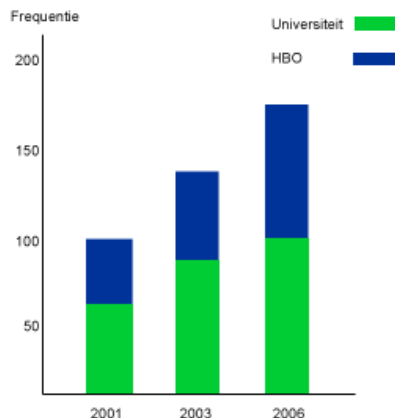
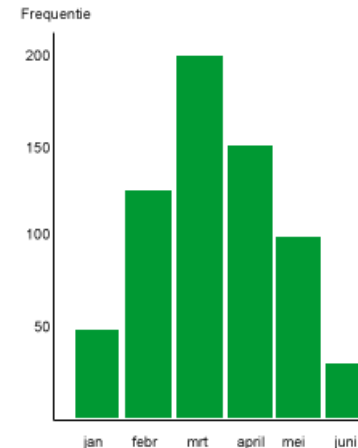


5.0 Voorkennis

Er zijn verschillende manieren om gegevens op een grafische wijze weer te geven:

1. Staafdiagram:

- De lengte van de staven komt overeen met de hoeveelheid;
- De staven staan meestal los van elkaar;
- De volgorde van de staven hoeft niet altijd van belang te zijn.



2. Stapeldiagram (Samengesteld staafdiagram):

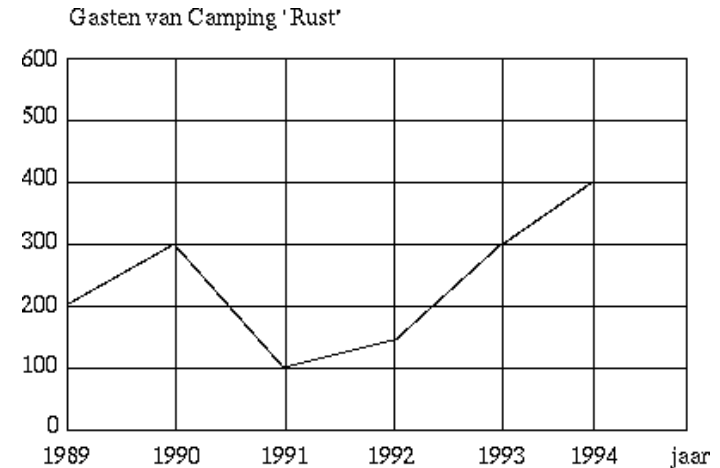
- De lengte van de staven komt overeen met de hoeveelheid;
- De staven staan meestal los van elkaar;
- De volgorde van de staven hoeft niet altijd van belang te zijn.

5.0 Voorkennis

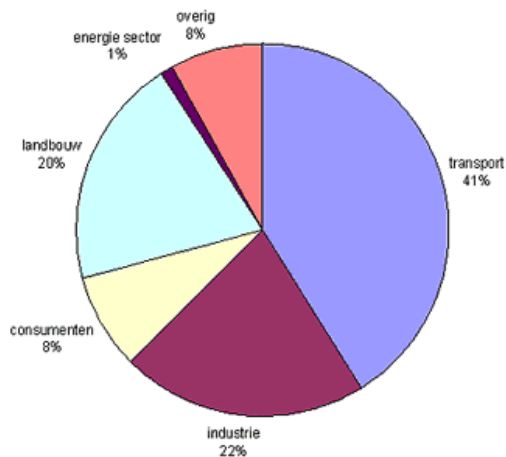
Er zijn verschillende manieren om gegevens op een grafische wijze weer te geven:

3. Lijndiagram

- Langs de horizontale as staat meestal de tijd;
- De opeenvolgende punten zijn verbonden door lijnstukken;
- Tussenliggende punten hebben geen betekenis;
- Je kunt indien nodig op de verticale as een scheurlijn gebruiken.



Bijdrage van de belangrijkste bronnen aan de emissie van fijn stof in Nederland (2002)

