



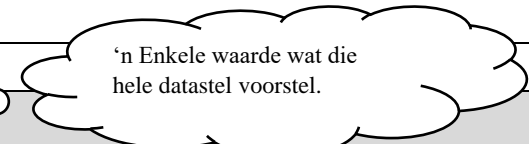
VAK en GRAAD	Wiskunde	Graad 12
KWARTAAL 3	Week 2	
ONDERWERP	Statistiek	
DOELSTELLINGS VAN LES	Hersiening van graad 10 en 11 statistiese konsepte.	
BRONNE	<b>Papiergebasseerde hulpbronne</b>	<b>Digitale hulpbronne</b>
	<i>Gaan asseblief na die Statistiek hoofstuk in jou handboek.</i>	Klik op  om 'n video te sien hoe om jou sakrekenaar te gebruik in die les Klik op  om 'n video te sien oor die konsep in die les.

**INLEIDING:**

Liewe leerder in vorige grade het jy geleer van die verskillende vlakke statistiese ondersoek. Die verskillende vlakke is: insameling van data, sortering van data, grafiese voorstelling van data, interpretasie van data en gevolgtrekkings wat gemaak kan word. In hierdie sessie gaan ons nog tegnieke ondersoek hoe om data te bestuur, op te som en hoe om gegroepeerde en ongegroepeerde data voor te stel.

**KONSEPTE EN VAARDIGHEDE:**

**Maatstawwe van sentrale neiging (ongegroepeerde data)**



Gemiddelde ( $\bar{x}$ )	Mediaan ( $Q_2$ )	Modus
<ul style="list-style-type: none"> <li>Die som van al die waardes (<math>x</math>) van die datastel gedeel deur die aantal waardes in die datastel (<math>n</math>).</li> <li><math>\bar{x} = \frac{\sum x}{n}</math> </li> </ul> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <a href="https://youtu.be/cHsGIHWOSaA">https://youtu.be/cHsGIHWOSaA</a> </div>	<ul style="list-style-type: none"> <li>Die middel waarde van die geordende datastel.</li> <li>Verdeel data in twee gelyke stelle. </li> <li>Posisie van <math>Q_2 = \frac{1}{2}(n + 1)</math>.</li> </ul> <div style="border: 1px solid black; padding: 5px; margin-top: 10px; text-align: center;"> Rangskik data in stygende orde. </div>	<ul style="list-style-type: none"> <li>Die waarde in die datastel wat die hoogste frekwensie het of die waarde wat die meeste voorkom.</li> </ul>



**Voorbeeld 1:**

Beskou die volgende stel punte (uit 50) vir 'n wiskunde toets.

25	26	21	20	30	32	45	37	41	32	33	42	25	26	26	33	43	35	36	38
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Bepaal die gemiddelde, mediaan en modus.

**Oplossing:**

**Gemiddelde**

$$\bar{x} = \frac{\sum x}{n} = \frac{25 + 26 + 21 + 20 + 30 + 32 + 45 + 37 + 41 + 32 + 33 + 42 + 25 + 26 + 26 + 33 + 43 + 35 + 36 + 38}{20} = \frac{646}{20} = 32,3$$

**Mediaan**

20, 21, 25, 25, 26, 26, 26, 30, 32, 32, 33, 33, 35, 36, 37, 38, 41, 42, 43, 45.

Rangskik data in stygende orde.

Posisie van  $Q_2 = \frac{1}{2}(n + 1) = \frac{1}{2}(20 + 1) = 10,5^{de}$  waarde.

$\therefore Q_2$  lê tussen die  $10^{de}$  en  $11^{de}$  waarde.

$\therefore Q_2 = \frac{32+33}{2} = 32,5$  (Indien die posisie van  $Q_2$  tussen twee waardes; bereken die som van die twee waardes en deel dit deur 2.)

**Modus**

26

**KAN JY?**

Bepaal die gemiddelde, mediaan en modus van die onderstaande data.

15	16	21	23	33	34	17
32	22	19	18	15	16	14

**Antwoorde:**

Gemiddelde: 21,07      Mediaan: 18.5      Mode: 15 en 16



**Maatstawwe van sentrale neiging (geordende data)**

**Voorbeeld 2:**

Bepaal die geskatte gemiddelde, mediaan en modus van die onderstaande data.

