DATAHANTERING.

Data is die insameling van inligting vir 'n spesifieke doel.

Wanneer data ingesamel word deur dit te tel, word dit **diskrete data** genoem. 'n Voorbeeld van diskrete data is bv skoengroottes : 'n Persoon dra 'n nommer 5 skoen, òf $5\frac{1}{2}$ òf 6 òf.....Daar is nie skoengroottes tussen hierdie nommers nie.

Data wat verkry word deur meting, word **kontinue data** genoem. As jou lengte 1,5 meter is, beteken dit nie dat dit presies 1,5 meter is nie, dit mag byvoorbeeld 1,531 meter wees, maar afgerond tot een desimale getal is dit 1,5 meter.

1. Ordening van data :

Data moet altyd georden word voordat enige afleidings gemaak kan word. Data word georden deur die data te rangskik van klein na groot.

15 Leerders het die volgende persentasies in 'n toets behaal :

55, 68, 54, 77, 88, 71, 68, 51, 92, 48, 80, 75, 61, 70, 69.

Die data, gerangskik van klein na groot, sal wees :

48, 51, 54, 55, 61, 68, 68, 69, 70, 71, 75, 77, 80, 88, 92.

Stingel-en-blaarvoorstellings :

Data kan ook georden word deur 'n stingel-en-blaarvoorstelling. Die tiene word as die stingel aangedui en die ene as die blare, bv. met die getal 26 sal 2 die stingel wees en 6 die blaar en met die getal 135 sal 13 die stingel wees en 5 die blaar. Die stingel word aan die linkerkant van 'n vertikale lyn geskryf en die blare aan die regterkant. Die blare word dan numeries gerangskik.

Voorbeeld 1

Die volgende punte is deur 'n rugbyspan in 24 wedstryde aangeteken :

15	22	12	31	41	17	28	16	25	36	42	47
12	34	44	14	19	21	7	18	26	24	35	13

Organiseer die data met behulp van 'n stingel-en-blaartabel.

Oplossing :

Stingel	Blaar
0	7
1	5,2,7,6,2,4,9,8,3
2	2,8,5,1,6,4
3	1,6,4,5
4	1,2,7,4

Rangskik nou die blare numeries.

Stingel	Blaar
0	7
1	2,2,3,4,5,6,7,8,9
2	1,2,4,5,6,8
3	1,4,5,6
4	1,2,4,7

Om die data nou van klein na groot te rangskik, is maklik.

7 12 12 13 14 15 16 17 18 19 21 22

24 25 26 28 31 34 35 36 41 42 44 47

2. Berekening van maatstawwe van sentrale neiging :

2.1 Ongegroepeerde data :

Die toetspunte van 20 leerders is 'n voorbeeld van ongegroepeerde data. Die data is nie in intervalle gegroepeer nie, maar daar word met die individuele toetspunte gewerk.

2.1.1 **Die gemiddelde** (rekenkundige gemiddelde) word bereken deur die som van die waardes deur die aantal waardes te deel. Dit word aangedui deur $\bar{x}$ (wat gelees word as streep x) en die volgende formule word gebruik :

$$\boxed{\bar{x} = \frac{\sum x}{n} = \frac{\textit{Die som van al die waarnemings } (\sum x)}{\textit{Die aantal waarnemings } (n)}}$$

2.1.2 **Die modus** is daardie getal wat die meeste voorkom. Wanneer daar twee getalle is wat ewe veel voorkom, sê ons die data is bimodaal. Wanneer meer as twee modusse voorkom, geld die modus nie.

2.1.3 **Die mediaan** , die middelste waarde, verdeel die data in twee dele, 50% van die data lê onder die mediaan en 50% van die data lê bo die mediaan.

Voorbeeld 2.

Die volgende tabel toon die massa van 12 graad 10 leerders.

51	48	55	41	61	58	62	53	66	59	46	51

2.1 Bereken die gemiddelde.
2.2 Wat is die modus ?
2.3 Bereken die mediaan.

Oplossing :

Data moet altyd eers georden word. Rangskik dus eers die data van klein na groot.

