



BAGGRUND OG DATABESKRIVELSE

Denne opgave beskæftiger sig med udbudsprisen på brugte biler. Opgaven tager udgangspunkt i datamaterialet i filen `BilTorvet.jmp`. Datamaterialet indeholder detaljeret information om et udsnit af de brugte biler, der har været udbudt til salg via bilportalen BilTorvet.dk i løbet af januar måned 2017.

OPGAVE 1 (BILERNES TYPE)

- a). **Tegn** et søjlediagram over fordelingen af biltype blandt de biler, der er udbudt til salg (variablen *Type*).

JMP-vink: "Analyze" -> "Distribution".

Bestem hvor stor en andel af de udbudte biler, der er personbiler.

- b). **Tegn** et lagkagediagram over fordelingen af biltype blandt de biler, der er udbudt til salg.

JMP-vink: "Graph" -> "Graph Builder".

Angiv på det optegnede lagkagediagram de procentvise andele hørende til hver biltype.

- c). **Tegn** separate søjlediagrammer over fordelingen af biltype for henholdsvis benzin- og dieslbiler (variablen *Drivmiddel*).

*JMP-vink: "Analyze" -> "Distribution" og placer *Drivmiddel* under "By".*

Angiv andelen af personbiler for hver af de to drivmidler.

- d). **Vurdér** om der ser ud til at være en sammenhæng mellem bilens type og drivmiddel.

OPGAVE 2 (BILERNES GRØNNE EJERAFGIFT)

- a). **Tegn** et histogram og et boxplot over fordelingen af den grønne ejeravgift blandt de udbudte biler (variablen *Årlig grøn ejeravgift*).

JMP-vink: "Analyze" -> "Distribution".

- b). **Lav** en summarisk fordelingsbeskrivelse af fordelingen af den grønne ejeravgift (inddrag relevante nøgletal samt figurerne fra delspg. a)).

- c). **Gør rede for** i hvilket omfang de beregnede værdier af median og gennemsnit indikerer en asymmetrisk fordeling af den grønne ejeravgift.



Gør rede for hvorvidt det er muligt at anvende den empiriske regel til beskrivelse af fordelingen af den grønne ejerafgift.

- d). **Angiv** og fortolk minimum, maksimum, 1. kvartil og 3. kvartil i fordelingen af den grønne ejerafgift.

- e). **Gør rede for** om det er korrekt, at mindst 50% af de udbudte biler har en årlig grøn ejerafgift på *mindst* 3.000 kr.

Angiv hvor meget der *højest* betales i årlig grøn ejerafgift på de 25% af de udbudte biler med den *laveste* årlige afgift.

Angiv hvor meget der *mindst* betales i årlig grøn ejerafgift på de 25% af de udbudte biler med den *højeste* årlige afgift.

Angiv hvor stor en andel af de udbudte biler, der har en årlig grøn ejerafgift på mindst 7.000 kr.

- f). **Tegn** separate histogrammer og boxplot over fordelingen af den grønne ejerafgift for henholdsvis benzin- og dieslbiler (dvs. tegn figurer af variabelen *Årlig grøn ejerafgift* opdelt efter værdien af variabelen *Drivmiddel*).

JMP-vink: "Analyze" -> "Distribution" og placer *Drivmiddel* under "By".

- g). **Vurdér** på baggrund af relevante nøgletal og figurer, om der ser ud til at være en forskel i størrelsen af den grønne ejerafgift på benzin- og dieslbiler.

OPGAVE 3 (BILERNES MOTORKRAFT)

- a). **Tegn** et histogram og et boxplot af antal hestekræfter på de udbudte biler (variabelen *Hestekræfter*).

- b). **Lav** en summarisk fordelingsbeskrivelse af fordelingen af antal hestekræfter (inddrag relevante nøgletal samt figurerne fra delspg. a)).

- c). **Angiv** det *maksimale* antal hestekræfter blandt de 25% af bilerne med *færrest* hestekræfter.

Angiv det *minimale* antal hestekræfter blandt de 10% af bilerne med *flest* hestekræfter.

- d). **Gør rede for** hvorvidt det er muligt at bruge den empiriske regel til beskrivelse af antallet af hestekræfter blandt de udbudte biler.

Gør rede for hvad den empiriske regel fortæller om forventningerne til antallet af



heste kræfter blandt biler udbudt til salg på Biltorvet.dk (idet vi her forudsætter, at den empiriske regel finder anvendelse).

- e). **Gør rede for** (ved hjælp af en passende valgt fremgangsmåde) om der ser ud til at være en sammenhæng mellem drivmiddel og antal heste kræfter for biler udbudt til salg på Biltorvet.dk.