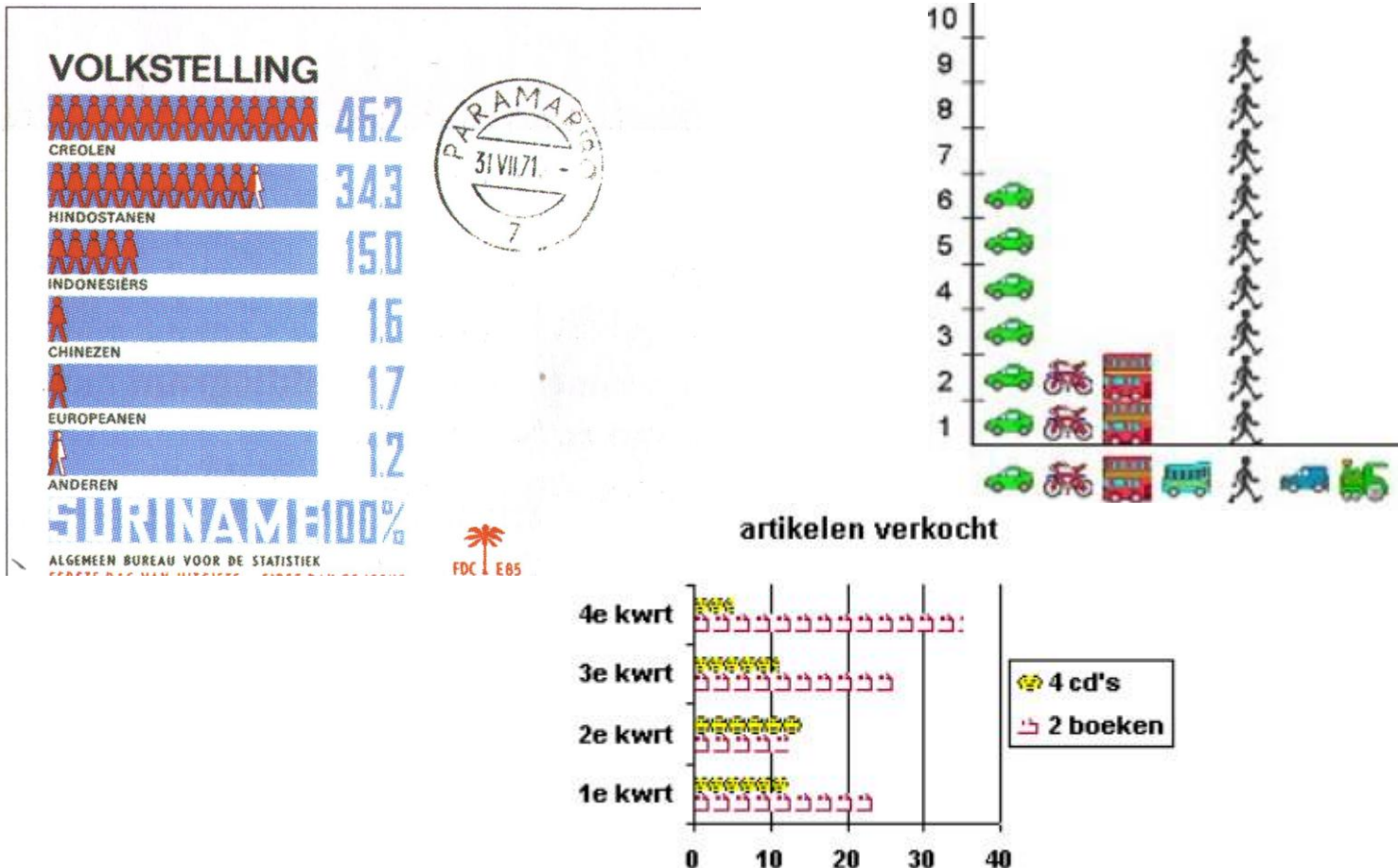


4. Cirkeldiagram

- Een cirkeldiagram brengt de relatieve verdeling in beeld;
- Bij een aandeel van $p\%$ hoort een sector met een hoek van $0,01p \cdot 360^\circ$.

5.0 Voorkennis

Gegevens kunnen ook weergegeven worden met een beelddiagram.
Beelddiagrammen kunnen er verschillend uit zien:



5.0 Voorkennis

Voorbeeld:

Defecten per week	1	2	3	4
Frequentie	14	16	15	18

Een **frequentietabel** geeft aan hoe vaak een bepaalde gebeurtenis zich voordoet.

De **totale frequentie** is $14 + 16 + 15 + 18 = 62$.

In de tabel zijn alle mogelijke gebeurtenissen weergegeven. Er is nu sprake van een **frequentieverdeling**.

5.0 Voorkennis

De onderstaande **frequentietabel** geeft aan hoeveel auto's er in een bepaald uur in een straat geteld zijn.

Aantal auto's per uur	15	16	17	18	19	20	21
frequentie	2	7	9	15	7	4	1

Het **waarnemingsgetal** 18 heeft de **frequentie** 15. Dit betekent dat het 15 keer is voorgekomen dat er in deze straat 18 auto's in een uur zijn geteld.

Je kunt het **gemiddelde** aantal auto's per uur nu als volgt berekenen:

1. Bereken de totale frequentie (het totaal aantal waarnemingen).

$$\text{Totale frequentie} = 2 + 7 + 9 + 15 + 7 + 4 + 1 = 45$$

2. Vermenigvuldig elk waarnemingsgetal met de bijbehorende frequentie en tel dit op.

$$2 \cdot 15 + 7 \cdot 16 + 9 \cdot 17 + 15 \cdot 18 + 7 \cdot 19 + 4 \cdot 20 + 1 \cdot 21 = 799$$

3. Deel de uitkomst van stap 2 door de uitkomst van stap 1

$$\text{Het gemiddelde aantal auto's per uur is } 799/45 \approx 17,76$$

5.0 Voorkennis

Voorbeeld:

Gegeven is de volgende rij getallen:

4 7 10 3 6 125 6 3 9

In deze rij zit één getal, dat erg afwijkt van de rest (het getal 125). Wanneer je nu het gemiddelde uit zou rekenen, zou dit erg vertekenen, door dit ene erg afwijkende getal.

Mediaan = Het middelste getal van alle waarnemingsgetallen nadat deze naar grootte zijn gerangschikt. Bij een even aantal waarnemingsgetallen is de mediaan het gemiddelde van de middelste twee getallen.

De getallen op volgorde zetten geeft; 3 3 4 6 6 7 9 10 125

Het middelste waarnemingsgetal is hier het 5^{de} getal. De mediaan is 6.

Modus = het waarnemingsgetal met de grootste frequentie. Zijn er twee of meer getallen met dezelfde grootste frequentie dan is er geen modus.

De getallen 3 en 6 komen allebei twee keer voor. Deze rij heeft dus geen modus.

5.0 Voorkennis

Voorbeeld:

De onderstaande tabel geeft aan hoe vaak de kinderen uit groep 4 van een lagere school het afgelopen jaar naar het zwembad zijn gegaan.

aantal	0	1	2	3	4	5	6
frequentie	5	7	12	15	13	8	6

Bereken het gemiddelde, de modus en de mediaan (de **centrummaten**)

Berekenen gemiddelde:

$$\text{Totale frequentie} = 4 + 7 + 12 + 15 + 13 + 8 + 6 = 66$$

$$4 \times 0 + 7 \times 1 + 12 \times 2 + 15 \times 3 + 13 \times 4 + 8 \times 5 + 6 \times 6 = 204$$

$$\text{Gemiddelde} = 204/66 \approx 3,09$$

5.0 Voorkennis

Voorbeeld:

De onderstaande tabel geeft aan hoe vaak de kinderen uit groep 4 van een lagere school het afgelopen jaar naar het zwembad zijn gegaan.

aantal	0	1	2	3	4	5	6
frequentie	5	7	12	15	13	8	6

Modus = 3 (Dit getal komt het meeste voor)

Mediaan:

Getallen 1 t/m 5: 0

Getallen 12 t/m 24: 2

Getallen 39 t/m 52: 4

Getallen 61 t/m 66: 6

Getallen 5 t/m 12: 1

Getallen 24 t/m 39: 3

Getallen 52 t/m 60: 5

Er is een even aantal getallen. Het 33^{ste} getal is 3. het 34^{ste} getal is 3.

De mediaan is nu: $(3 + 3)/2 = 3$.

5.1 Frequentieverdelingen [1]

Kwantitatieve gegevens: gegevens, die een hoeveelheid uitdrukken;
Gewicht, salaris, aantal kinderen, leeftijd.

Kwalitatieve gegevens: gegevens, die een kenmerk (en dus geen hoeveelheid) uitdrukken.
Bloedgroep, geslacht, merk auto, soort woning.

