-3	Eleverne kommer ind i CoTinker for første gang, introduceres til platformen og aktiviteten, og ser to eksempler på visualiseringer (step 2). Elevernes refleksioner over de to visualiseringer kan med fordel samles op på klassen.
16-25	Step 4-6	Eleverne arbejder i grupper med at lære om data, datatyper og hvordan musik kan forstås om data
26-35	Step 6 [VALGFRI]	Diskussion 1 (fælles): <i>"Hvad er egentlig data – og hvornår bliver noget til data?"</i>
36-40	Step 7	Eleverne prøver at lave deres første datavisualisering i CoTinker med blokværktøjet
41-50	Step 8-9	Eleverne lærer om søljediagrammer og farver, mens de selv laver et søljediagram om genrefordeling i datasættet.
51-60	Step 10-11	Eleverne lærer om punktdiagrammer og farver heri, mens de selv laver et punktdiagram til at vise sammenhængen mellem to variabler.
61-65	Step 12	Eleverne lærer om linjediagrammer og hvordan det kan bruges til tidsserier, mens de selv laver et linjediagram for tempoudvikling i to genre.
66-75	Step 13	Eleverne bliver præsenteret for de tre visualiseringer, de har lavet. De skal reflektere over deres fordele og ulemper
76-85	Step 13 [VALGFRI]	Diskussion 2 (fælles): <i>"Hvornår er en visualiseringstype (sølje, punkt eller linje) bedre end en anden? Og for hvem?"</i>
86-95	Step 14-15	Eleverne arbejder med at bruge datavisualiseringer til historiefortælling for specifikke målgrupper.
96-105	Step 16-17	Eleverne arbejder med at lave manipulerende datavisualiseringer og reflektere over, hvilke teknikker der virkede bedst.
106-115	Step 17 [VALGFRI]	Diskussion 3 (fælles): <i>"Hvornår går en visualisering fra at være 'tilpasset' til at være 'manipulerende'?"</i>
116-120	Step 18	Eleverne skal reflektere over deres arbejde i forhold til driving question fra step 2. Eleverne skal til sidst eksportere alt deres arbejde (skriftlige besvarelser og visualiseringer) til en PDF.

Ekstramateriale

Følgende er ekstramateriale, som enten kan inddrages i undervisningen eller bruges som baggrundsviden i forberedelse:

Organ, Nancy. *Data visualization for people of all ages*. AK Peters/CRC Press, 2024.

En god grundlæggende introduktion til datavisualisering målrettet begyndere. Den har et meget tilgængeligt sprog og er letforståelige. Bogen indeholder også en række analoge øvelser, som kan inddrages i undervisningen.

Tufte, Edward R. *The Visual Display of Quantitative Information*. 2. ed., Graphics Press, 2001.

En mere dybdegående bog der går mere teoretisk til værks i sin gennemgang af datavisualisering og især i forhold til designprincipper gode data visualisering. Bogen regnes af mange for et grundværk inden for litteratur i data visualisering.

The best stats you've ever seen – Hans Rosling TED Talk, 2006

En efterhånden meget berømt TED Talk af den svenske læge Hans Rosling, som på underholdende vis (og ved hjælp af forskellige visualiseringer) fortæller om globale tendenser inden for sundhed og økonomi. Videoen kan være en god første introduktion til datavisualiseringer for eleverne.

Link: https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen?referrer=playlist-the_best_hans_rosling_talks_yo&autoplay=true

Our World in Data

Hjemmeside af non-profit organisationen Our World in Data som indeholder en masse open source datavisualiseringer inden for emner som sundhed, befolkning, klima mm. Hjemmesiden kan bruges til at vise eksempler på datavisualiseringer og til at starte diskussioner omkring, hvad datavisualiseringer kan bruges til.

