

Marsbewoners

Bekijk het bijgeleverde bestand: "marsbewoners.txt". Hierin staan de lengtes van 10.000 marsbewoners die gemeten werden (in cm) door de mars exploration rovers. Probeer dit bestand in te laden. Het resultaat zou in een variabele "mars" moeten komen van het type Series (dus geen DataFrame). Je zal de kolommen dus moeten samenvoegen.

Geef antwoord op de volgende vragen:

- a) wat is de gemiddelde lengte van een marsbewoner?
- b) wat is de mediaan? Wat betekent dit?
- c) wat is de gemiddelde absolute afwijking? Wat betekent dit?
- d) wat is de standaardafwijking? Wat betekent dit?
- e) bepaal de kwartielen en de decielen.
- f) 20% van de marsbewoners is kleiner dan ... cm (gebruik de decielen)
- g) het percentage marsbewoners dat kleiner of gelijk is aan 60 cm is ... (experimenteer met de kwantielen om dit te vinden)
- h) teken een boxplot. Welke informatie zie je hier?
- i) maak klassen van 5 cm en bepaal de frequenties (laat de klassen beginnen bij 0 cm). Wat is de modus?
- j) maak klassen van 5 cm en teken een histogram. Wat zie je? Hoe zou je dit verklaren? Zie je dit ook als je klassen maakt van 10 cm?

Open nu de bestanden "mannetjes.txt" en "vrouwtjes.txt". Hierin staan de lengtes van dezelfde marsbewoners, maar nu per geslacht.

Geef antwoord op de volgende vragen:

- k) wat zijn de gemiddelde lengtes van een marsmannetje en een marsvrouwtje?
- l) bereken ook de medianen
- m) bereken de standaardafwijkingen
- n) bepaal de kwartielen
- o) maak een boxplot waar je deze twee verdelingen naast elkaar ziet (hint: je kan een list van datasets meegeven aan boxplot())
- p) bepaal de grenzen voor uitschieters voor de mannetjes en de vrouwtjes.
- q) hoeveel uitschieters zijn er voor ieder geslacht?
- r) zijn er extreme uitschieters? Welk zijn de grenzen daarvoor?