41	46	48	51	51	53	55	58	59	61	62	66

2.1 $\textit{Die gemiddeld } (\bar{x}) = \dfrac{\textit{Die som van al die waarnemings}}{\textit{Die aantal waarnemings}}$

$= \dfrac{41+46+48+51+51+53+55+58+59+61+62+66}{12}$

$= 54{,}25$

2.2 *Die modus* $= \underline{51}$ (51 *kom twee keer voor.)*

Die mediaan kan op twee maniere bereken word :

2.3 *Metode 1 (Deur die data te tel) :*

Die mediaan is die middelste waarde. Daar is 12 leerders, die mediaan lê dus tussen 53(die 6de waarde) en 55(die 7de waarde)

↓

41 46 48 51 51 53 55 58 59 61 62 66

dus : die gemiddeld van die 6de en die 7de waarde $\therefore \dfrac{53+55}{2} = 54$

$\therefore$ *Die mediaan* $= 54$

Metode 2 (Deur die formule te gebruik) :

Posisie van die mediaan $= \frac{1}{2}(n+1) \quad = \frac{1}{2}(12+1) = 6\frac{1}{2}$

Onthou : $6\frac{1}{2}$ *is slegs waar die mediaan lê, nie die waarde van die mediaan nie.*
Die mediaan lê dus tussen die sesde en die sewende getal.

Die sesde getal $= 53$ *en die sewende getal* $= 55$

$\therefore$ *Die mediaan* $= \dfrac{53+55}{2} = \underline{54}$

2.2 Gegroepeerde data :

Wanneer met 'n groot aantal data gewerk word, word die data in klasintervalle gegroepeer.

2.2.1 **Die rekenkundige gemiddelde :** Om die rekenkundige gemiddelde van gegroepeerde data te bereken, word die gemiddelde van elke klasinterval bereken. Hierdie gemiddeld van die **klasinterval** (x) word vermenigvuldig met die **frekwensie** (f) van die klasinterval.

Die rekenkundige gemiddeld van gegroepeerde data is dan :

$$\bar{x} = \frac{\sum f.x}{n} = \frac{\textit{Die som van (gemiddeld van klasinterval} \times \textit{frekwensie)}}{\textit{totale aantal waarnemings}}$$

2.2.2 **Die modus** of modale interval is die klasinterval met die hoogste frekwensie.

2.2.3 **Die mediaan :** Die posisie van die mediaan is $\frac{1}{2}n$ waar n = die aantal waarnemings.

Voorbeeld 3

Die volgende tabel toon die massa van 50 Graad 10 meisies.

Klasinterval	Klasmiddelwaarde(x)	Frekwensie(f)	$(x) \times (f)$
$35 \leq x < 40$	37,5	1	37,5
$40 \leq x < 45$	42,5	13	552,5
$45 \leq x < 50$	47,5	18	855
$50 \leq x < 55$	52,5	11	577,5
$55 \leq x < 60$	57,5	5	287,5
$60 \leq x < 65$	62,5	2	125
		Som van $(f)(x)$	2435

3.1 Bereken die rekenkundige gemiddeld.
3.2 Wat is die modale interval?
3.3 In watter klasinterval lê die mediaan.

Oplossing :

3.1 *Rekenkundige Gemiddeld* $= \frac{\textit{Die som van } (x \times f)}{\textit{totale aantal waarnemings}}$

$$= \frac{37,5 + 552,5 + 855 + 577,5 + 278,5 + 125}{50}$$

$$= \frac{2435}{50} \qquad = \underline{48,7}$$

3.2 *Die Modale interval* $=$ *die interval* $45 \leq x < 50$.

3.3 *Die Mediaan* : *Posisie van die mediaan* $= \frac{1}{2}n = \frac{1}{2}(50) = 25$

Die 25ste waarde kom voor in die interval $[45; 50)$.

