

# Kwantitatief onderzoek

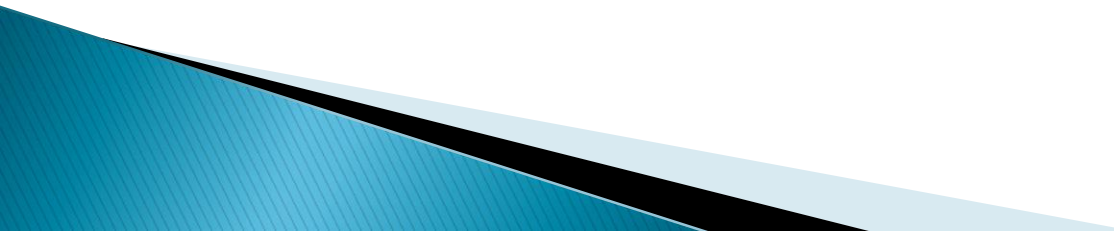
College 11  
Correlatie en Regressie

# Vorige les

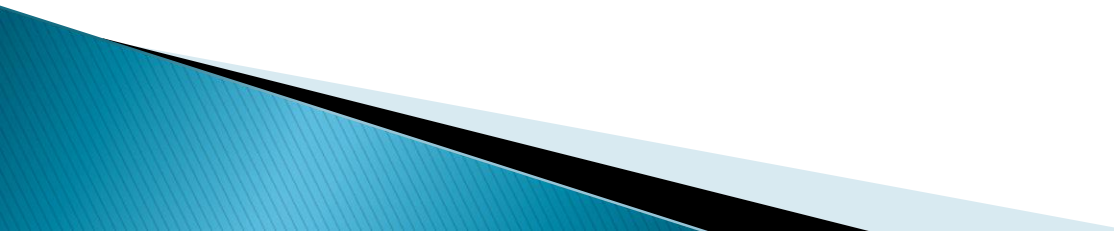
Herhalen T-test in experiment opdracht

- ▶ **Leerdoel 1:** Studenten kunnen omschrijven wat een experiment is.
- ▶ **Leerdoel 2:** Studenten kunnen een experiment uitvoeren, analyseren en interpreteren aan de hand van een t-test in SPSS.

ANOVA

- ▶ **Leerdoel 3:** studenten kunnen uitleggen wat een ANOVA is en wanneer je deze gebruikt.
  - ▶ **Leerdoel 4:** Studenten kunnen een One-Way ANOVA uitvoeren en interpreteren.
  - ▶ **Leerdoel 5:** Studenten kunnen een post hoc analyse uitvoeren en interpreteren.
- 

# Deze les

- ▶ **Leerdoel 1:** studenten kunnen het verschil tussen een verband (correlatie) en een causale relatie uitleggen.
  - ▶ **Leerdoel 2:** studenten kunnen een correlatie analyse uitvoeren en interpreteren.
  - ▶ **Leerdoel 3:** Studenten kunnen een enkelvoudige en meervoudige regressie analyse uitvoeren en interpreteren.
- 

# Zetten games aan tot geweld?

[https://www.youtube.com/watch?v=\\_tLFd9L40oU](https://www.youtube.com/watch?v=_tLFd9L40oU)



# Opdracht 1: Wat denk jij?

- ▶ Is er een relatie tussen gaming en geweld?
- ▶ Klassikaal

http://www.game-en-co.nl/gamen-en-geweld-correlatie/ Gamen en geweld - correlat...

Bestand Bewerken Beeld Favorieten Extra Help

ALGEMEEN (P)REVIEWS CONSOLE&PC BROWSER EN APP TV WINNEN! CONTACT NIEUWS Zoek op deze website... Zoek

## Gamen en geweld - correlatie of niet?

26 februari 2013 door Willem Groeneveld

Gamen en geweld gaat hand in hand, rapmuziek zet aan tot verkrachtingen en iedere junk begint met een jointje. Maar zit er wel een correlatie tussen?