Klasinterval	Frekwensie	Middelpunt van klasinterval	Frekwensie × Middelpunt
$0 \leq x \leq 20$	6	10	60
$20 < x \leq 40$	8	30	240
$40 < x \leq 60$	10	50	500
$60 < x \leq 80$	3	70	210
$80 < x \leq 100$	2	90	180
	29		1190

**Oplossing:**

**Gemiddeld**

$$\bar{x} = \frac{\sum x}{n} = \frac{1190}{29} = 41,03$$

**Mediaan**

Posisie van  $Q_2 = \frac{1}{2}(n + 1) = \frac{1}{2}(29 + 1) = 15^{de}$  waarde.

Die  $15^{de}$  waarde val in  $40 < x \leq 60$  klasinterval.

$\therefore Q_2 = 50$  (middelpunt van klasinterval)

**Modus** =  $40 < x \leq 60$

**Berekening avn geskatte gemiddeld:**

1. Skep addisionele kolomme vir **Middelpunt van klasinterval** en **Frekwensie × Middelpunt**.


2. Bereken die **Middelpunt van klasinterval** en **Frekwensie × Middelpunt**.

2.1 **Middelpunt van klasinterval** =  $\frac{\text{Onderste limiet van klas} + \text{boonste limiet van klas}}{2}$

2.2 Vermenigvuldig die frekwensie kolom met die middelpunt van elke klasinterval.

3. Bereken die som van die **Frekwensie** kolom om "n" te vind.

4. Bereken die som van **Frekwensie × Middelpunt** kolom om " $\sum x$ " waarde te vind.

5. Vervang stap 3 en 4 in  $\bar{x} = \frac{\sum x}{n}$  om geskatte gemiddeld te kry. 

<https://youtu.be/MgKljN7mpqk>

**Mediaan**

- Bepaal die posisie van  $Q_2$ .
- Bepaal in watter klasinterval die posisie van  $Q_2$  val.
- Gebruik die klasmiddelpunt vir die waarde van  $Q_2$ .

**Modus**

Die klasinterval met die hoogste frekwensie.

<https://youtu.be/V1GqYnJu1o0>



**KAN JY?**

Bepaal die geskatte gemiddelde, mediaan en modale interval van die onderstaande data.

Klasinterval	Frekwensie
$0 \leq x \leq 10$	1
$10 < x \leq 20$	2
$20 < x \leq 30$	11
$30 < x \leq 40$	9
$40 < x \leq 50$	14
$50 < x \leq 60$	3