OPGAVE 4 (BILERNES PRIS)

I denne opgave ser vi udelukkende på biler udbudt til salg af modellen Kia Picanto.

JMP-vink: "Rows" -> "Data Filter".

- a). **Tegn** et histogram og et boxplot af fordeling af årgangen af biler af mærket Kia Picanto udbudt til salg.

Angiv hvor mange biler af mærket Kia Picanto, der er udbudt til salg indenfor hver årgang.

JMP-vink: Lav variabelen "Årgang" om fra kvantitativ til nominal ved at højreklikke på den blå trekant til venstre for variabelen i søjlen yderst til venstre i datafilen. Vælg dernæst "Analyze" -> "Distribution" og vælg herefter "Histogram Options" -> "Show Counts" i JMPs histogram output.

- b). **Tegn** et histogram af fordelingen af udbudsprisen på biler af mærket Kia Picanto, årgang 2016.

- c). **Gør rede for** hvorvidt det er muligt at bruge den empiriske regel til beskrivelse af prisen på brugte biler af mærket Kia Picanto, årgang 2016.

Gør rede for hvad den empiriske regel fortæller om forventningerne til udbudsprisen på en brugt bil af mærket Kia Picanto, årgang 2016 (idet vi her forudsætter, at den empiriske regel finder anvendelse).



DATAKILDE

Datamaterialet i ovenstående opgaver er venligst stillet til rådighed af bilportalen **Biltorvet.dk** (biltorvet.dk). Virksomheden giver følgende beskrivelse af sig selv:

Biltorvet A/S er et IT-firma, som producerer løsninger til bilforhandlere. Medarbejderskaren er på 40 personer bestående af softwareudviklere, projektledere, kommunikationsfolk, konsulenter mm. Analyse og processering af "big data" såvel som "small data" er et felt, vi gerne vil dygtiggøre os indenfor, og vi anser det for vigtigt at have profiler, som mestrer dette både teoretisk og praktisk. Vi er derfor altid på udkig efter dygtige mennesker med en profil indenfor ovenstående.

DISCLAIMER:

Datamaterialet i nærværende opgaver må udelukkende anvendes i forbindelse med undervisningen i faget Dataanalyse på CBS. Datamaterialet er fortroligt og må ikke videredistribueres eller offentliggøres andet end overfor studerende, der følger faget Dataanalyse på CBS, hverken skriftligt, mundtligt eller på anden vis, og ej heller de resultater og indsigter, der måtte komme ud af arbejdet med datamaterialet.



KORTFATTEDE TALLØSNINGER

OPGAVE 1

- a): andel personbiler = 93,2%
- c): andel personbiler (benzin) = 98,1%; andel personbiler (diesel) = 86,7%

OPGAVE 2

- b): gennemsnit = 2.720 kr.; median = 2.380 kr.; standardafvigelse = 1.956 kr.; interkvartilbredde = 2.700 kr.
- d): minimum = 260 kr.; maksimum = 12.000 kr.; 1. kvartil = 1.200 kr.; 3. kvartil = 3.900 kr.
- e): Ejerne af de 25% af bilerne med lavest grøn ejerafgift betaler højst 1.200 kr. årligt; Ejerne af de 25% af bilerne med højest grøn ejerafgift betaler mindst 3.900 kr. årligt; Ca. 3% af bilerne har en grøn ejerafgift på mindst 7.000 kr.

OPGAVE 3

- b): gennemsnit = 106 hk; median = 100 hk; standardafvigelse = 30 hk; interkvartilbredde = 41 hk
- c): De 25% af bilerne med færrest hestekræfter har maksimalt 84 hk;
De 10% af bilerne med flest hestekræfter har mindst 145 hk
- d): ca. 68% af bilerne vil have et antal hestekræfter i intervallet [76 hk; 136 hk];
ca. 95% af bilerne vil have et antal hestekræfter i intervallet [46 hk; 166 hk];
næsten alle bilerne vil have et antal hestekræfter i intervallet [16 hk; 196 hk]

OPGAVE 4

- a): 2009: 1; 2010: 2; 2011: 0; 2012: 6; 2013: 2; 2014: 2; 2015: 3; 2016: 27
- c): ca. 68% af bilerne vil have en udbudspris i [97.258 kr.; 112.824 kr.];
ca. 95% af bilerne vil have en udbudspris i intervallet [89.475 kr.; 120.606 kr.];
næsten alle bilerne vil have en udbudspris i intervallet [81.692 kr.; 128.389 kr.]