Link: <https://ourworldindata.org/data-insights>

Det Fælles Designsystem - Digitaliseringsstyrelsen & Erhvervsstyrelsen

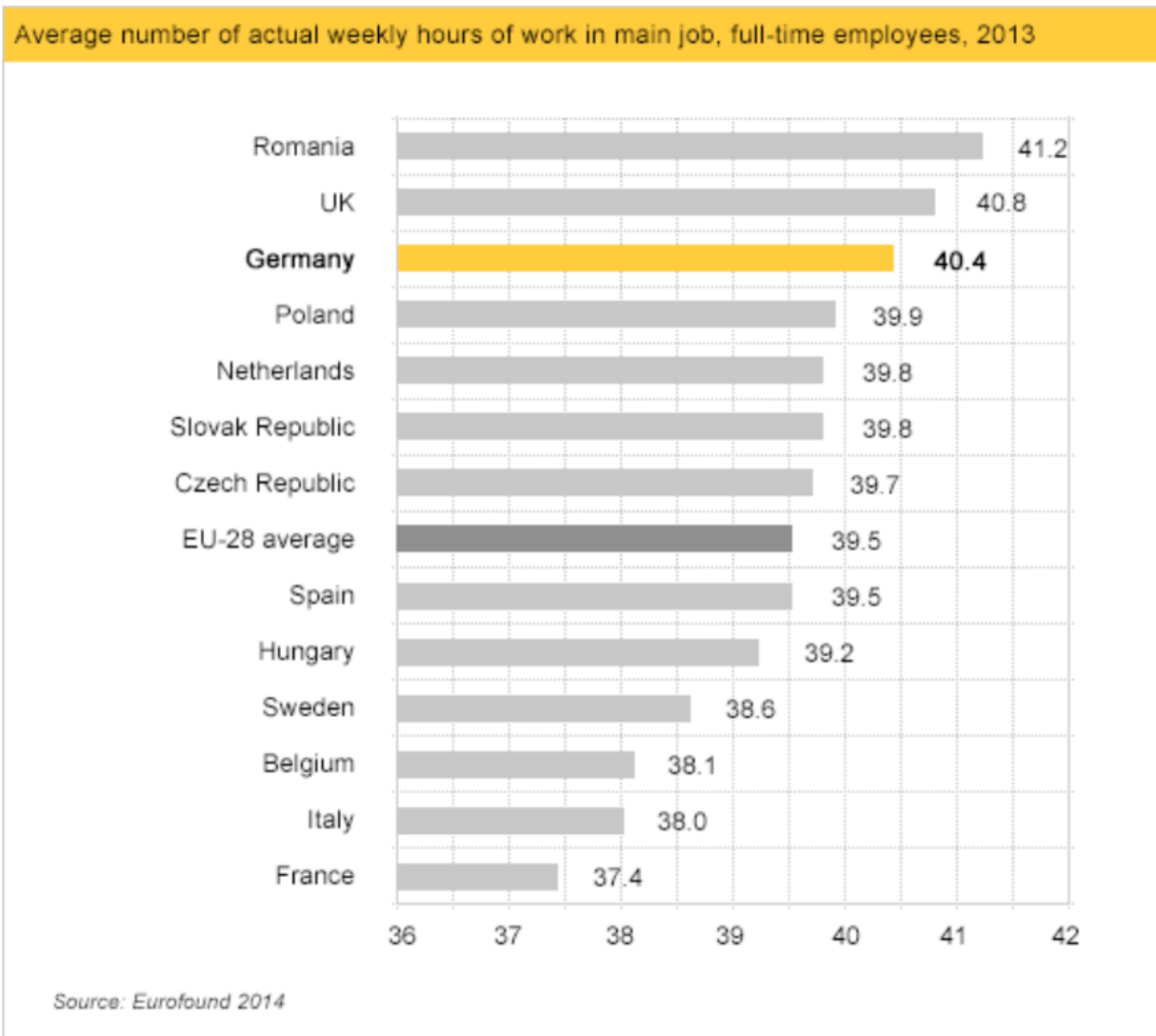
En designguide fra Digitaliseringsstyrelsen & Erhvervsstyrelsen til at lave datavisualiseringer, som er tilgængelige. Guiden kan være en god introduktion til, hvordan man designer tilgængelige datavisualiseringer i forhold til brugen af visuelle elementer og forskellige designteknikker.