**Antwoorde:**

Geskatte gemiddelde 35,5

mediaan : 35

modale interval  $40 < x \leq 50$



Maatstawwe van verspreiding		
<p><b>Variasiewydte</b>                      Maksimum waarde – Minimum waarde.  <math>IQR = Q_3 - Q_1</math>  <b>Semi – Interkwartielvariasiewydte</b>  <math>= \frac{Q_3 - Q_1}{2}</math></p>	<p><b>Kwartiele</b></p> <ul style="list-style-type: none"> <li>• Verdeel geordende data in kwarte..</li> <li>• Onderste kwartiel (<math>Q_1</math>)</li> <li>• Mediaan (<math>Q_2</math>)</li> <li>• Boonste kwartiel (<math>Q_3</math>)</li> <li>• Die posisie van <math>Q_1 = \frac{1}{4}(n + 1)</math>.</li> <li>• Die posisie van <math>Q_3 = \frac{3}{4}(n + 1)</math>.</li> </ul>	<p><b>Persentiele</b></p> <ul style="list-style-type: none"> <li>• Verdeel geordende datastel in 'n 100 gelyke dele.</li> <li>• Die posisie van die <math>k^{ste}</math> persentiel = <math>\frac{k}{100}(n + 1)</math>.</li> </ul> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin-top: 10px;"> <a href="https://youtu.be/K3wsOqlqA6k">https://youtu.be/K3wsOqlqA6k</a> </div>
<p><b>Voorbeeld 3:</b>                      Bepaal die onderste kwartiel, boonste kwartiel, variasiewydte, interkwartielvariasiewydte, semi - interkwartielvariasiewydte en die 80<sup>ste</sup> persentiel van die onderstaande data. 20, 21, 25, 25, 26, 26, 26, 30, 32, 32, 33, 33, 35, 36, 37, 38, 41, 42, 43, 45</p>		
<p><b>Oplossing</b></p>		
<p><b>Onderste kwartiel (<math>Q_1</math>)</b>                      Die posisie van <math>Q_1 = \frac{1}{4}(n + 1)</math>  <math>= \frac{1}{4}(20 + 1) = 5,25^{de}</math> posisie.  <math>\therefore Q_1 = \frac{26 + 26}{2} = 26</math></p>	<p><b>Boonste kwartiel (<math>Q_3</math>)</b>                      Die posisie van <math>Q_3 = \frac{3}{4}(n + 1)</math>  <math>= \frac{3}{4}(20 + 1) = 15,75^{th}</math> posisie.  <math>\therefore Q_3 = \frac{37 + 38}{2} = 37,5</math></p>	<p><b>80<sup>ste</sup> persentiel</b>                      Die posisie van <math>80^{ste} = \frac{80}{100}(20 + 1) = 16,8^{ste}</math> posisie.  <math>\therefore 80^{ste}</math> persentiel = <math>\frac{38+41}{2} = 39,5</math></p>
<p><b>Variasiewydte</b> = <math>45 - 20 = 25</math></p>	<p><b>IQR</b> = <math>Q_3 - Q_1 = 37,5 - 26 = 11,5</math></p>	<p><b>Semi - IQR</b> = <math>\frac{Q_3 - Q_1}{2} = \frac{37,5 - 26}{2} = 5,75</math></p>
<p><b>KAN JY?</b>                      Bepaal die onderste kwartiel, boonste kwartiel, variasiewydte, interkwartielvariasiewydte, semi - interkwartielvariasiewydte en die 60<sup>ste</sup> persentiel van die onderstaande data. Wat impliseer die onderste kwartiel van die toets punte                      'n Groep van 21 leerders het n toets geskryf wat uit 50 was. Hulle uitslae was as volg:                      17; 8; 19; 9; 12; 28; 11; 16; 20; 14; 29; 23; 37; 23; 26; 4; 35; 26; 18; 45; 7</p>		
<p><b>Antwoorde</b>                      Onderste kwartiel: 11.5 Boonste kwartiel: 27 variasiewydte: 41 interkwartielvariasiewydte: 15.5 semi interkwartielvariasiewydte : 7.75                      60<sup>th</sup> persentiel: 23 Q1 -25% van die uitslae is minder as 11,5</p>		



**Die vyfgetal opsomming**

Die vyfgetal opsomming word voorgestel deur die volgende maatstawwe van verspreiding:

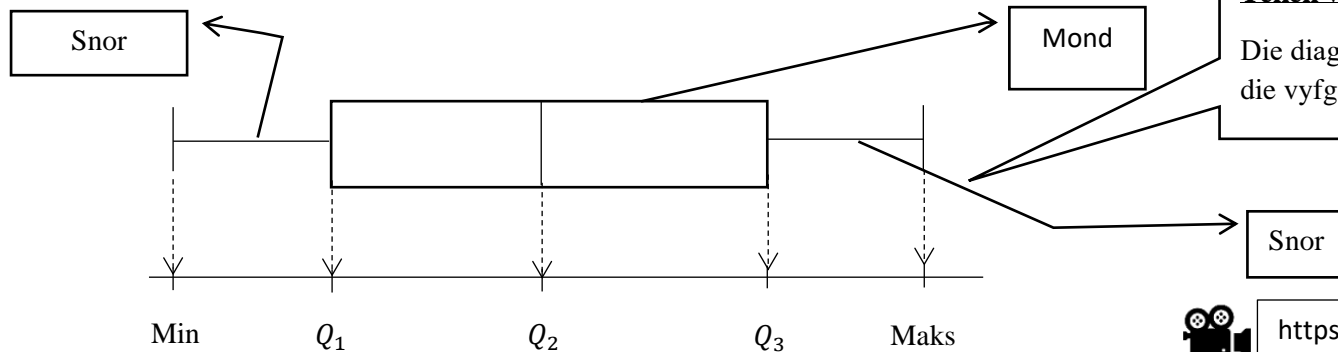
- ❖ Die minimum waarde van die datastel.
- ❖  $Q_1$
- ❖  $Q_2$
- ❖  $Q_3$
- ❖ Die maksimum waarde van die datastel.