Voorbeeld:

Defecten per week	1	2	3	4
Frequentie	14	16	15	18

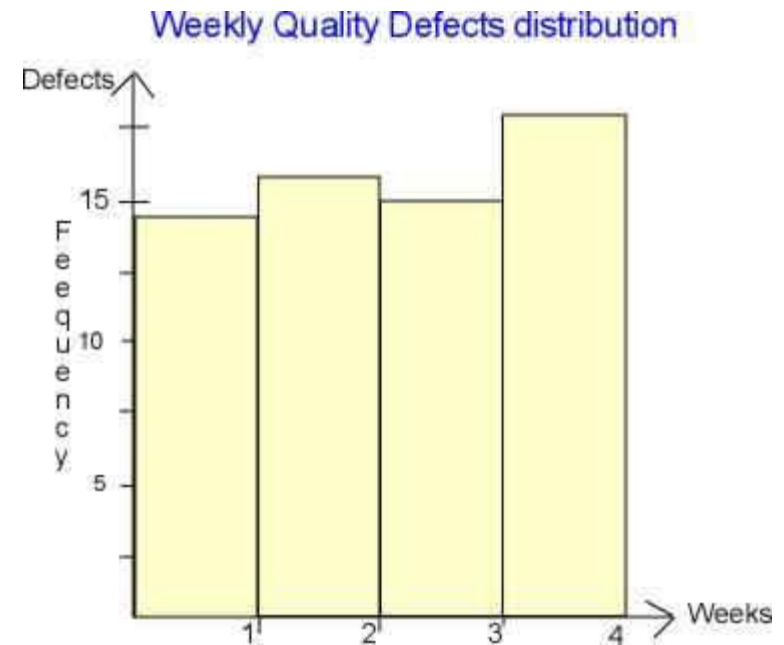
Een **frequentietabel** geeft aan hoe vaak een bepaalde gebeurtenis zich voordoet.
De totale **frequentie** is $14 + 16 + 15 + 18 = 62$.
In de tabel zijn alle mogelijke gebeurtenissen weergegeven. Er is nu sprake van een **frequentieverdeling**.

5.1 Frequentieverdelingen [1]

Defecten per week	1	2	3	4
Frequentie	14	16	15	18

De gegevens uit de frequentietabel staan in het **histogram** rechts.

Een histogram is een staafdiagram bij een frequentietabel met kwantitatieve gegevens op de horizontale as. De staven liggen tegen elkaar aan.

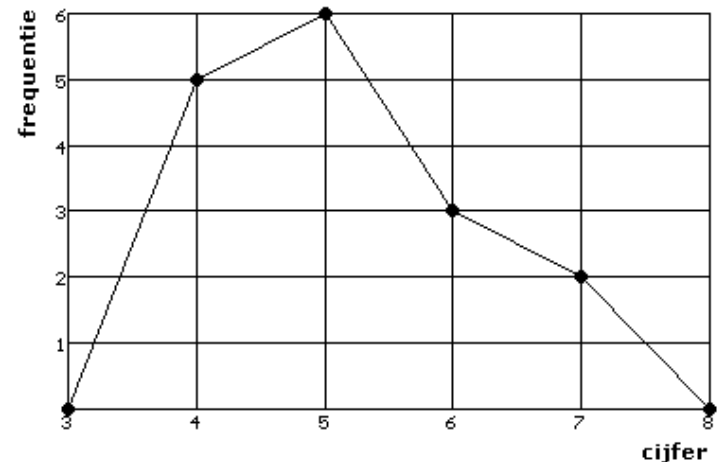


5.1 Frequentieverdelingen [1]

Cijfer	3	4	5	6	7	8
Frequentie	0	5	6	3	2	0

De gegevens uit de frequentietabel staan in de **frequentiepolygoon** rechts.

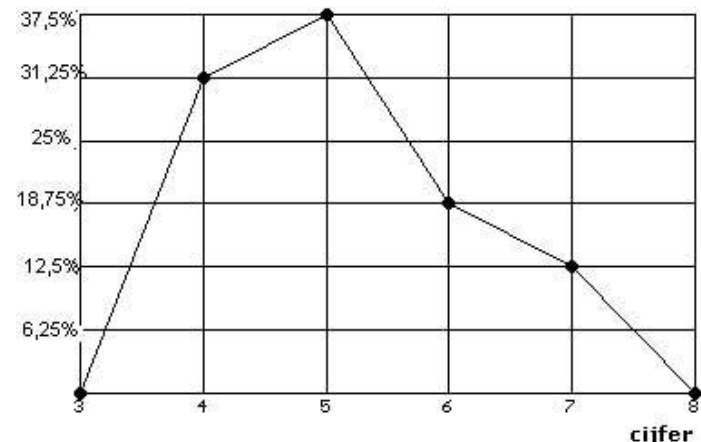
Er is nu gebruik gemaakt van **absolute frequenties**.



Dit is een **relatieve-frequentiepolygoon**.

Er is nu gebruik gemaakt van **relatieve frequenties**.

$$\text{Relatieve frequentie} = \frac{\text{frequentie}}{\text{tot. frequentie}} \cdot 100\%$$



5.1 Frequentieverdelingen [2]

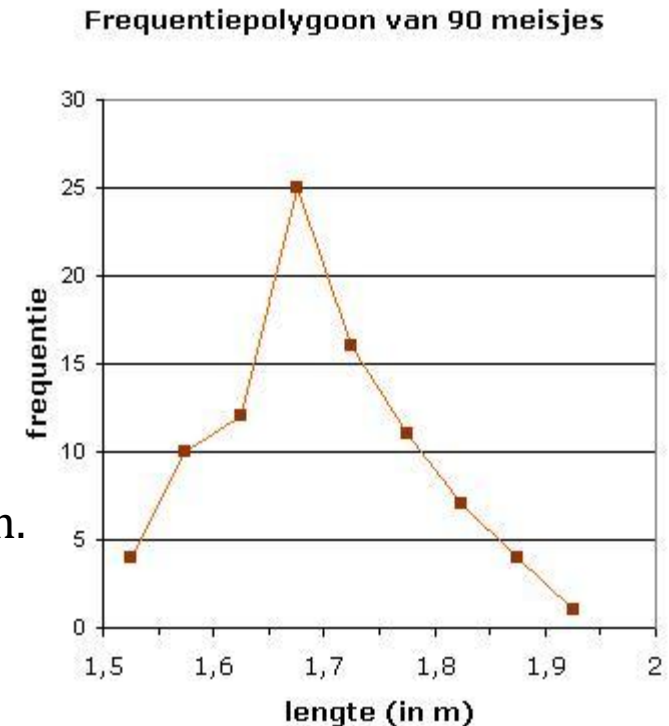
Lengte (m)	1,50 - < 1,55	1,55 - < 1,60	1,60 - < 1,65	1,65 - < 1,70	1,70 -< 1,75	1,75 - < 1,80	1,80 - < 1,85	1,85 - < 1,90	1,90 -< 1,95
Freq.	4	10	12	25	16	11	7	4	1

Wanneer je een verzameling gegevens hebt, waarin vaak de frequentie 1 voorkomt, ga je de gegevens indelen in **klassen**.