Link: <https://designsystem.dk/styleguide/datavisualisering-infografik/tilgaengelig-grafik/>

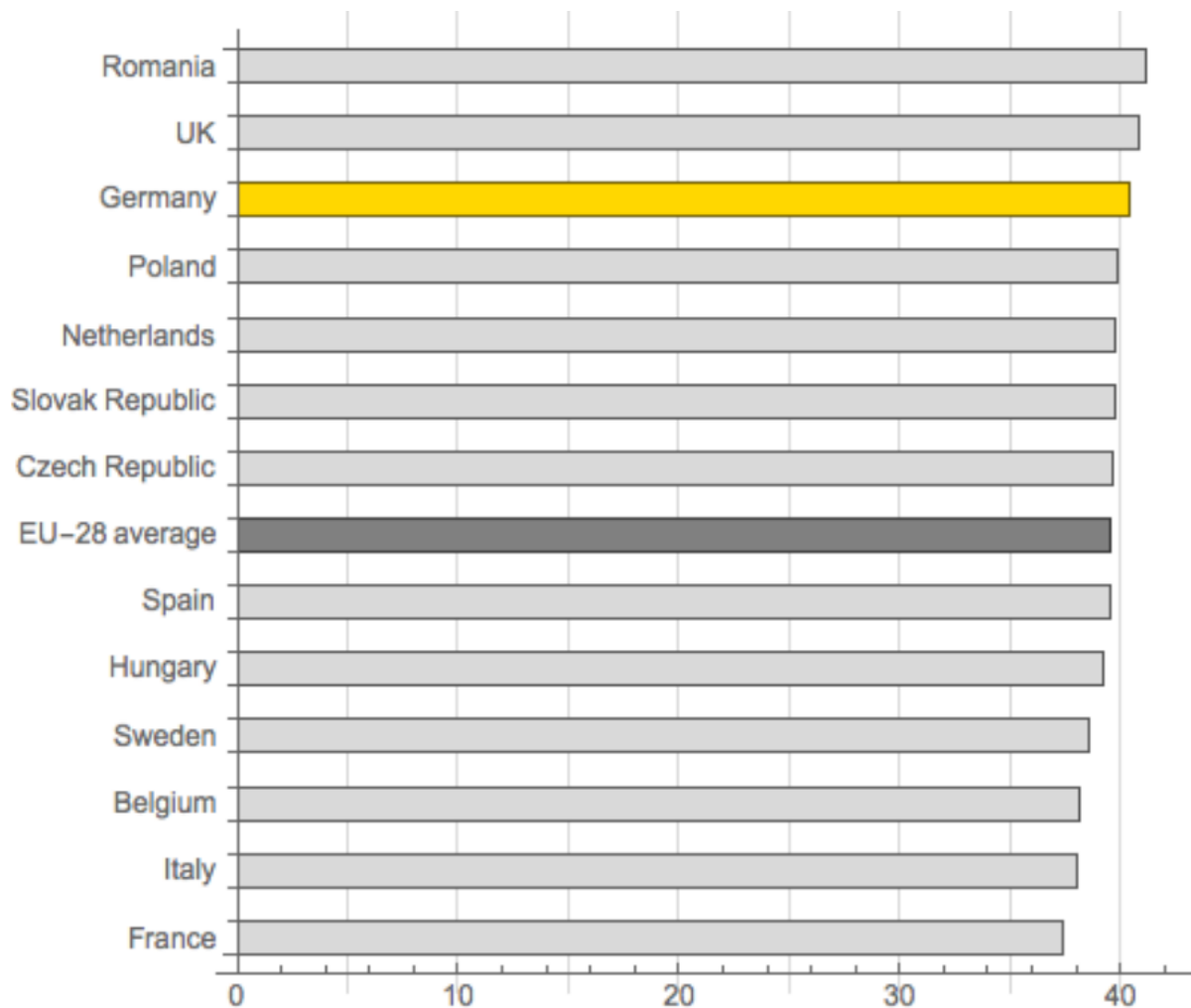
Eksempler på manipulerende datavisualiseringer

Forskydning af akser

Forskydning af akser er en hyppig årsag til fejllæsning af visualiseringer og er en anvendt teknik til at manipulere. Se på nedstående visualisering over gennemsnitlig antal arbejdstimer pr. uge for fuldtidsjob fordelt på land.