Zo maar wat algemene waarheden die door de kracht van herhaling een eigen leven leiden, maar waar amper wetenschap aan ten grondslag ligt. De weinige grondige onderzoeken die over dit onderwerp gaan zijn allesbehalve unaniem. Het ene **wetenschappelijk rapport** ziet een voorzichtige link, terwijl het andere **onderzoek** dat juist tegensprekt. Gamen en geweld is een onderwerp dat aandacht nodig heeft, maar vooral behoefte heeft aan een nuance.

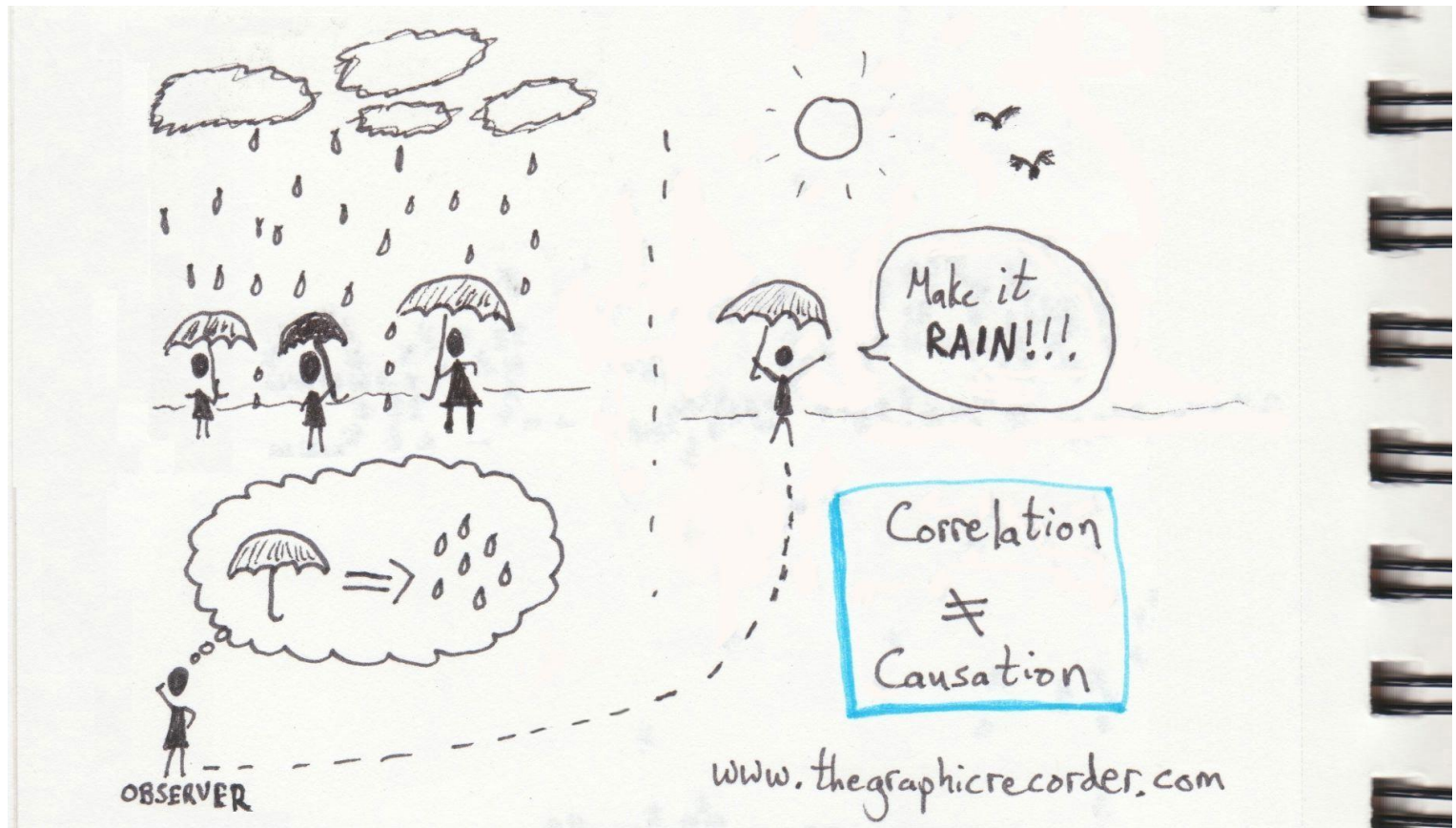
Powered by: MediaMarkt OOG ONDERDEEL VAN DE NEDERLANDSE STUC