In dit voorbeeld is er een **klassenindeling** gemaakt met een **klassenbreedte** van 0,05 meter.

Van de klassen 1,55 -< 1,60 zijn 1,55 en 1,60 de **klassengrenzen**.

De gegevens zijn weergegeven in een frequentiepolygoon. De punten staan nu steeds in het midden van de klasse.



5.1 Frequentieverdelingen [3]

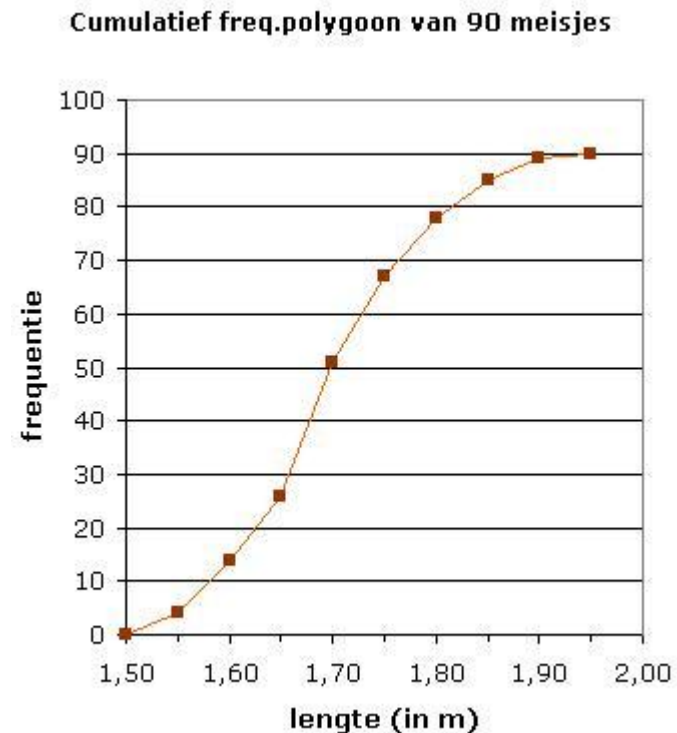
Klasse	Frequentie	Cumulatieve frequentie	Relatieve cumulatieve frequentie
1,50 -< 1,55	4	4	4,44%
1,55 -< 1,60	10	14 (10 + 4)	15,56%
1,60 -< 1,65	12	26 (14 + 12)	28,89%
1,65 -< 1,70	25	51	56,67%
1,70 -< 1,75	16	67	74,44%
1,75 -< 1,80	11	78	86,67%
1,80 -< 1,85	7	85	94,44%
1,85 -< 1,90	4	89	98,89%
1,90 -< 1,95	1	90	100%

- De **cumulatieve frequentie** van een klasse krijg je door de frequentie van deze klasse en alle voorgaande klassen op te tellen;
- De **relatieve cumulatieve frequentie** van een klasse krijg je door de cumulatieve frequentie van deze klasse te delen door het aantal waarnemingen $\cdot 100\%$.

5.1 Frequentieverdelingen [3]

Als je de cumulatieve frequenties in een polygoon tekent, krijg je een **cumulatief frequentiepolygoon**.

- Op de horizontale as staan de klassen;
- Boven de rechtergrens van elke klasse wordt de cumulatieve frequentie gezet;
- Boven de linkergrens van de eerste klasse wordt de cumulatieve frequentie 0 gezet;
- Verbind de punten door lijnstukjes.



Als je de relatieve cumulatieve frequenties in een polygoon tekent, krijg je een **relatief cumulatief frequentiepolygoon**.

5.2 Centrummaten en variabelen [1]

Kwantitatieve gegevens: gegevens, die een hoeveelheid uitdrukken;
Gewicht, salaris, aantal kinderen, leeftijd.

Een **kwantitatieve variabele** wordt uitgedrukt in een getal.
Het verschil tussen twee van deze opeenvolgende getallen heeft een eenduidige betekenis.

Kwalitatieve gegevens: gegevens, die een kenmerk (en dus geen hoeveelheid) uitdrukken.
Bloedgroep, geslacht, merk auto, soort woning.

Een **kwalitatieve variabele** wordt vaak niet uitgedrukt in een getal.
Is dit toch het geval, dan heeft het verschil tussen twee van deze opeenvolgende getallen geen eenduidige betekenis.

5.2 Centrummaten en variabelen [1]

Gegeven zijn de volgende 10 waarnemingsgetallen: 1, 3, 3, 3, 4, 5, 6, 8, 8, 9

Het **gemiddelde** is:
$$\frac{1 + 3 + 3 + 3 + 4 + 5 + 6 + 8 + 8 + 9}{10} = \frac{50}{10} = 5$$

De **mediaan** is het middelste waarnemingsgetal als de getallen naar grootte Geordend zijn. Bij een even aantal waarnemingsgetallen is dit het gemiddelde van de middelste twee getallen.

De **mediaan** is:
$$\frac{4 + 5}{2} = 4,5$$

De **modus** is het waarnemingsgetal dat het meeste voorkomt. Komen meerdere waarnemingsgetallen het meeste voor dan is er geen modus.

De **modus** is: 3

5.2 Centrummaten en variabelen [2]

Branduren	Frequentie
800 -< 900	6
900 -< 1000	15
1000 -< 1100	26
1100 -< 1200	21
1200 -< 1300	7

Voorbeeld: Bereken het gemiddelde aantal branduren

Stap 1:

Noteer de klassenmiddens: 850, 950, 1050, 1150, 1250

Stap 2:

Bereken de **totale frequentie**: $6 + 15 + 26 + 21 + 7 = 75$

Stap 3:

Bereken het gemiddelde:
$$\frac{6 \cdot 850 + 15 \cdot 950 + 26 \cdot 1050 + 21 \cdot 1150 + 7 \cdot 1250}{75} \approx 1060,67$$

5.2 Centrummaten en variabelen [2]

Branduren	Frequentie
800 -< 900	6
900 -< 1000	15
1000 -< 1100	26
1100 -< 1200	21
1200 -< 1300	7

Voorbeeld: Bereken de modale klasse en de mediaan.

De modale klasse is de klasse die het meeste voorkomt.
Dit is de klasse 1000 -< 1100.

Er zijn 75 getallen. De mediaan is het 38^e getal.