Ved første blik ser det ud til, at Tyskland arbejder væsentlige mere end lande som Sverige, Belgien og Italien, og næsten tre gange så meget som Frankrig! Akserne starter dog ikke ved nul, hvilket forvrænger billedet. Hvis akserne i stedet starter ved nul, får vi:



Tyskland arbejder stadig mere i gennemsnit end Sverige, Belgien, Italien og Frankrig, men forskellen er ikke helt så dramatisk, som den første visualisering.

NB: Denne visualisering bruges på step 2 i læringsaktiviteten.

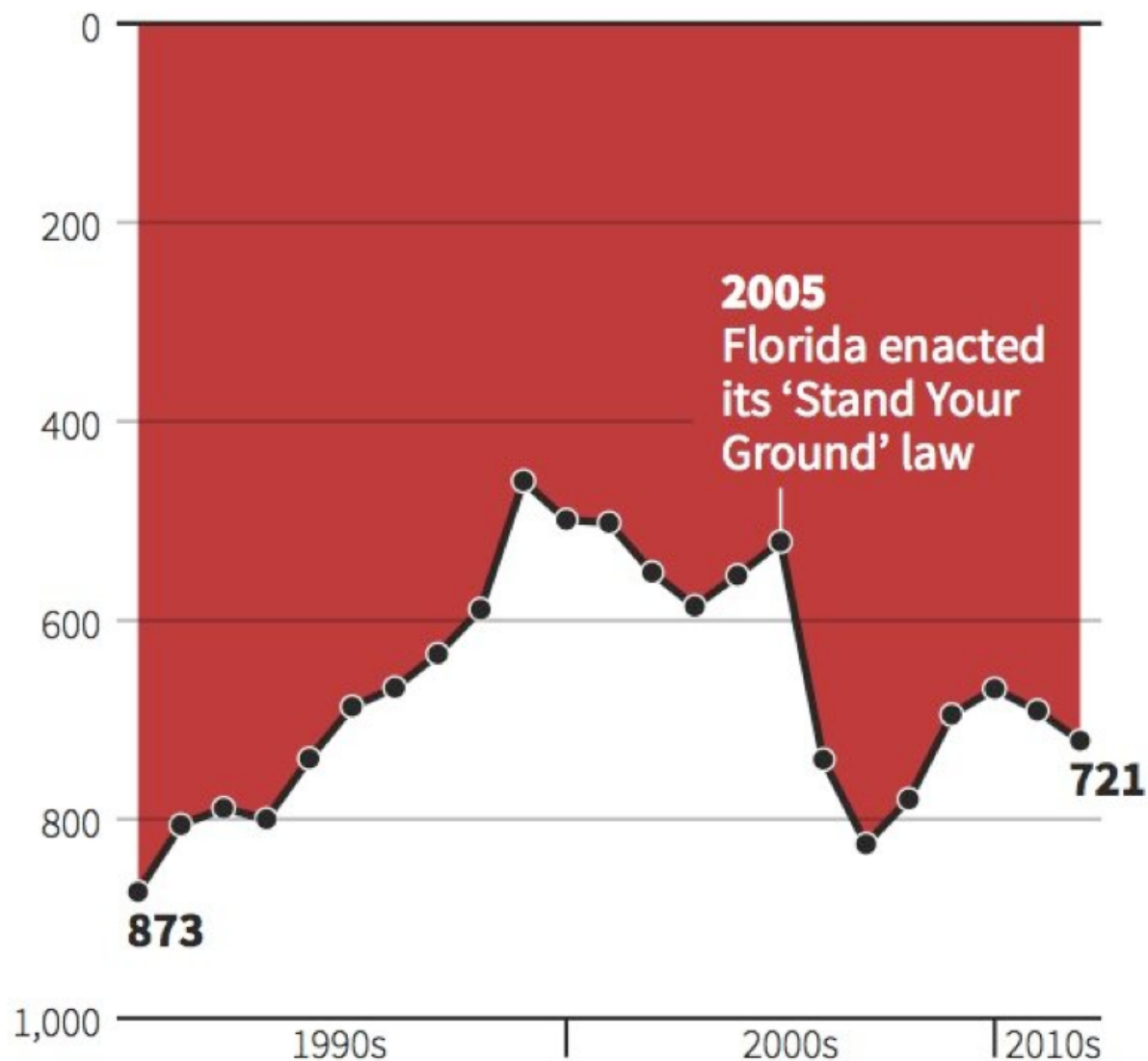
Kilde: https://www.callingbullshit.org/tools/tools_misleading_axes.html