### TOP BERICHTEN & PAGINA'S

- Minecraft: de basis uitgelegd door Jens en Jonas | video
- Theme Park gratis en nu in het Nederlands
- World of Tanks review: knallen voor nop
- Officiële kaart GTA V bekend en circuleert op internet
- "Nu racen, kut!" New Kids het spel

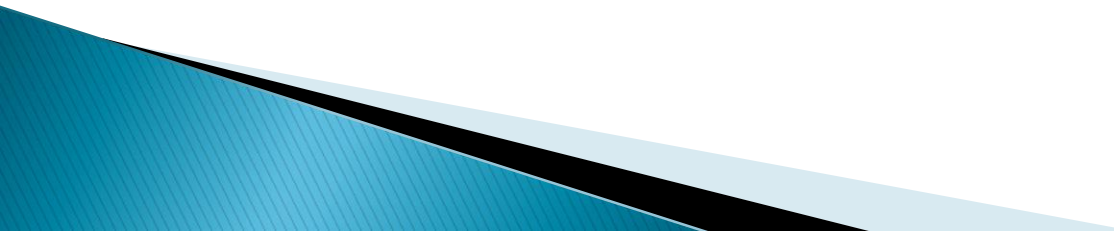


# Correlatie vs. causaal verband



# Opdracht 2

**Leerdoel 1:** studenten kunnen het verschil tussen een verband (correlatie) en een causale relatie uitleggen.

- ▶ Bedenk zelf een voorbeeld waarbij er wel sprake is van een correlatie, maar niet van een causaal verband.
  - ▶ Hoe: in groepjes
  - ▶ Tijd: 5 minuten
  - ▶ Korte presentatie per groepje (max 2 min)
- 

# Correlatie vs. causaal verband



**AMSTERDAM - Britse neurologen gebruiken vreemd verband om probleem in de wetenschap aan te kaarten.**

Zij schrijven dat maandag in Practical Neurology.

De neurologen reageren met hun bevinding op een recente publicatie die een verband legde tussen de consumptie van chocola en het winnen van Nobelprijzen. De flavonoïden uit de chocola zouden zorgen voor betere hersenen.

Ze zeggen dat ook al is er een verband, of correlatie, tussen twee zaken, dan wil dat nog niet zeggen dat er ook een oorzakelijk of causaal verband is. De Britten gebruiken het als grap bedoelde artikel om een veelvoorkomend probleem in de wetenschap aan de kaak te stellen, namelijk dat niet-bewezen verbanden wel vaak als oorzakelijk worden opgevat.

## Reactie

De keus voor het verband tussen melk en de belangrijke wetenschapsprijs komt als een reactie op het chocoladeonderzoek. Melk vertoont namelijk een sterker verband met de Nobelprijs dan chocola. Via deze luchtige publicatie met serieuze ondertoon proberen ze het verschil tussen correlatie en causatie op een humoristische wijze duidelijk te maken.

Het verband is wel echt, maar het punt van de neurologen is dat er veel factoren te vinden zijn die dit verband kunnen verklaren. Melk wordt bijvoorbeeld vaak gedronken in Westerse landen, en daar is het onderwijssysteem over het algemeen goed. Er is dan dus wel een gewoon, maar geen oorzakelijk verband tussen melk drinken en het winnen van de prijs.