**Uitskieters**

Dit is waardes wat baie groter of kleiner is as die gegewe data items

**Mond- en - snor diagram**

Die mond – en – snor diagram is die grafiese voorstelling van die vyfgetal opsomming.



**Teken wenke**

Die diagram moet altyd ‘n mond en ‘n snor hê. Toon die vyfgetal opsomming duidelik op die as aan.



[https://youtu.be/pSC8ToA9\\_no](https://youtu.be/pSC8ToA9_no)

**Belangrike gevolgtrekkings van af die mond – en – snor digram.**

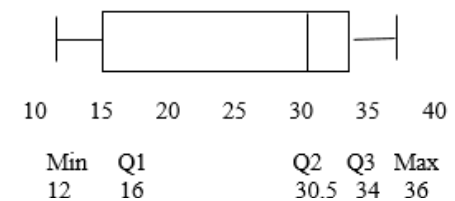
- 25% van die data lê tussen die minimum waarde en  $Q_1$ .
- 25% van die data lê tussen  $Q_1$  en  $Q_2$ .
- 25% van die data lê tussen  $Q_2$  en  $Q_3$ .
- 25% van die data lê tussen  $Q_3$  en die maksimum waarde.

**KAN JY?**

Teken ‘n mond – en – snor diagram van die onderstaande data.

12, 12, 13, 14, 16, 22, 23, 24, 30, 31, 31, 31, 32, 34, 35, 35, 36

**Antwoord**





### Verspreiding van data

#### Positief skeef/Skeef na regs

- 'n Hoë volume van die data is geleë by die laer waardes van die data stel.
- Die hoër waardes van die data stel is wyer verspreid.
- Gemiddelde > Mediaan.

#### Simmetriese verspreiding

- 'n Hoë volume data rond om gedie gemiddelde geleë.
- Die data is simmetries versreid.
- Gemiddelde = Mediaan.

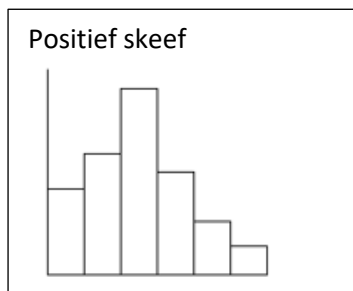
#### Negative skew/Skewed to the left

- 'n Hoë volume van die data is geleë by die hoër waardes van die data stel.
- Die laer waardes van die data stel is wyer verspreid.
- Gemiddelde < Median.

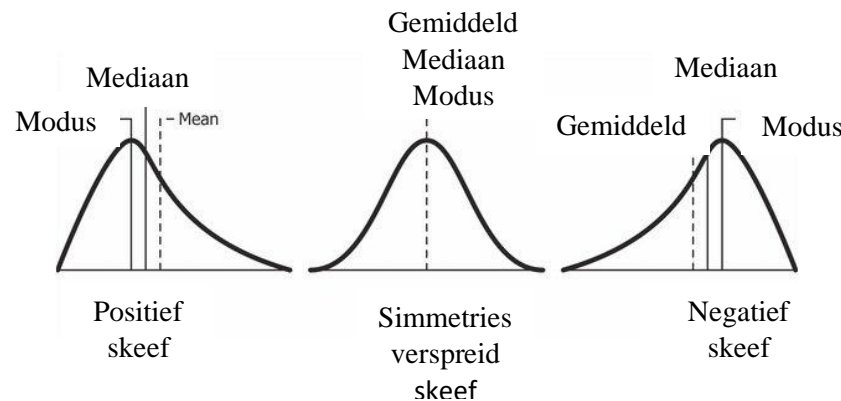
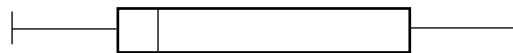
### Grafiese voorstelling van verspreiding



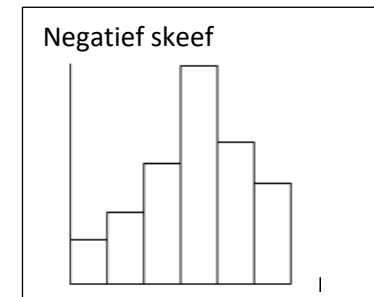
[https://youtu.be/pfujiA5Mk\\_U](https://youtu.be/pfujiA5Mk_U)