$6 + 15 = 21$ EN $6 + 15 + 26 = 47$. De mediaan ligt in de klasse 1000 -< 1100.
De mediaan is het 17^e getal van deze klasse.

$$\text{Mediaan} = 1000 + \frac{17}{26} \cdot 100 \approx 1065,38$$

5.2 Centrummaten en variabelen [2]

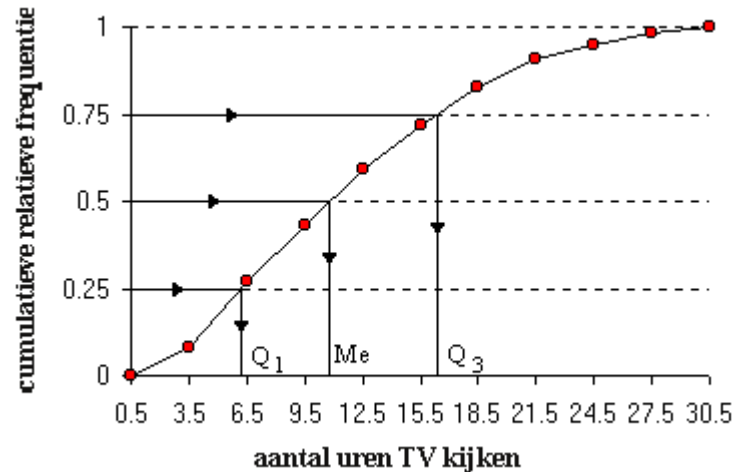
Continue variabele:

- Alle waarden zijn mogelijk
- Bv.: Lengte van mannen, Gewicht van vrouwen, alles wat normaal verdeeld is;

Discrete variabele X:

- Alleen een aantal losse waarden zijn mogelijk;
- Bv.: Aantal auto's op een weg per minuut, de schoenmaat van volwassenen, het aantal branduren van een lamp.

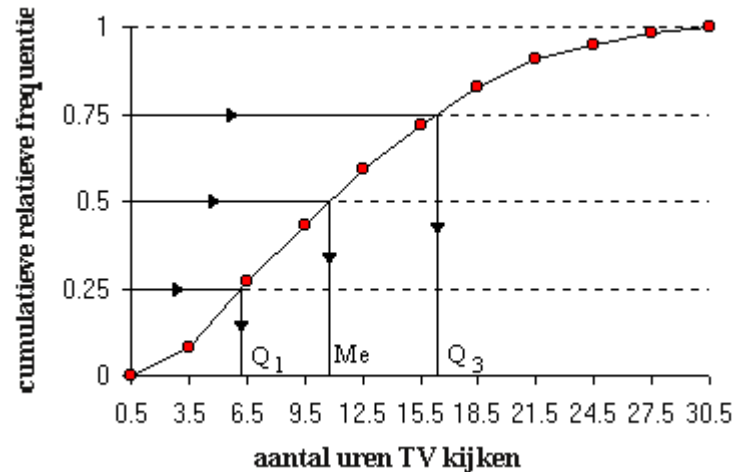
5.3 Spreidingsmaten [1]



In het plaatje is een **relatieve cumulatieve frequentiepolygoon** getekend van het aantal uren dat door een groep mensen tv gekeken wordt.

De **mediaan** (of Q_2) is het middelste waarnemingsgetal. Dit is te vinden bij 50% van de waarnemingsgetallen. Dit is hier rond de 10,5.

5.3 Spreidingsmaten [1]

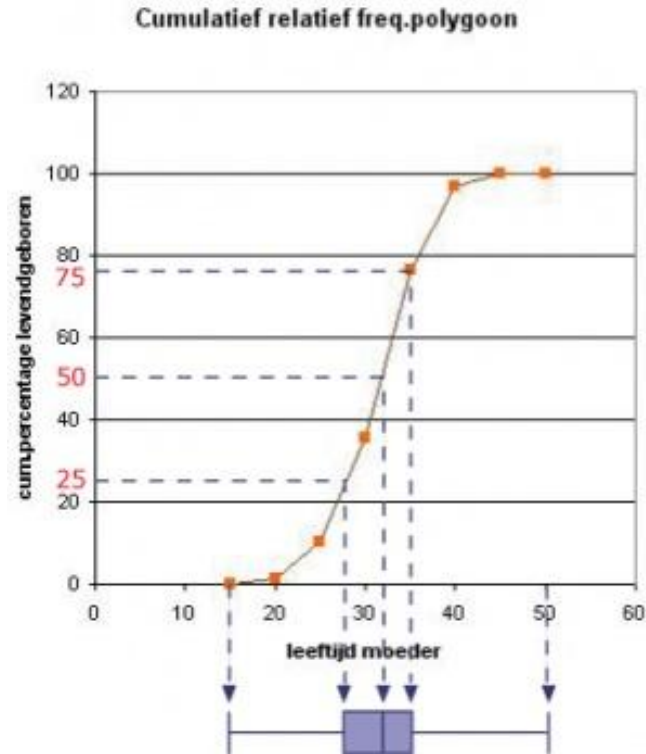


Het **eerste kwartiel** (Q_1) is het waarnemingsgetal dat op één kwart ligt. Dit is te vinden bij 25% van de waarnemingsgetallen. Dit is hier rond de 6.

Het **derde kwartiel** (Q_3) is het waarnemingsgetal dat op drie kwart ligt. Dit is te vinden bij 75% van de waarnemingsgetallen. Dit is hier rond de 16.

Mediaan, eerste kwartiel en derde kwartiel kun je schematisch weergeven in een **boxplot**.

5.3 Spreidingsmaten [1]



In het boxplot zie je het eerste kwartiel (Q_1), de mediaan (MED), het derde kwartiel (Q_3) en het eerste en laatste waarnemingsgetal.

5.3 Spreidingsmaten [2]

Voorbeeld:

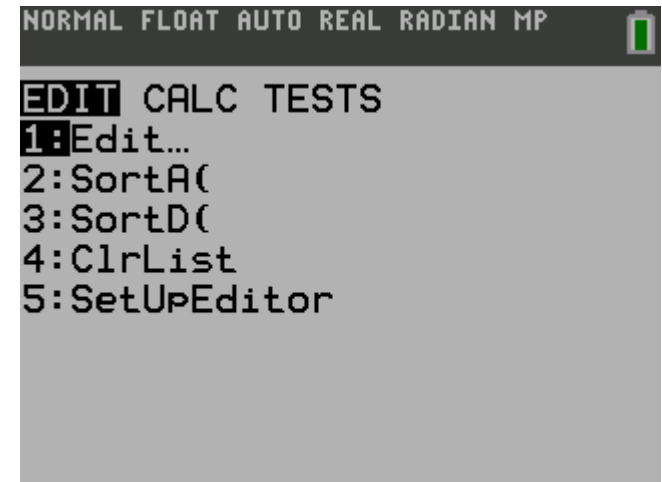
Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

Maak een boxplot bij deze verdeling:

Stap 1:

Voer de afstanden in je GR in.