# Correlatie vs. causaal verband

http://www.nu.nl/lifestyle/2857381/mannen-huishou... x Mannen die het huishoude... x

ewerken Beeld Favorieten Extra Help

Pagina Beveiliging Extra ? N S

**nu.nl** Zondag 22 maart 2015 Het laatste nieuws het eerst op NU.nl

9 °C 1 NS 499,12 TV gids 4 Live

**Net binnen** Meest gelezen

14:43 - Zonderland pakt brons op brug en valt op... >  
14:23 - Mercedes probeerde Duitse Grand Prix te... >  
14:21 - PEC Zwolle komt niet verder dan gelijksp... >  
14:14 - Farah loopt Europees record op halve ma... >

Meer nieuws >

**Voorpagina**  
Net binnen  
NUweekend  
Algemeen  
Binnenland  
Buitenland  
Politiek  
Economie  
Geld  
Ondernemen  
Beurs  
Belastingaangifte  
Sport  
Voetbal  
Scorebord  
Wielrennen  
Formule 1  
Meer sport  
Tech  
Internet  
Mobiel  
Gadgets  
Games  
Entertainment  
Achterklap  
Films en series  
Muziek  
Boek en cultuur  
Media

**NU.nl > Lifestyle**

**Mannen die het huishouden doen hebben vaker seks**

Foto: Thinkstock

Gepubliceerd: 12 juli 2012 13:53  
Laatste update: 12 juli 2012 13:53

**AMSTERDAM -** Mannen die regelmatig helpen in het huishouden hebben vaker seks dan mannen die niets doen in het huishouden.

Dit blijkt uit een recent onderzoek van een datingsite. Aan het onderzoek deden 5.760 mensen mee in tien Europese landen en in Brazilië.

72 procent van de Nederlandse mannen die huishoudelijke klusjes doen, heeft elke dag seks. Dit in tegenstelling tot de mannen die niets doen, van wie slechts 10 procent elke dag tussen de lakens duikt.

**Huishoudelijke taken**

Nederlandse vrouwen zien hun mannen voornamelijk graag stofzuigen, gevolgd door de afwas en de was. Stofzuigen zorgt ervoor dat mannen de meeste seks krijgen. 79 procent van de mannen die dit huishoudelijke klusje op zich nemen, heeft elke dag seks.

Mannen die elke dag seks te veel vinden, maar een keer in de week wel voldoende, kunnen beter de afwas doen. Van mannen die de afwas doen, heeft 84 procent een keer per week seks. Van de mannen die niets doen, heeft 8

**NGO MEE VOOR LJOENEN**  
valt er gegarandeerd Bingo!

**Dagelijkse nieuwsbrief**

Elke dag rond 12.00 uur het belangrijkste nieuws in uw mailbox?  
**Schrijf u in voor de NU.nl-nieuwsbrief >**

**Gaultier kleedt Manneken Pis**

Jean-Paul Gaultier heeft vanwege de Brusselse Gay Pride een feestelijk nakie ontwerp voor