Positief skeef



Simmetriese verspreiding



Negatief skeef



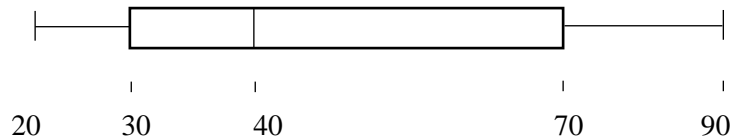
#### Belangrike gevolgtrekkings:

Skeefheid beïnvloed die gemiddelde. Hoe skewer die data, hoe minder kan die gemiddelde gebruik word vir sentrale neiging. Indien die data skeef na links is, beteken dit dat die gemiddelde te laag is. Indien die data skeef na regs is, beteken dit dat die gemiddelde te hoog is. Die beste manier om sentrale neiging te toets is om die mediaan te gebruik omdat dit nie verander nie, al verander die skeefheid van die data.



**Voorbeeld 4:**

Die mond – en – snor diagram vir punte (uit 90) vir nege leerders word in onderstaande diagram getoon.



**Beantwoord die volgende vrae.**

1. Lewer kommentaar oor die skeefheid van die data.
2. Hoekom sal die gemiddelde nie 'n goeie maatstaf van sentrale neiging wees nie?
3. Dink jy die toets was te maklik of te moeilik? Gee 'n rede vir jou antwoord. answer.

**Oplossing:**

1. Skeef na regs of positief skeef.
2. Die gemiddelde was te hoog.
3. Die toets was te moeilik want 50% van leerders het minder as 40% gekry.

**KAN JY?**

Die onderstaande data toon die aantal besoeke aan die plaaslike kliniek vir die afgelope 6 dae : 12, 29, 13, 63, 12, 3

1. Teken 'n mond – en – snor diagram.
2. Gebruik jou sakrekenaar om die gemiddelde van die data te bepaal.
3. Bereken die gemiddelde – mediaan.
4. Lewer kommentaar oor die skeefheid van die data.

**Antwoorde**

2. gemiddelde  $\approx 14,1$
3. gemiddelde - mediaan =  $14,1 - 17,5 = -3,4$
4. Die data is nouer gegroepeer aan die regterkant en meer verspreid aan die linkerkant.. Dit is negatief skeef. Die verskil tussen die gemiddelde en die mediaan is negatief wat aandui dat die data negatief skeef is.

**Kumulatiewe frekwensie grafiek (ogief)**

**Definisie:**

Dit is die totale frekwensie en alle frekwensies tot dusver in 'n frekwensie verspreiding. Dit is die “lopende totaal” van frekwensies.

**Wat moet ek kan doen?**

- Gebruik die frekwensietabel om 'n ogief te teken.
- Bepaal waardes deur gebruik te maak van 'n ogief.
- Gebruik 'n ogief om die frekwensietabel te voltooi.
- Gebruik die ogief om persentiele en kwartiele te bereken.

**Wat is belangrik wanneer 'n ogief geteken word?**

1. Dit moet 'n opskrif hê.
2. Die asse moet benoem word.
3. Plot die punte. (Boonste grens van klas; kumulatiewe frekwensie van klas)
4. ANKER die ogief by (onderste grens van die eerste interval; 0)

<https://youtu.be/gDPo0R27wzg>



Wanneer jy die kurwe skets mag jy nie 'n liniaal gebruik nie. Dit is 'n S – vormige



**Voorbeeld 5:**

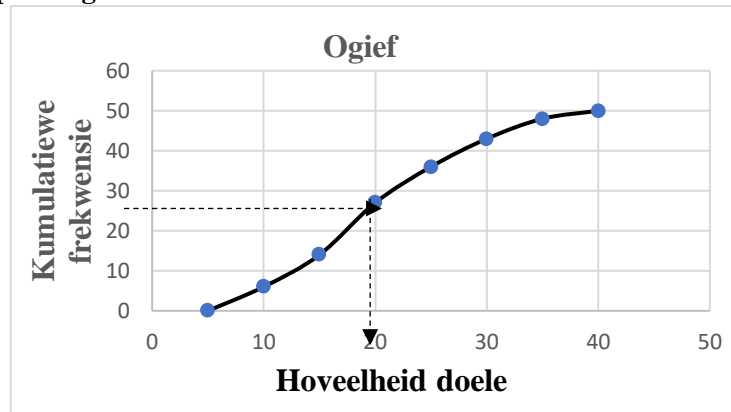
Die volgende frekwensietabel toon die hoeveelheid doele wat die Protea netbalspan aangeteken het in die 50 wedstryde wat hulle gespeel het in 2018.