- **STAT | EDIT 1:EDIT | ENTER**



5.3 Spreidingsmaten [2]

Voorbeeld:

Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

Maak een boxplot bij deze verdeling:

Step 1:

- Controleer of L1 leeg is.
- L1 niet leeg. Ga op L1 staan **CLEAR | ENTER**

L1	L2	L3	L4	L5	1
0	-----	-----	-----	-----	
1					
2					
3					
4					
5					
6					

L1(8)=

[illegible]

Ga onder L1 staan en voer de afstanden $\{0, 1, 2, 3, 4, 5, 6\}$ in

5.3 Spreidingsmaten [2]

Voorbeeld:

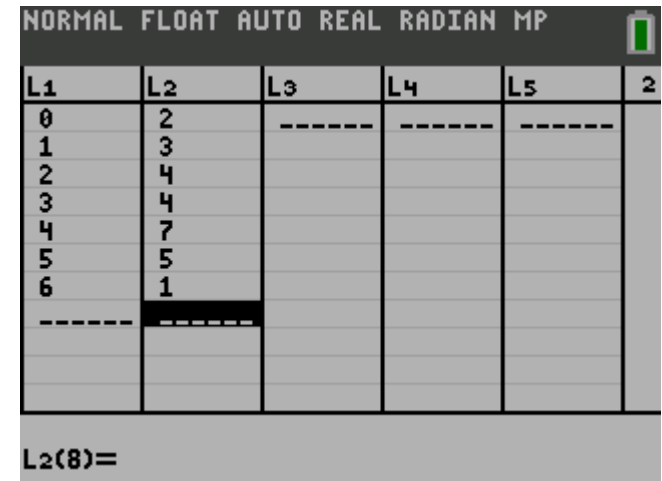
Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

Maak een boxplot bij deze verdeling:

Step 2:

Voer de frequenties in je GR in:

- Controleer of L2 leeg is.
- L2 niet leeg. Ga op L2 staan **CLEAR | ENTER**
- Ga onder L2 staan en voer frequenties $\{2, 3, 4, 4, 7, 5, 1\}$ in.



5.3 Spreidingsmaten [2]

Voorbeeld:

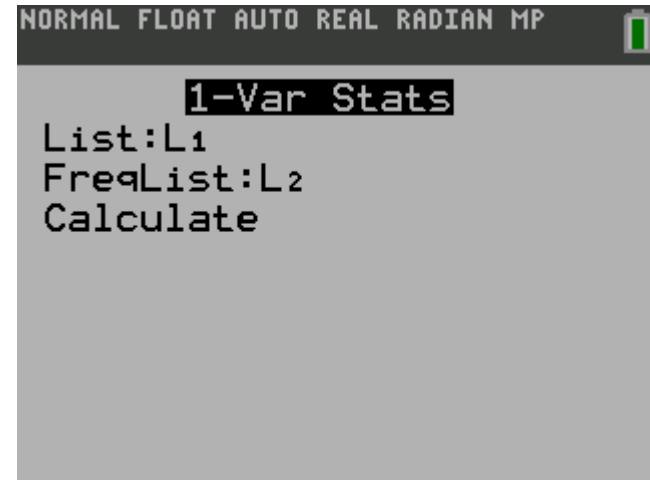
Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

Maak een boxplot bij deze verdeling:

Stap 3:

Bereken de benodigde waarden voor het boxplot:

- **STAT | CALC 1:1-Var Stats | ENTER**
- Vul bij **List**, lijst L1 met waarnemingsgetallen in;
- Vul bij **FreqList**, Lijst L2 met frequenties in.
- Op ENTER drukken geeft de benodigde waarden.



Let op:

- Als L1 al ingevuld staat bij List kun je dit gewoon laten staan.
- Invullen van L2 bij FreqList gaat via: 2ND | 2.

5.3 Spreidingsmaten [2]

Voorbeeld:

Stap 4:

Lees nu de benodigde gegevens af:

MinX (Minimale waarde) = 0;

Q_1 (Eerste kwartiel) = 2

Med (Mediaan) = 3,5;

Q_3 (Derde kwartiel) = 4

MaxX (Maximale waarde) = 6

$\bar{x}$ geeft het gemiddelde (3,15)

σ_x geeft de standaardafwijking (1,63)

```
NORMAL FLOAT AUTO REAL Radian MP
1-Var Stats
x̄=3.153846154
Σx=82
Σx²=328
Sx=1.665948563
σx=1.633596968
n=26
minX=0
↓Q1=2
```

```
NORMAL FLOAT AUTO REAL Radian MP
1-Var Stats
↑Sx=1.665948563
σx=1.633596968
n=26
minX=0
Q1=2
Med=3.5
Q3=4
maxX=6
```

5.3 Spreidingsmaten [2]

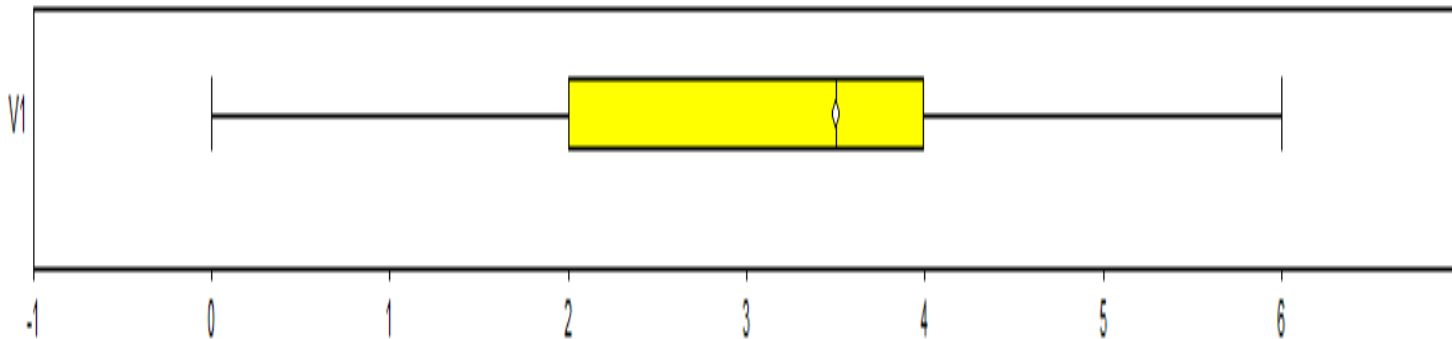
Voorbeeld:

Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

MinX (Minimale waarde) = 0; Q_1 (Eerste kwartiel) = 2

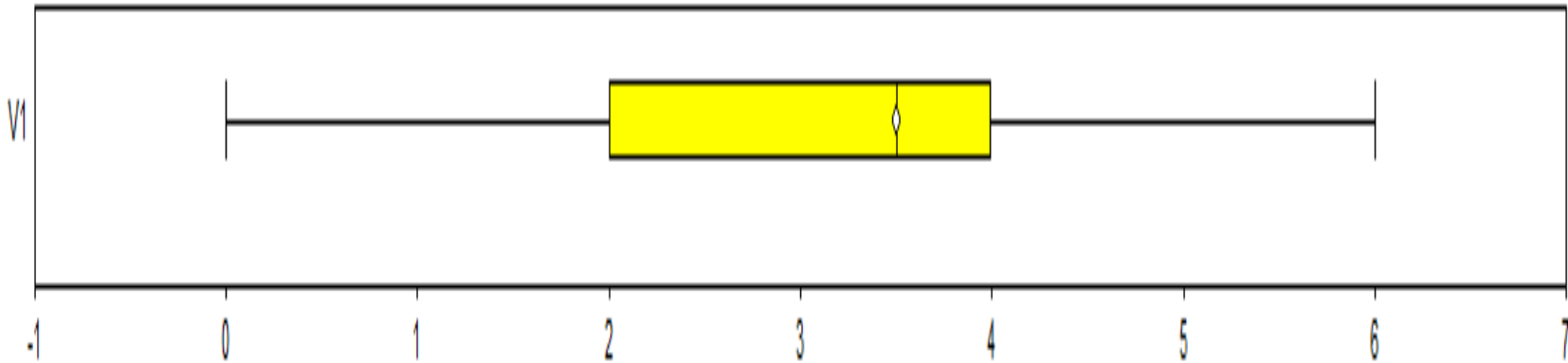
Med (Mediaan) = 3,5; Q_3 (Derde kwartiel) = 4

MaxX (Maximale waarde) = 6



5.3 Spreidingsmaten [3]

Voorbeeld:



Spreidingsbreedte = verschil tussen grootste en kleinste waarnemingsgetal
($6 - 0 = 6$)

Kwartielfstand = verschil tussen derde en eerste kwartiel ($4 - 2 = 2$)

De spreidingsbreedte en de kwartielfstand geven informatie over de spreiding van de getallen. Hoe groter deze **spreidingsmaten** zijn, hoe meer de getallen verspreid zijn.

Meestal wordt echter de **standaardafwijking** als spreidingsmaat gebruikt.

5.3 Spreidingsmaten [4]

Voorbeeld:

Afstand	0	1	2	3	4	5	6
frequentie	2	3	4	4	7	5	1

Bereken de standaardafwijking:

Stap 1:

Voer de lijst $L1 = \{0, 1, 2, 3, 4, 5, 6\}$ en lijst $L2 = \{2, 3, 4, 4, 7, 5, 1\}$ in.

Stap 2:

Gebruik 1-Var Stats L1, L2

Stap 3:

De standaardafwijking is nu te vinden bij: σx en is 1,63.

5.4 Onderzoeken en presenteren [1]

Bij een statistisch onderzoek probeert men antwoord te vinden op een vraag te vinden door het verzamelen, verwerken en interpreteren van gegevens.

De Stemming van 19 juli 2015

De afgelopen week laat een aantal beperkte verschuivingen zien. De VVD zakt 2 weg en de SP 1. PVV wint er 1, GroenLinks 1 en 50PLUS 1. Daardoor staan nu 4 partijen (VVD, SP, CDA en PVV) nog maar 3 zetels van elkaar af. Via een extra analyse gaan we dieper op de relatieve positie van de verschillende partijen in.

<i>in zetels Tweede Kamer</i>	TK2012	5-7 2015	12-7 2015	Vershil in 1 week	Versch. TK2012
VVD	41	26	24	-2	-17
PvdA	38	9	9		-29
PVV	15	20	21	1	6
SP	15	23	22	-1	7
CDA	13	22	22		9
D66	12	17	17		5
ChristenUnie	5	6	6		1
GroenLinks	4	13	14	1	10
SGP	3	3	3		
Partij voor de Dieren	2	5	5		3
50PLUS	2	4	5	1	3
VNL		2	2		2
Totaal	150	150	150		

Regelmatig wordt in Nederland onderzocht wat de politieke voorkeur van de Nederlanders is.

5.4 Onderzoeken en presenteren [1]

Bij een goed **statistisch onderzoek** horen de volgende fasen:

- 1) Het onderzoeksontwerp;
- 2) Het verzamelen van gegevens;
- 3) Het verwerken en analyseren van de verzamelde gegevens;
- 4) Het trekken van conclusies.

Het **onderzoeksontwerp** omvat een:

- **Onderzoeksonderwerp;**
- **Onderzoeksvraag;**
- **Probleemstelling;**
- **Hoofdvraag.** (**Deelvragen** helpen om antwoord op hoofdvraag te geven)

Ook dient er aandacht te zijn voor de relevantie van het onderzoek en de haalbaarheid. Bepaal welke variabelen een rol spelen en hoe de gegevens verzameld gaan worden.

De **populatie** is de totale groep waarop het onderzoek zich richt.

5.4 Onderzoeken en presenteren [2]

Als bepaalde zaken onderzocht worden zoals:

- Het stemgedrag van de Nederlandse bevolking;
- Het beste soort wasmiddel;
- De kwaliteitscontrole van een product.

dan wordt meestal gebruik gemaakt van een **steekproef**. Bij een steekproef wordt een selectie gemaakt uit de volledige groep (bijvoorbeeld 10.000 personen (**respondenten**) uit de Nederlandse bevolking wordt naar hun politieke voorkeur gevraagd). Op basis van deze selectie kunnen dan bepaalde conclusies getrokken worden.

Een steekproef moet **representatief** zijn, om een juiste afspiegeling van de volledige groep (**populatie**) te zijn.

1) De steekproef moet voldoende groot zijn;

(Er moeten voldoende mensen ondervraagd worden)

2) De steekproef moet aselekt zijn.