Omvendte akser

Normalvis går akserne fra den mindste værdi (typisk nul som vi så ovenfor) til den højeste værdi. Nogen vender akserne om, hvilket skaber ikke bare forvirring, da vi som læser forventer noget andet, men det kan også være direkte manipulerende! Se på denne graf over antal skuddøde i Florida efter en ny våbenlov blev vedtaget i 2005:

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

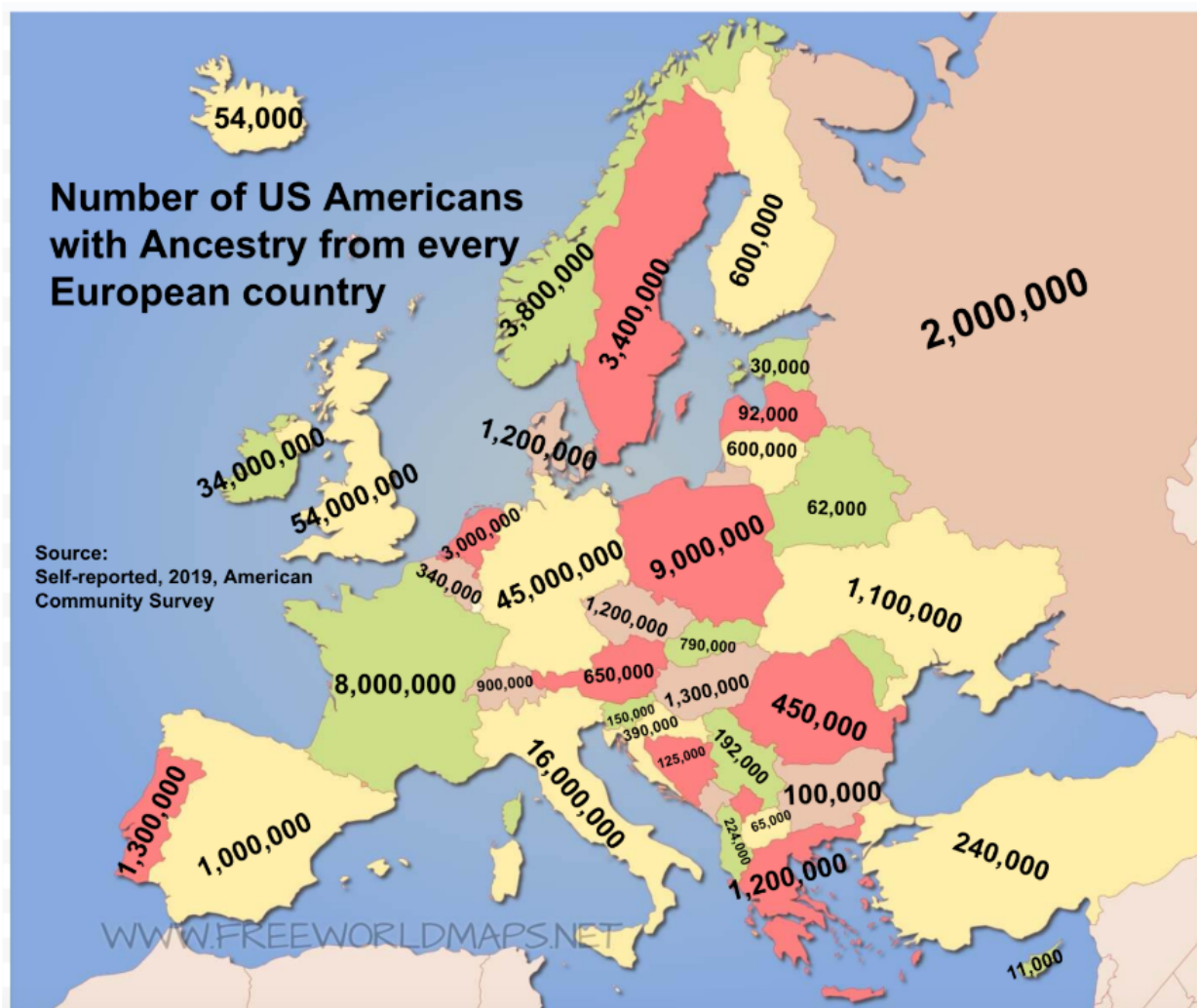
Umiddelbart ligner det, at den nye lov fik antallet af skuddøde til at falde, hvilket jo er fantastiske nyheder og en politisk succes. Hvis man dog dobbelttjekker akserne, tegner der ser et andet billede: Antallet af skuddøde steg rent faktisk efter loven blev vedtaget!