$\therefore$ *Mediaan kom voor in die interval* $[45; 50)$

3. Maatstawwe van verspreiding :

Die mediaan verdeel die data in twee dele. Die data kan egter ook in **kwartiele** verdeel word, waar 25% van die data onder die eerste kwartiel lê en 75% van die data onder die derde kwartiel lê.

3.1 **Berekening van maatstawwe van verspreiding met ongegroepeerde data :**

3.1.1 **Die eerste kwartiel, Q_1**: Posisie van **eerste kwartiel,** Q_1, is $\frac{1}{4}(n+1)$.

3.1.2 **Die derde kwartiel, Q_3**: Posisie van **derde kwartiel,** Q_3, is $\frac{3}{4}(n+1)$.

3.1.3 **Die variasiewydte of omvang** = grootste waarde – die kleinste waarde.

3.1.4 **Die interkwartielvariasie-wydte** = $Q_3 - Q_1$.

Behalwe dat die data in kwartiele verdeel kan word, kan dit ook in **persentiele** verdeel word. Persentiele verdeel die data in **honderdstes**. 10% van die data sal byvoorbeeld onder die 10de persentiel val. Bereken persentiele net soos die kwartiele. Die posisie van die 10de persentiel sal dus wees : $\frac{1}{10}(n+1)$.

Voorbeeld 4

Kyk weer na die data in voorbeeld 2 :

41	46	48	51	51	53	55	58	59	61	62	66

Bereken :

4.1 Die eerste kwartiel
4.2 Die derde kwartiel.
4.3 Die omvang.
4.4 Die interkwartielvariasie-wydte.

Oplossing :

Die kwartiele kan net soos die mediaan op twee maniere bereken word.

4.1 *Metode 1 :*

41 46 48 ↓ 51 51 53 ↓ 55 58 59 ↓ 61 62 66

Die mediaan verdeel die data in twee gelyke dele.

Die eerste kwartiel verdeel die eerste helfte van die data in twee gelyke dele.
Daar is 6 getalle links van die mediaan.

Die eerste kwartiel lê dus tussen die 3de en 4de getal.

$\therefore Q_1 = \frac{48+51}{2} = \underline{49,5}$

Metode 2 :

Posisie van $Q_1 = \frac{1}{4}(n+1) = \frac{1}{4}(13) = 3,25$

Dit is die posisie van Q_1. *Bereken nou* Q_1.

$\therefore Q_1 = 3de\ getal + 0,25(Verskil\ tussen\ 4de\ getal\ en\ 3de\ getal.)$

$\therefore Q_1 = 48 + 0,25(51-48)$ $(4de\ getal = 51\ :\ 3de\ getal = 48.)$

$= \underline{48,75}$

4.2 *Metode 1 :*

Die derde kwartiel verdeel die tweede helfte van die data in twee gelyke dele.
Daar is 6 getalle regs van die mediaan.

Die derde kwartiel lê dus tussen die 9de en 10de getal.

$\therefore Q_3 = \frac{59+61}{2} = \underline{60}$

Metode 2 :

Posisie van $Q_3 = \frac{3}{4}(n+1) = \frac{3}{4}(13) = 9,75$

$\therefore Q_3 = 9de\ getal + 0,75(10de\ getal - 9de\ getal.)$

$= 59 + 0,75(61-59)$ $= \underline{60,5}$

4.3 *Omvang = Hoogste waarde – laagste waarde*

$= 66 - 41$ $= \underline{25}$

4.4 *Interkwartielvariasie-wydte* $= Q_3 - Q_1$

$$= 60 - 49,5 = \underline{10,5}$$

en as die waardes van die tweede metode gebruik word :

$$= 60,5 - 48,75 \qquad = \underline{11,75}$$

Al verskil die antwoorde effens sal albei metodes as reg aanvaar word.

3.2 **Berekening van maatstawwe van verspreiding met gegroepeerde data** :

3.2.1 **Die eerste kwartiel, Q_1** : Die posisie van die eerste kwartiel, Q_1, is $\frac{1}{4}n$.