75%

14:58  
22-3-2015

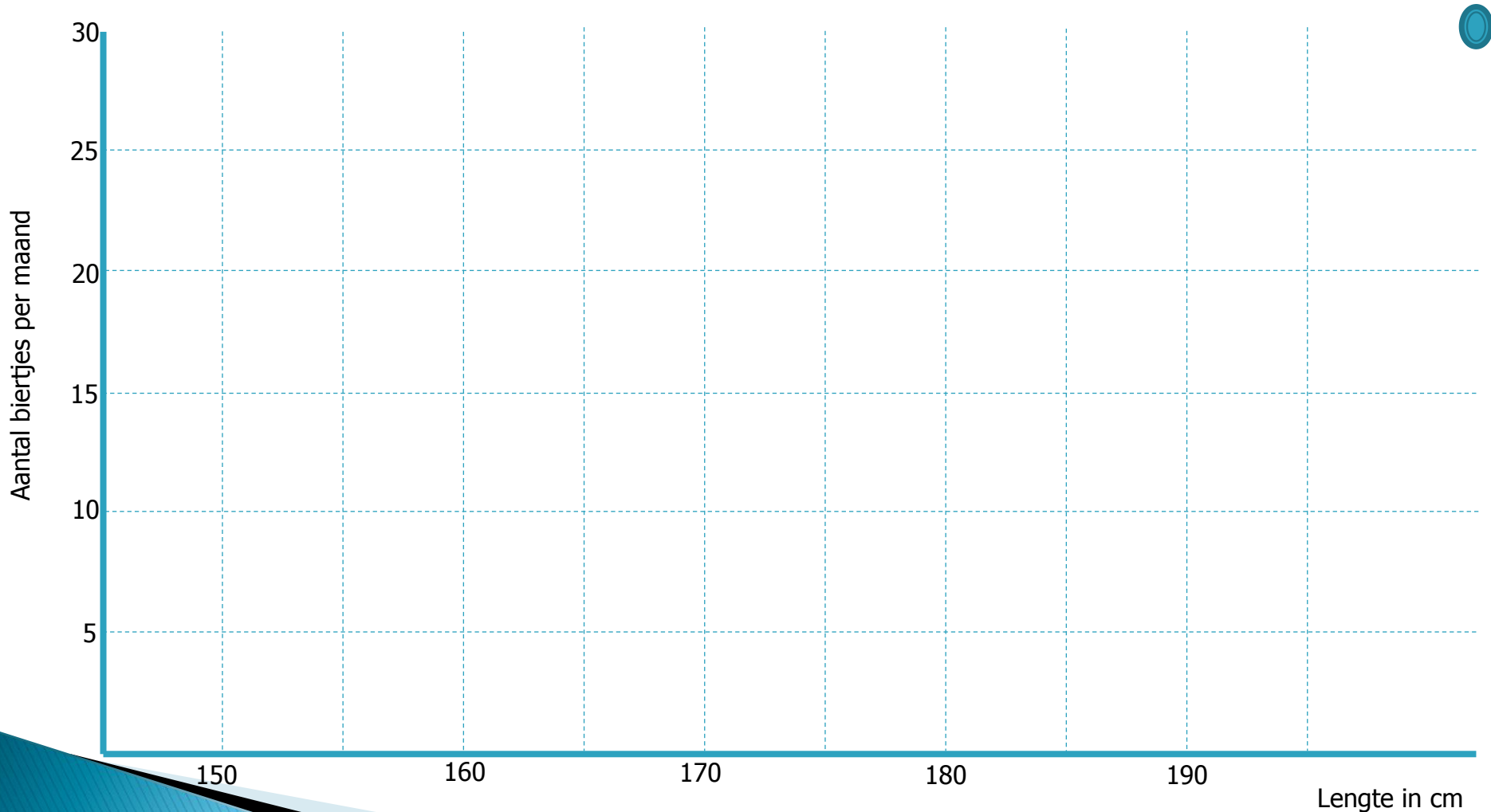
# Wanneer welke toets?

Onafhankelijke variabele	Afhankelijke variabele	Soort techniek
Nominaal / Ordinaal	Nominaal / Ordinaal	Chi kwadraat toets
Nominaal / Ordinaal	Interval / Ratio	T-test of ANOVA
Interval / Ratio	Interval / Ratio	Regressie of correlatie

