



Brainstorm-spørgsmål

Brug WTFcsv til at finde spørgsmål, som du kan stille dit regneark

Hvad er WTFcsv?

WTFcsv giver dig et hurtigt overblik over, hvad der er i din CSV-fil. Det hjælper dig til hurtigt at nå frem til nogle spørgsmål, du kan stille dit regneark, så du kan finde historierne i det. Denne øvelse hjælper deltagerne med at opbygge deres datakompetence, mens de dykker ned i sjove datasæt og brainstormer sig frem til spørgsmål om dem.

Læringsmål

- Forståelse for, at der findes mange slags data bl.a. tal, tekst og datoer.
- Styrke evnen til at opstille hypoteser om og finde historier ud fra metadata.
- Forståelse for, at man kan findes mange historier i et datasæt.
- Forståelse for, at flere datasæt kan kombineres for at besvare mere komplicerede spørgsmål.

Øvelsen

Problemløsning

Åbn et regneark i Excel eller et andet program og spørg deltagerne, hvor mange af dem, der arbejder med data på denne måde.

Introducer CSV-filer som én standardiseret måde at gemme regnearksdata. At finde historier i en CSV-fil kan være svært, specielt når der er mange kolonner og rækker. Det kan være en hjælp at begynde med at tænke over hvilke spørgsmål, du kan stille dine data.

WTFcsv hjælper med hurtigt at analysere en CSV-fil for at forstå, hvad den indeholder, så du kan begynde at stille spørgsmål til den.

Hav de gode eksempler med

Vis en åben dataportal i nærheden af stedet, workshoppen afholdes (fx <http://data.gov>). Fortæl om det store spektrum af datasæt, som er tilgængelige online i portaler, som den I kigger på. Download vores UFO-datasæt (<https://databasic.io/wtfcsv/ufo.csv>) og åbn det i Excel. Vis at hver kolonne indeholder forskellige typer information om hver række.

Tidsramme

30 til 45 minutter

Deltagere

3 -100 personer. Alder: 12+.

Udviklet for 6.-9. klasse, ungdomsuddannelser, nyhedsorganisationer, non-profit og frivillige organisationer. Tidligere erfaring med data er ikke nødvendig.

Lokale

- Projektor og computer.
- Mulighed for at kunne arbejde i grupper af tre omkring en computer.
- Store borde, gulv- eller væglads til placering af post-its og tegninger.

Udstyr

- En computer pr. gruppe.
- Kopier af WTFcsv spørgsmålsark (3. ark i denne vejledning).
- Skriveredskaber.

Øvelsen (fortsat)

Demonstrér hvordan man kan skifte mellem det visuelle og tekstuelle overblik ved at vende et kort. Giv eksempler på spørgsmål, man kan stille, hvilke andre data man har brug for for at besvare spørgsmålene, og hvordan man får fat i dem. Et spørgsmål man fx kunne stille UFO-datasættet er: "Bliver der oftere set UFO'er i byer med mange indbyggere end i byer med færre indbyggere?" For at kunne besvare dette må man udover UFO-datasættet også have data om indbyggertal for alle byer. Disse data kan man hente hos US Census records.

Igangsæt øvelsen:

1. Deltagerne får 15 minutter.
2. Deltagerne arbejder i 3-mandsgrupper.
3. Hver gruppe får et WTFcsv spørgsmålsark.
4. Hver gruppe bruger WTFcsv til at analysere en af eksempel-Csv-filerne og brainstorme over mulige spørgsmål, som man kan stille den.
5. Hver gruppe skriver det mest interessante spørgsmål, de fandt på, ned på deres spørgsmålsark.

Præsentationsrunde:

Mens du går rundt blandt grupperne, er det en god idé at være opmærksom på forskellige typer af spørgsmål, der kan fremhæves. Hør om nogle af grupperne har lyst til at fremlægge det spørgsmål, de har valgt. Giv dem max 1 minut, så alle grupper kan komme til orde. Mens grupperne fremlægger deres spørgsmål, sørger du for, at de grupper, du tidligere bemærkede, også får tid.

Fokuser på disse spørgsmål og emner:

- Er alle svarene til jeres spørgsmål indeholdt i jeres datasæt? *Sikkert ikke; nogle gange er det nødvendigt at finde et andet datasæt eller ringe til nogen for at få den information, som man mangler.*
- Hvor kommer dataene fra? Hvis kilden til dataene ikke fremgår, skal man være skeptisk. *(Vi downloadede UFO-dataene fra the National UFO Reporting Center (<http://www.nuforc.org>)).*
- Ser du steder i dataene hvor der mangler værdier? Manglende værdier er én måde, hvorpå data kan være "rodede"! *Spørg deltagerne hvordan data kan være "rodede" og hvad "rene" data betyder?*
- WTFcsv viser de aggregerede (opsummerede) data, men nogle gange kan det være interessant at stille spørgsmål om fx en enkelt række i et datasæt og lede efter en "outlier" (noget der skiller sig ud). *Fandt deltagerne nogle historier om outliers?*

Husk

- Vi skal ikke besvare spørgsmålene, men kun lave en liste af spørgsmål, som man kan stille.
- Der findes ikke dumme spørgsmål.
- Nogle data skal "normaliseres" for at de kan sammenlignes. Hvis man fx skal sammenligne antallet af UFO-observationer pr. by, er det nødvendigt at "normalisere" disse data ud fra befolkningstal, så man kan sammenligne pr. person.

Følgende termer skal introduceres

csv - En "komma-separeret-værdi"-fil ("comma-separated-values"); en standardiseret måde hvorpå tabeller eller information kan gemmes i en tekstfil.

DataTyper - Data klassificeres ofte i typer. Eksempler på almindelige datatyper er tal, datoer og tekst.

Metadata

Metadata er oplysninger om data, så du ved, hvordan de skal forstås.

Normalisering

En process som bruges til at gøre data sammenlignelige.

Outlier ("noget det skiller sig ud")

En del af datasættet som ser meget anderledes ud end resten af dataene.



Spørgsmålsark til uddeling

Vores spørgsmål:

Hvilket spørgsmål er mest interessant at stille jeres datasæt?

Andre data som vi har brug for:

Har I brug for andre datasæt for at kunne besvare jeres spørgsmål?

Kilder til at finde disse andre data:

Hvordan kan I finde de andre data, som I har brug for, for at kunne besvare spørgsmålet?

data
BASIC.io



DataBasic er en række af nemme begynder-værktøjer, der introducerer til arbejdet med data. DataBasic er et projekt skabt af Engagement Lab på Emerson Universitet og MIT Center for Civic Media. Projektet er støttet af Knight Foundation.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