Kilde: https://www.callingbullshit.org/tools/tools_misleading_axes.html

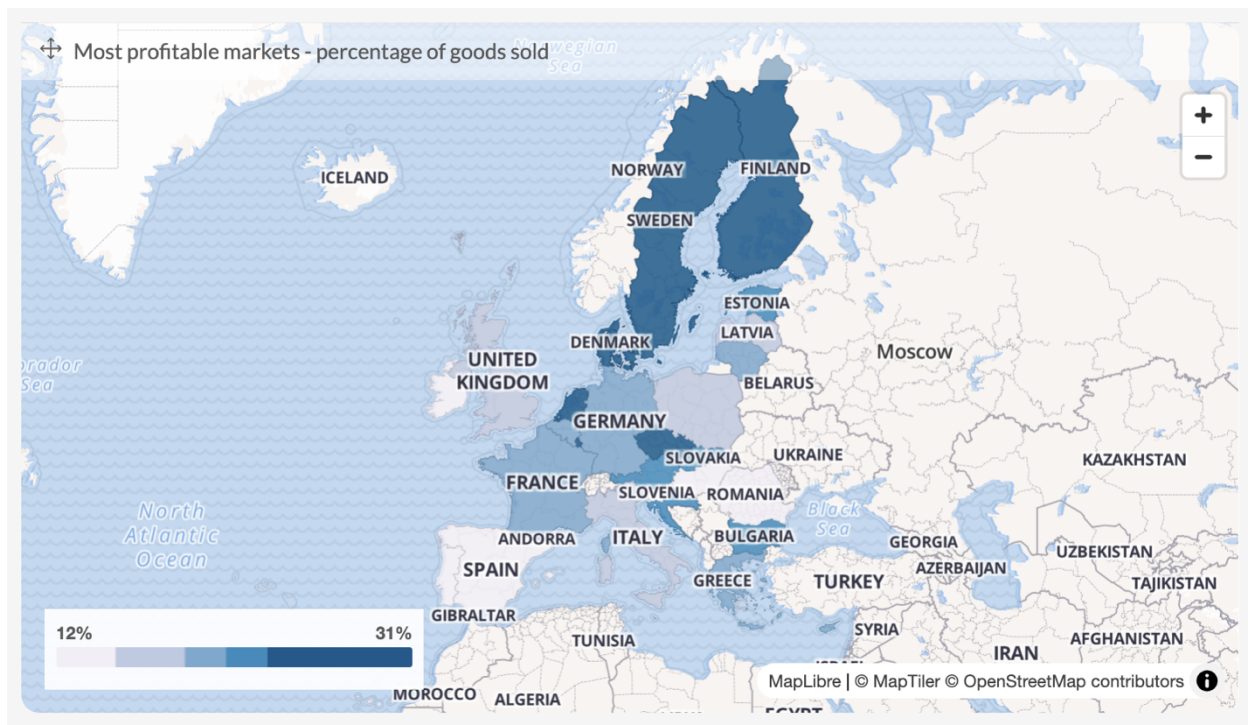
NB: Denne visualisering bruges på step 2 i læringsaktiviteten.