(Iedereen moet evenveel kans hebben om in de steekproef voor te komen)

5.4 Onderzoeken en presenteren [3]

Variabelen die onderzocht worden, kunnen op verschillende manieren gemeten worden:

1. **Nominaal meetniveau:**

Alleen het onderscheid tussen verschillende categorieën is van belang.
Bv. Geslacht, woonplaats, beroep en postcode.

2. **Ordinaal meetniveau:**

Niet alleen het onderscheid maar ook de volgorde tussen verschillende categorieën van belang.
Bv. aantal sterren van restaurant, rang in het leger.

3. **Intervalniveau:**

Onderscheid en volgorde tussen verschillende categorieën zijn van belang en er zijn gelijke verschillen tussen de categorieën.
Bv. Jaartallen.

5.4 Onderzoeken en presenteren [3]

Variabelen die onderzocht worden, kunnen op verschillende manieren gemeten worden:

4. **Rationiveau:**

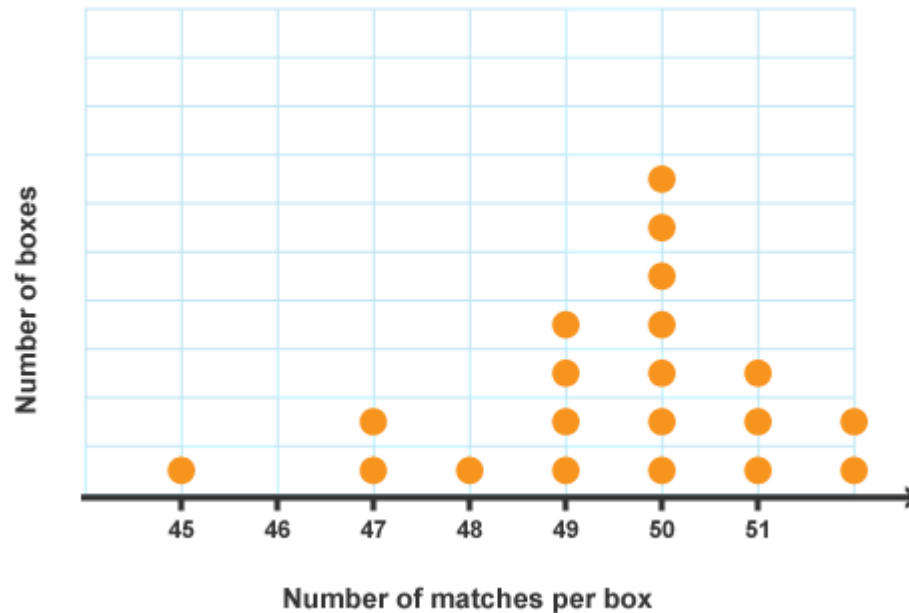
Onderscheid en volgorde tussen verschillende categorieën zijn van belang.
Er zijn gelijke verschillen tussen de categorieën en er is een natuurlijk nulpunt.

Bij leeftijd is de geboorte een natuurlijk nulpunt.

Bij inkomen is geen inkomen een natuurlijk nulpunt.

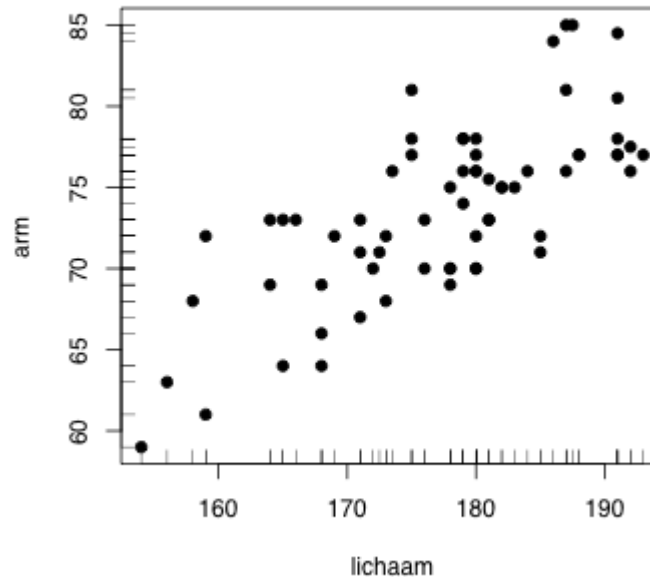
Bij jaartallen is het jaar nul geen gegeven nulpunt, want het is op een gegeven moment gekozen. (Jaartallen zijn een meting op intervalniveau)

5.4 Onderzoeken en presenteren [4]



In het plaatje hierboven is m.b.v. een **dotplot** aangegeven hoe vaak een bepaald aantal lucifers in een doosje zit.

5.4 Onderzoeken en presenteren [4]



Dit plaatje is een **spreidingsdiagram**. Van een groep personen is onderzocht hoe lang ze zijn en wat de lengte van hun armen is.

De punten zijn de elementen (personen) van de onderzochte populatie. Elk element is op twee kenmerken (lichaamslengte en lengte armen) onderzocht.

5.4 Onderzoeken en presenteren [4]

Voorbeeld:

2,6% van alle Nederlanders, die mogen stemmen, is lid van een politieke partij. Dit is de **populatieproportie** (p). Dit is het gedeelte van de populatie dat een bepaald kenmerk heeft.

$$p = \frac{\text{aantal elementen met een bepaald kenmerk in de populatie}}{\text{totaal aantal elementen in de populatie}}$$

Bij een steekproef van lengte 1.400 blijkt dat 35 mensen lid zijn van een politieke partij.

De **steekproefproportie** ($\hat{p}$) is gelijk aan: $\frac{35}{1400} = 0,025$

$$\hat{p} = \frac{\text{aantal elementen met een bepaald kenmerk in de steekproef}}{\text{totaal aantal elementen in de steekproef}}$$

5.4 Onderzoeken en presenteren [5]

Veel voorkomende fouten bij statistisch onderzoek:

- 1) Een conclusie is gebaseerd op een niet representatieve steekproef;
- 2) Een conclusie is suggestief, nietszeggend of zet je op het verkeerde been;
- 3) Voor een conclusie is slechts een gedeelte van de onderzoeksresultaten gebruikt;
- 4) Er wordt een gevolgtrekking gedaan die te kort door de bocht is.

Er is een causaal (oorzakelijk) verband tussen twee variabelen als:

- 1) Er een statistisch verband is tussen de variabelen;
- 2) De variabele die als veroorzakende variabele wordt beschouwd (onafhankelijke variabele), moet in de tijd voorafgaan aan de andere variabele (afhankelijke variabele);
- 3) Er moeten geen andere variabelen in het spel zijn die het verband veroorzaken.