

Anvendt Statistik for Erhvervsøkonomer

Workshop 2: Opgavesæt

Sådan fungerer workshops: I løser opgaverne i jeres studiegruppen. For hver opgave starter I med at diskutere svaret på opgaven i gruppen. Gruppediskussionen øger jeres forståelse af begreberne samt gør at I bedre kan huske dem. Det er vigtigt at I snakker om eksempler, så I får forståelse for den praktiske anvendelse af det teoretiske begreb. Når I har diskuteret i gruppen, skriver I hver især svaret på opgaven ned med jeres egne ord, dette er god træning til eksamen, som er skriftlig og individuel. Workshopinstruktørerne er til stede, så ræk hånden op hvis I har problemer med en opgave. **De opgaver i ikke når at løse på workshoppen er hjemmeopgaver til næste gang.** Hver workshop starter med at løsningen på opgaverne fra sidste gang gennemgås af jeres instruktører.

2.1 Giv to eksempler på henholdsvis *nominal*, *ordinal* og *interval* data

2.2 Angiv datatypen for hver af følgende dataeksempler:

- Antal kilometer i ugen løbet af joggere
- Startløn for Cand.merc. uddannede
- De måneder hvor virksomhedens medarbejdere vælger at holde ferie
- Socialøkonomisk status: høj, medium, lav

2.3 Angiv datatypen for hver af følgende dataeksempler:

- Den ugentlige lukkekurs for Amazon.com aktien
- Måneden med højest ledighedsrate på Hotel D'Angleterre
- Størrelsen på sodavanden købt af en stikprøve af McDonald's kunder
- Antal Ranch Rovers importeret til Danmark de seneste 10 år
- Legolands årlige omsætning

2.8 En stikprøve af fodboldfans bliver spurgt ind til forskellige aspekter i sportsgrenen. Afgør hvilken type data der er tale om:

- Hvor mange kampe ser du på stadion årligt?
- Hvordan vurderer du kvaliteten af serviceoplevelsen på stadion? (fantastisk, meget god, god, okay, dårlig)
- Har du et sæsonkort
- Hvordan vurderer du kvaliteten af maden på stadion (super, god, dårligt)

2.12 En stikprøve af bilejere bliver stillet følgende spørgsmål. Angiv hvilken datatype der er tale om:

- Bilmærke
- Bilens alder i måneder
- Årligt beløb betalt til bilforsikring
- Antal kørte kilometer i bilen

2.13 Hvornår løber verden tør for olie? En måde at vurdere dette på er at kigge på oliereserverne. Datasæt **Xr02-13** indeholder de kendte oliereserver for top-15 mest olieproducerende lande. Beskriv dette datasæt grafisk i et søjlediagram i Excel og forklar resultatet.

2.14 Brug datasættet fra forrige opgave. Den totale oliereserve i verden er 1.689.078.618.100 tønder. Den totale oliereserve for de 15-lande der indgår i datasættet, er 1.563.350.000.000 tønder. Anvend et cirkeldiagram til at illustrerer den procentvise fordeling af olie på tværs af landene i Excel og forklar resultatet.

2.15 Datasæt **Xr02-15** indeholder det gennemsnitlige olieforbrug per dag for top-20 mest olieforbrugende lande. Brug et søjlediagram og et cirkeldiagram til at illustrere denne information i Excel og forklar resultaterne.

2.16 Olieproduktionen fordeler sig på en række forskellige produktkategorier. Information om andelen af olieproduktionen på tværs af produktkategorier findes i datasæt **Xr02-16**, lav et cirkeldiagram for at illustrer disse tal i Python og forklar resultatet.

2.17 Verden elektricitetsforbrug er vigtig information ift. at sikre tilstrækkelig energiforsyning. Tabellen i **Xr02-17** viser det totale elektricitetsforbrug målt i Kilowatt-timer på tværs af en række lande. Illustrer disse tal grafisk. Bonussspørgsmål: hvilken en variabel ville i tilføje til datasættet for at gøre informationen om elektricitetsforbrug mere sammenligning på tværs af land ¹. Find denne information på Google og tilføj den til datasættet. Lav nu en ny grafisk illustration af datasættet i et søjlediagram i Python og forklar resultatet.

2.24 Hvert år gennemfører Michigan State University Collegiate Employment Research Institute (CERI) en spørgeskemaundersøgelse for at kortlægge startlønnen efter endt universitetsuddannelse. Datasæt **Xr02-24** viser de 18 universitetsuddannelser, som har den højeste startløn. Lønnen er den gennemsnitlige startløn målt i dollars på tværs af universitetsuddannede fra mere end 200 amerikanske universiteter. Anvend et søjlediagram i Python og forklar resultatet.

3.13 Andelen af kunder, der kom ind i en bank i løbet af den første time af åbningstiden, blev registreret over 200 dage. Lav et histogram i datasæt Xr03-13, beskriv formen og forklar resultatet.

3.20 Antallet af bøger, Amazon.com sendte ud dagligt, blev registreret i 100 dage. Lav et histogram i datasæt Xr03-20 og forklar resultatet.

¹ Hint: befolkningstal

3.30 Et prisindeks er en række tal, der bruges til at vise den månedlige ændring i priserne på et fast sæt af varer. Generelt skal ethvert prisindeks bestå af et sæt priser for de udvalgte varer samt en tilsvarende vægt tildelt hver vare i indekset. Besvar følgende spørgsmål:

- a) Hvordan beregnes Consumer Price Index (CPI)?
- b) Hvordan bruges CPI?
- c) Hvorfor vægtes bevægelserne på disse priser?
- d) Hvornår er der ikke behov for vægtning af priserne?
- e) Er CPI det bedste mål for inflation?
- f) Hvad er begrænsningerne ved CPI?
- g) List de indeks, du kender:

3.32 Den amerikanske regering giver 'Medicare' til amerikanere på 65 år og ældre. Medicare Part A betaler udgifterne til hospitals- og sygepleje, men ikke lægernes regninger. Datafilen **Xr03-32** viser Medicare Part A-omkostninger siden 1966. Lav et kurvediagram og forklar, hvad det viser.

3.56 En statistikprofessor havde en teori om, at studerende, der færdiggjorde deres eksamener hurtigere også klarede sig bedre end studerende, der afleverede senere. For at udvikle data til at undersøge gyldigheden af hendes teori, registrerede hun tiden (i minutter), som de studerende brugte inden aflevering af deres eksamener (tidsbegrænsning på 90 minutter) samt den efterfølgende karakter på et udsnit af 12 elever. Lav et scatterdiagram i datasæt **Xr03-56** og beskriv, hvad det viser dig om professorens teori.