Dårlig brug af farver

Farver kan hjælpe med at gøre visualiseringer lettere at aflæse og forstå. Man skal dog være varsom med at bruge uden en bestemt hensigt. Se på nedstående kort, der viser antallet af amerikanere med rødder fra europæiske lande:



Brugen af farver er ikke meningsfuld her: Italien, hvor 16.000.000 amerikaner har rødder fra, har samme farve som Spanien, hvor 1.000.000 amerikanere har rødder fra – altså 16 gange så meget, hvilket farver ikke viser. En bedre brug af farver havde været en gradient, hvor den lyseste farve er den laveste værdi og den mørkeste er den højeste værdi. Se på nedstående eksempel for eksport af varer:

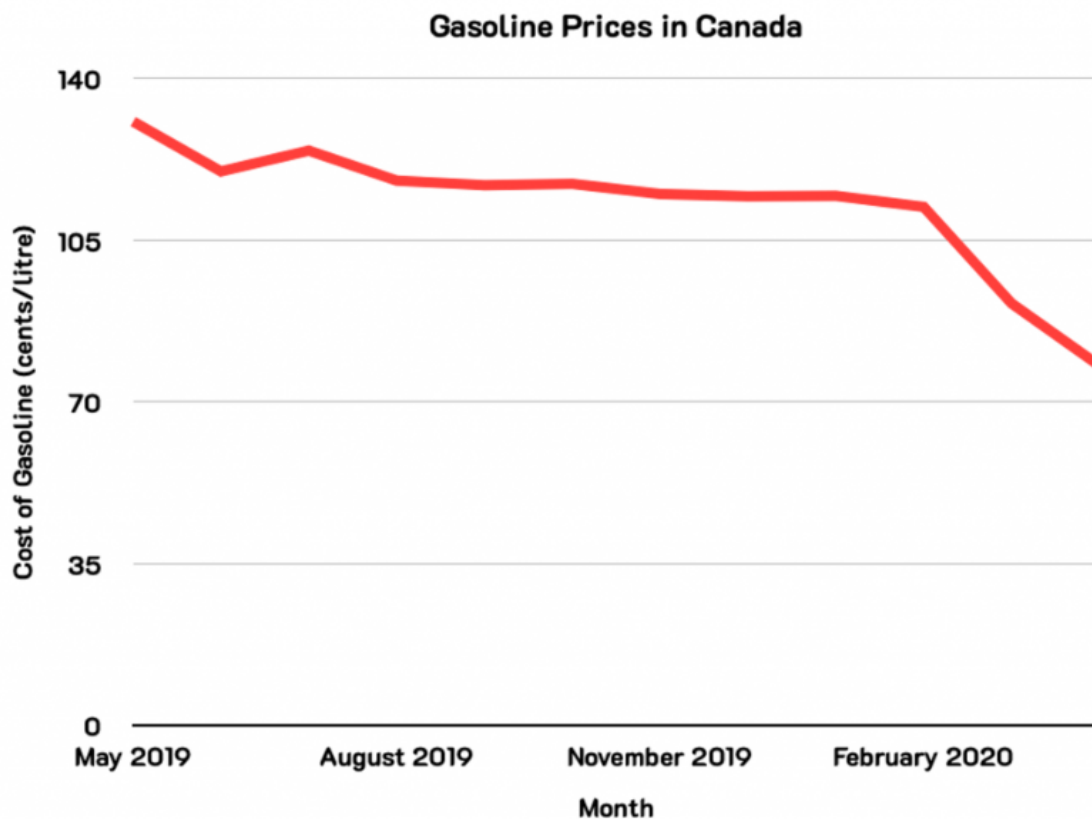


Farverne kommunikerer mere effektivt forskellen i værdier, og er med til at gøre visualiseringen lettere at fortolke.