3.2.2 **Die derde kwartiel, Q_3** : Die posisie van die derde kwartiel, Q_3, is $\frac{3}{4}n$.

Voorbeeld 5

Kyk na die data in voorbeeld 3.

5.1 In watter interval lê die eerste kwartiel?
5.2 In watter interval kom die derde kwartiel voor?

Oplossing :

5.1 *Posisie van* $Q_1 = \frac{1}{4}n \qquad = \frac{1}{4}(50) \qquad = 12,5$

Die 12,5*de waarnemig kom voor in die interval* [40 – 45).

$\therefore$ *Q_1 lê in die interval* [40 – 45)

5.2 *Posisie van* $Q_3 = \frac{3}{4}n \qquad = \frac{3}{4}(50) \qquad = 37,5$

Die 37,5*de waarneming kom voor in die interval* [50 – 55).

$\therefore$ *Q_3 lê in die interval* [50 – 55)

3. **Berekening van 'n vyfsyferopsomming :**

Die vyfsyferopsomming verskaf die kleinste waarde, die eerste kwartiel (Q_1), die mediaan, die derde kwartiel (Q_3) en die hoogste waarde.

Voorbeeld 6

Die volgende is die toetspunte van 10 leerders :

68	49	54	65	42	87	44	47	65	85

Gee 'n vyfsyferopsomming van die data.

Oplossing :

Orden altyd eers die data, m.a.w. rangskik van klein na groot.

42	44	47	49	54	65	65	68	85	87

Metode 1 *(Deur die data te tel) :*

1. *Die kleinste waarde* = 42

2. *Die eerste kwartiel :*

Verdeel jou data in 4 *gelyke dele.*

		↓		↓			↓		
42	44	47	49	54	65	65	68	85	87

Daar is 5 *getalle links van die mediaan.*

$\therefore$ *Q_1 is die derde getal* [*Die middelste getal van die* 5 *getalle.*]

$\therefore Q_1 = 47$

3. *Die mediaan :*

Daar is 10 *getalle, die mediaan lê dus tussen die* 5*de en die* 6*de getal.*

$\therefore$ *Die mediaan* $= \dfrac{54+65}{2} = 59,5$

4. *Die derde kwartiel :*

Daar is 5 *getalle regs van die mediaan.*

$\therefore$ Q_3 *is die agste getal* [*Die middelste getal van die* 5 *getalle.*]

$\therefore$ $Q_3 = 68$

5. *Die hoogste waarde* = 87

$\therefore$ *Die vyfsyferopsomming is* : 42; 47; 59,5; 68; 87.

Metode 2 *(Gebruik van die formules) :*

1. *Die kleinste waarde* = 42

2. *Die eerste kwartiel :*

Posisie van eerste kwartiel is $\frac{1}{4}(n+1)$ $\therefore$ $\frac{1}{4}(10+1) = 2\frac{3}{4}$

$\therefore$ Q_1 = 2*de getal* + $\dfrac{3}{4}$(3*de getal* – 2*de getal.*)

$\therefore$ $Q_1 = 44 + \frac{3}{4}(47-44)$ [2*de getal* = 44 *en* 3*de getal* = 47]

$= 46,25.$

3. *Die mediaan :*

Posisie van die mediaan is $\frac{1}{2}(n+1)$ $\therefore$ $\frac{1}{2}(10+1) = 5\frac{1}{2}$

Die mediaan lê dus tussen die 5*de en* 6*de getal.*

$\therefore$ *Die mediaan* $= \dfrac{54+65}{2} = \underline{59,5}$

4. *Die derde kwartiel :*

Posisie van derde kwartiel is $\frac{3}{4}(n+1)$ $\therefore$ $\frac{3}{4}(10+1) = 8\frac{1}{4}$

Q_3 *is dus getal* 8 + $\dfrac{1}{4}$(9*de getal* – 8*ste getal*)

$\therefore$ $Q_3 = 68 + \frac{1}{4}(85-68)$ [8*ste getal* = 68 *en die* 9*de getal* = 85.]
$= 72,25.$

5. *Die hoogste waarde* = 87

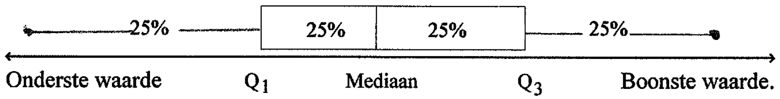
$\therefore$ *Die vyfsyferopsomming is* : 42; 46,25; 59,5; 72,25; 87.