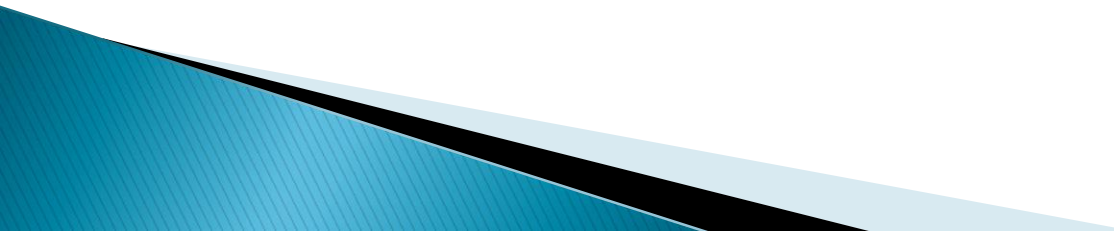
# CORRELATIE



## Opdracht 3; verband lengte en bierdrinkgedrag



# Correlatie Kenmerken

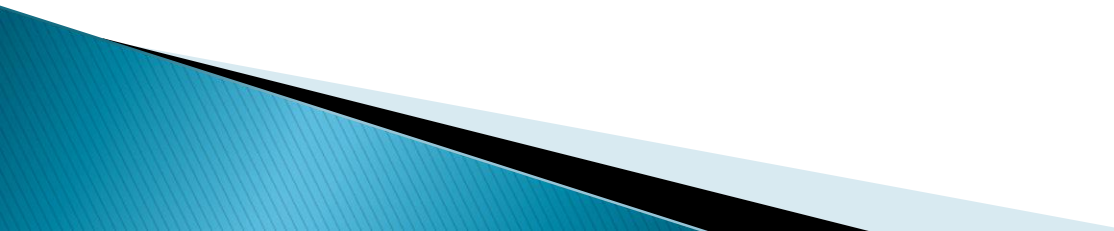
- ▶ Parametrische test
  - ▶ Toetst of er een (lineaire) samenhang is tussen twee variabelen.
  - ▶ Geschikt voor variabelen gemeten op Interval of Ratio niveau.
  - ▶ Geeft niet een **oorzakelijk** verband weer maar een **statistisch** verband!
- 



# Opdracht 4

- ▶ Stel dat je onderzoek doet in de muziekbranche. Je hebt in kaart hoe veel minuten mensen gemiddeld per week naar muziek luisteren. Je weet ook hoe vaak mensen per jaar een festival bezoeken. Je bent benieuwd of er een verband is tussen muziek luisterfrequentie en festival bezoekfrequentie.
- ▶ Hoe luiden  $H_0$  en  $H_1$ ?

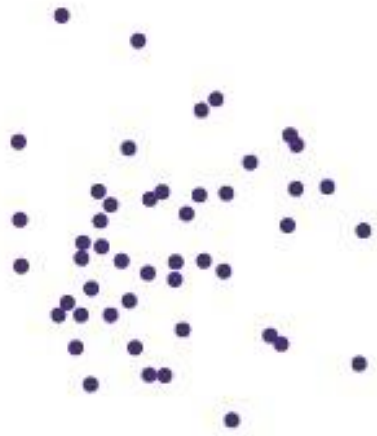
# Opdracht 4; uitwerking

- ▶ Stel dat je onderzoek doet in de muziekbranche. Je hebt in kaart hoe veel minuten mensen gemiddeld per week naar muziek luisteren. Je weet ook hoe vaak mensen per jaar een festival bezoeken. Je bent benieuwd of er een verband is tussen muziek luisterfrequentie en festival bezoekfrequentie.
  - ▶  $H_0$ : Er is geen verband tussen muziek luisterfrequentie en festival bezoekfrequentie.
  - ▶  $H_0$ : Er is wel een verband tussen muziek luisterfrequentie en festival bezoekfrequentie.
- 

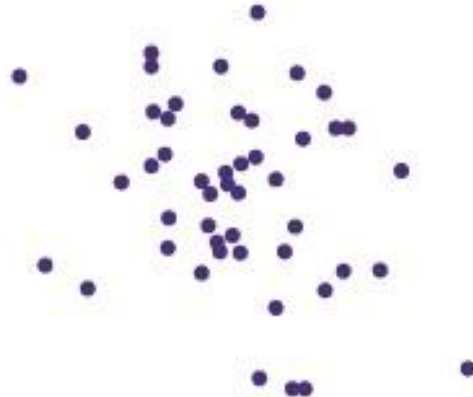
# Correlatiecoëfficiënt $r$

- ▶ Correlatiecoëfficiënt ' $r$ ' geeft mate van samenhang weer.
- ▶  $r$  kan variëren van  $-1$  tot  $+1$ .  $r$ :
  - 0: Geen correlatie tussen de variabelen
  - $-1$ : een perfecte negatieve correlatie tussen de variabelen
  - $+1$ : een perfecte positieve correlatie tussen de variabelen
- 0 tot 0,2: nauwelijks verband
- 0,2 tot 0,4: zwak verband
- 0,4 tot 0,6: redelijk verband
- 0,6 tot 0,8: sterk verband
- boven 0,8: zeer sterk verband