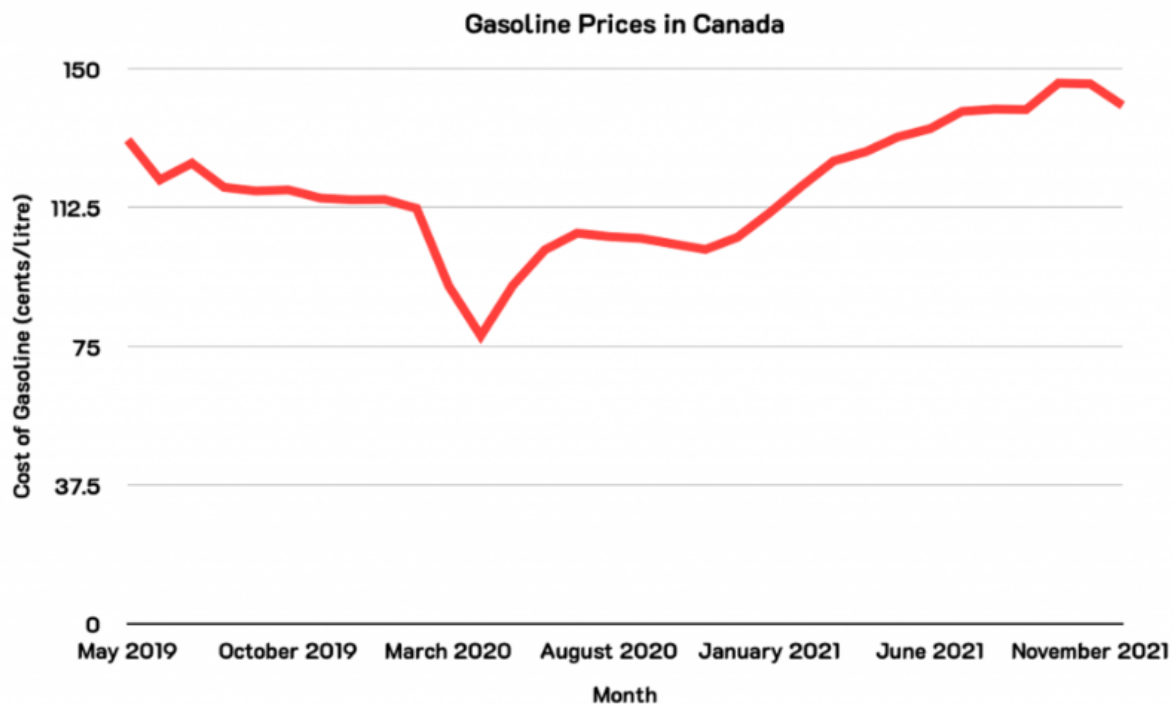
Kilde: <https://www.luzmo.com/blog/bad-data-visualization>

Cherry picking af data

Ved kun at vise en omhyggeligt udvalgt del af data kan man fortælle næsten hvilken som helst historie. Se på nedstående graf:



Prisen på benzin i Canada ser ud til at være faldende siden maj 2019, hvilket mange bilister og politikere nok vil se som positivt. Det er dog ikke hele billedet, da visualiseringen har udeladt noget af data:



Prisen på benzin er faktisk stigende siden maj 2019! Det er en typisk manipulationsteknik at undlade data, som ikke passer ind i ens historie. Som læser skal man være opmærksom på, om der kan være data, som ikke er med i visualiseringen – f.eks. hvis akserne er forskudte.

Kilde: <https://pressbooks.library.torontomu.ca/criticaldataliteracy/chapter/misleading-data-visualizations/>