Hoeveelheid doele	Frekwensie	Kumulatiewe frekwensie	Koördinate
$5 < x \leq 10$	6	6	(10 ; 6)
$10 < x \leq 15$	8	14	(15 ; 14)
$15 < x \leq 20$	13	27	(20 ; 27)
$20 < x \leq 25$	9	36	(25 ; 36)
$25 < x \leq 30$	7	43	(30 ; 43)
$30 < x \leq 35$	5	45	(35 ; 45)
$35 < x \leq 40$	2	50	(40 ; 50)

1. Teken die kumulatiewe frekwensie grafiek(ogief) van die gegewe data.
2. Gebruik die grafiek om die mediaan te bepaal.

**Oplossing:**

1.



2. Posisie van  $Q_2 = \frac{1}{2}(n + 1) = \frac{1}{2}(50 + 1) = 25,5^{ste}$  waarde.  $\therefore Q_2 = 19$

**Riglyne om 'n ogief te skets.**

1. Skep 'n kumulatiewe frekwensie kolom.
2. Skep 'n koördinate kolom.
3. Anker ogief by (onderste grens; 0).
4. Plot punte in koördinate kolom.
5. Konnekteer die punte met gladde kurwe.

**Hoe word kwartiele en persentiele grafies bepaal.**

1. Vind die posisie.
2. Vind die posisie op die kumulatiewe frekwensie as.
3. Teken 'n gebroke lyn na die ogief.
4. Teken 'n gebroke lyn van die ogief na die  $x -$  as.
5. Lees die waarde van die  $x -$  as af.

**KAN JY?**

'n Steekproef is gedoen op 50 mense om die afstande te bepaal wat hulle daaglik na hul werk reis. Die volgende tabel toon die resultate van die steekproef:

Afstand in km	Frekwensie
$0 < d \leq 5$	2
$5 < d \leq 10$	7
$10 < d \leq 15$	4
$15 < d \leq 20$	13
$20 < d \leq 25$	16
$25 < d \leq 30$	8

1. Teken die kumulatiewe frekwensie grafiek(ogief) van die gegewe data.
2. Gebruik die grafiek om die 60<sup>ste</sup> persentiel te bepaal.
3. Gebruik jou grafiek om die geskatte aantal mediaan kilometers wat gereis was per week te bepaal.

**Antwoorde:** (2)  $\approx 54$  (3) Mediaan tussen 25ste en 26steleerderr  $\approx 50$  km per week






**Standardafwyking en Variansie**

Standaardafwyking( $\sigma$ ) is 'n getal wat wys met hoe die data om die van die **gemiddelde** van die data verspreid is. Hoe groter die waarde van die standaardafwyking, hoe verder lê die data van die gemiddelde ( $\bar{x}$ ). Dit gee 'n aanduiding dat die gemiddelde nie 'n betroubare maatstaf van sentrale neiging is nie, en dat die mediaan meer geskik sal wees. Hoe kleiner die standaardafwyking is, hoe nader lê die data aan die gemiddelde. Dit wil sê dat die gemiddelde 'n betroubare maatstaf van sentrale neiging sal wees. Die variansie is die kwadraat van die standaardafwyking ( $\sigma^2$ ). Die interval van data wat  $k$  standaardafwykings van die gemiddelde is ( $\bar{x} - k\sigma; \bar{x} + k\sigma$ ).

**Hoe om standaardafwyking te bepaal met jou casio sakrekenaar.**

1. Druk "mode" en selekteer STAT.
2. Selekteer "1 - VAR".
3. Voer waardes in. Na elke individuele waarde ingesleutel is druk "=" voor die volgende waarde ingesleutel word.
4. Na alle individuele waardes ingesleutel is druk "AC" 
5. Om  $\sigma$  te vind druk "SHIFT STAT(by 1)".
6. Druk dan "Var" om  $\sigma$  te vind.