# Correlatiecoëfficiënt: $r$



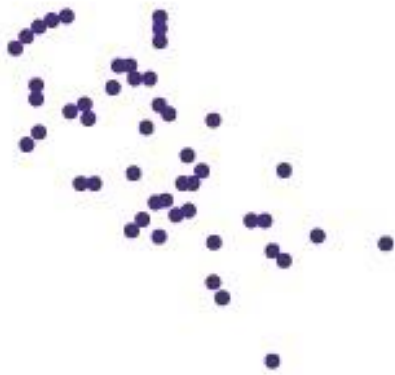
Correlation  $r = 0$



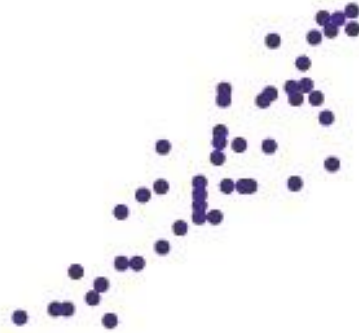
Correlation  $r = -0.3$



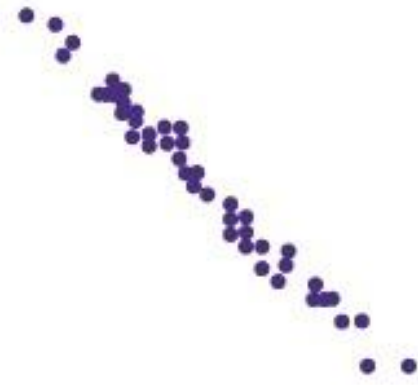
Correlation  $r = 0.5$



Correlation  $r = -0.7$



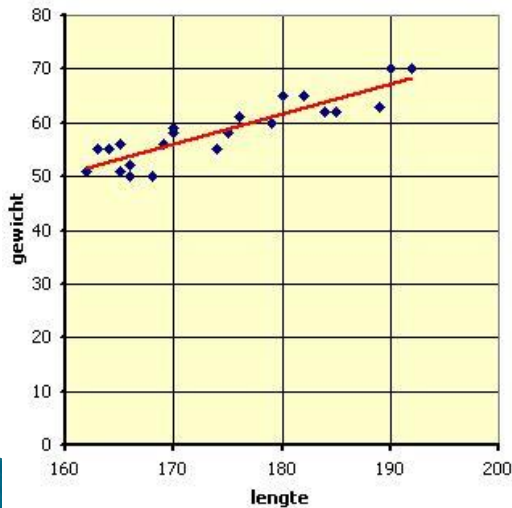
Correlation  $r = 0.9$



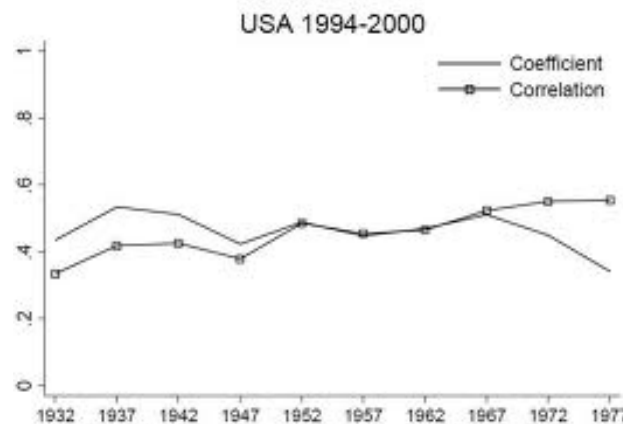
Correlation  $r = -0.99$

# Opdracht 5

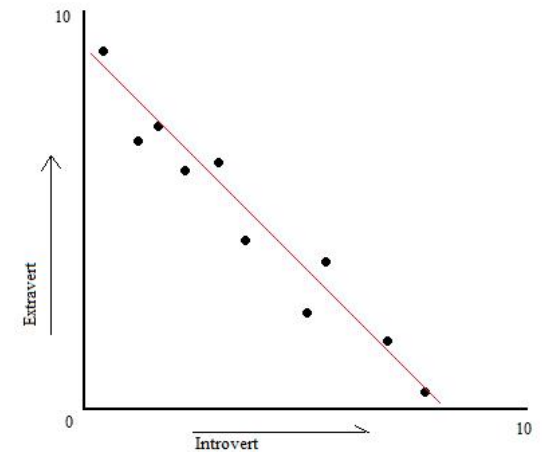
- Is bij onderstaande grafieken sprake van een negatieve correlatie, positieve correlatie, of geen correlatie? Waarom?
- Hoe: Individueel.
- Tijd: 3 minuten.