Daar is weer 'n klein verskil tussen die waardes wat verkry is. Doen egter die metode wat jy in die klas geleer het. Albei antwoorde sal as reg aanvaar word.

4. **Teken van 'n houer-en-punt-diagram :**

Die houer-en-punt-diagram is 'n grafiese voorstelling van die vyfsyfer-opsomming. 'n Houer-en-punt-diagram het 4 dele en elke deel stel 'n kwart van die data voor.

Kyk na die volgende voorstelling van die houer-en-punt-diagram :



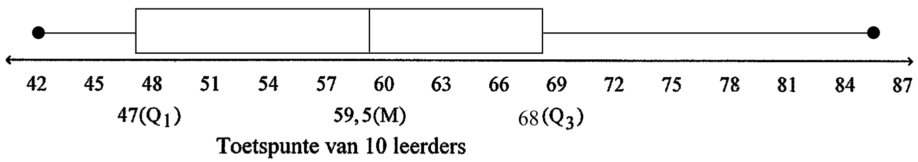
Wanneer jy die eerste kwartiel, die mediaan en die derde kwartiel bereken het, is dit baie maklik om die houer-en-punt-diagram te teken.

Voorbeeld 7

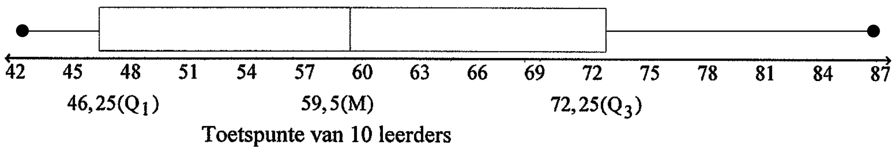
Teken 'n houer-en-punt-diagram van die data in voorbeeld 6.

<u>Oplossing :</u>

<u>*As jy metode* 1 *gebruik het* :</u>



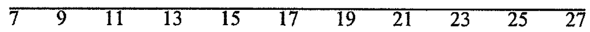
<u>*As jy metode* 2 *gebruik het* :</u>



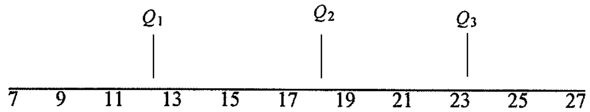
Om 'n houer-en-punt diagram te teken :

<u>Vyfsyferopsomming: 9; 12; 18; 23; 27</u>

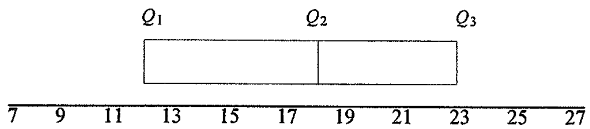
1. Gebruik 'n geskikte skaal en trek'n horisontale as wat begin by die minimum waarde en eindig by die maksimun waarde.



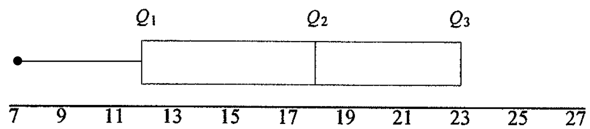
2. Trek kort vertikale lyne bo Q_1, Q_2 en Q_3 net bo die horisontale as.



3. Verbind die vertikale lyne bo Q_1 en Q_3 om 'n houer te vorm.



4. Trek 'n lyn aan die linkerkant van die houer na die minimum waarde.



5. Trek 'n lyn van die regterkant van die houer na die maksimum waarde.