[https://youtu.be/iW1e\\_FJeYS](https://youtu.be/iW1e_FJeYS)

**Bereken standaardafwyking met 'n formule.**



[https://youtu.be/ueWU\\_nzyPg](https://youtu.be/ueWU_nzyPg)

**Voorbeeld 6:**

Die volgende steekproef van weeklikse lone van tien mense wat vir 'n drukkers maatskappy werk, is geneem.

R2 250; R2 250; R3 000; R3 300; R3 300  
R3 600; R3 900; R4 350; R4 350; R5 250

1. Bereken die gemiddelde weeklikse loon.
2. Bereken die standaardafwyking van die weeklikse lone.
3. Bepaal die persentasie mense wat binne EEN standaardafwyking vanaf die gemiddelde lê.

**Oplossing:**

1.  $\bar{x} = \frac{\sum x}{n} = \frac{35550}{10} = 3555$

2.  $\sigma = 900,12$

3.  $(\bar{x} - k\sigma; \bar{x} + k\sigma)$   
 $= (3555 - 900,12; 3555 + 900,12)$   
 $= (2654,88; 4455,12)$

$\therefore 7$

$\therefore \frac{7}{10} \times 100 = 70\%$  van mense lê binne een standaardafwyking.

**KANJY?**

Die volgende toets uitslae(uit 40) was aangeteken:

12; 19; 40; 33; 25; 24; 15; 38

1. Bereken die standaardafwyking van die toets uitslae.
2. Bepaal hoeveel toets uitslae buite EEN standaardafwyking van af die gemiddl

**Antwoorde:** 1. 9,74      2. 4



**Konsolidasie**

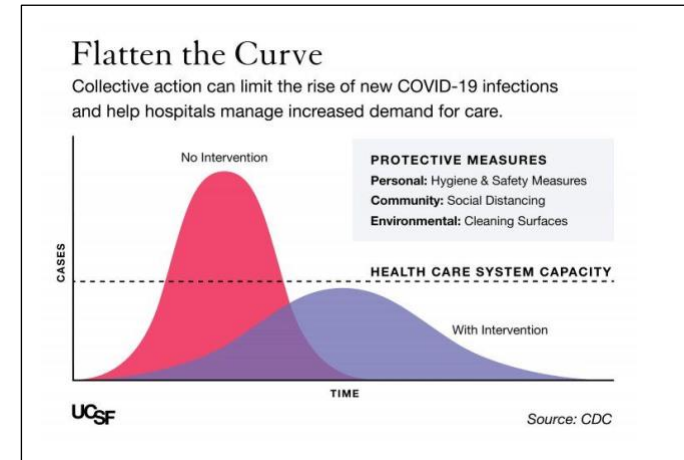
- Om grafieke van gegewe data te teken en statistiese konsepte te verstaan is baie belangrik maar ook...
- Om grafieke te analiseer en kommentaar te lewer oor die skeefheid van data is n belangrike vaardigheid wat sal bydra tot interpretasie van data.

**Flatten the curve !!! Maak die kurwe platter !!Wat beteken dit?**

Die "curve / kurwe " waaroor navorsers praat verwys na die geprojekteerde aantal mense wat COVID 19 oor 'n tydperk sal opdoen.

(Onthou dit is nie n definitiewe voorspelling van presies hoeveel mense die virus sal opdoen nie, maar 'n teoretiese getal wat gebruik word om die verspreiding van die virus voor te stel.)

Die bedoeling van die maak-die-kurwe-platter benadering is om die opeenhoping van pasiënte wat die mediese vermoë van die land kan oorlaai, te vermy deur infeksies oor tyd uit te rek ten einde binne die land se kapasiteit tot mediese versorging te bly.



<b>AKTIWITEITE</b>	Mind Action Series Hersienings oefening bl. 288 – bl. 291	Via Africa Gemengde oefening bl. 106	Classroom Mathematics Hersienings oefeninge bl 304 - 305	Siyavula hersienings oefening bl 370 - 372	Platinum hersienings oefening bl 242
--------------------	---	--	--	--	--