A.



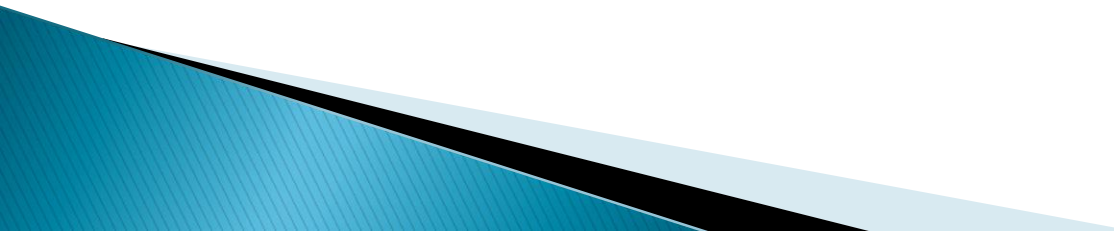
B.



C.

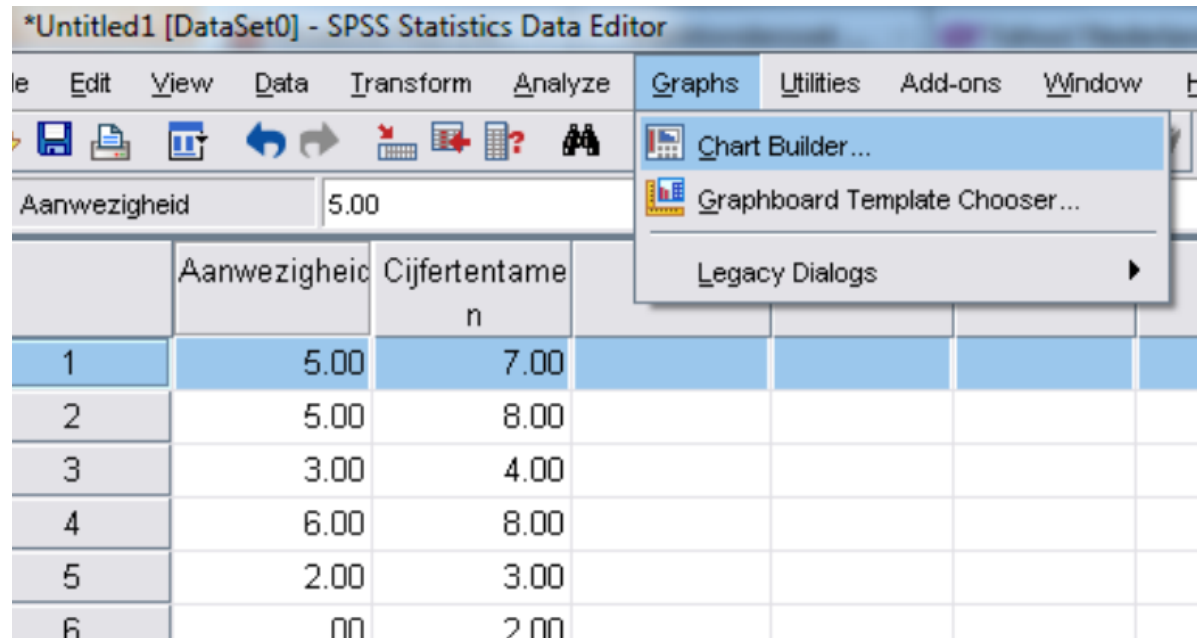


# Correlatie– stappen

1. Formuleer de  $H_0$  en de  $H_1$  hypothese
  2. Maak een scatter plot grafiek.
  3. Voer de correlatie toets uit in SPSS
  4. Interpreteer de uitkomsten
  5. Formuleer de conclusie
- 

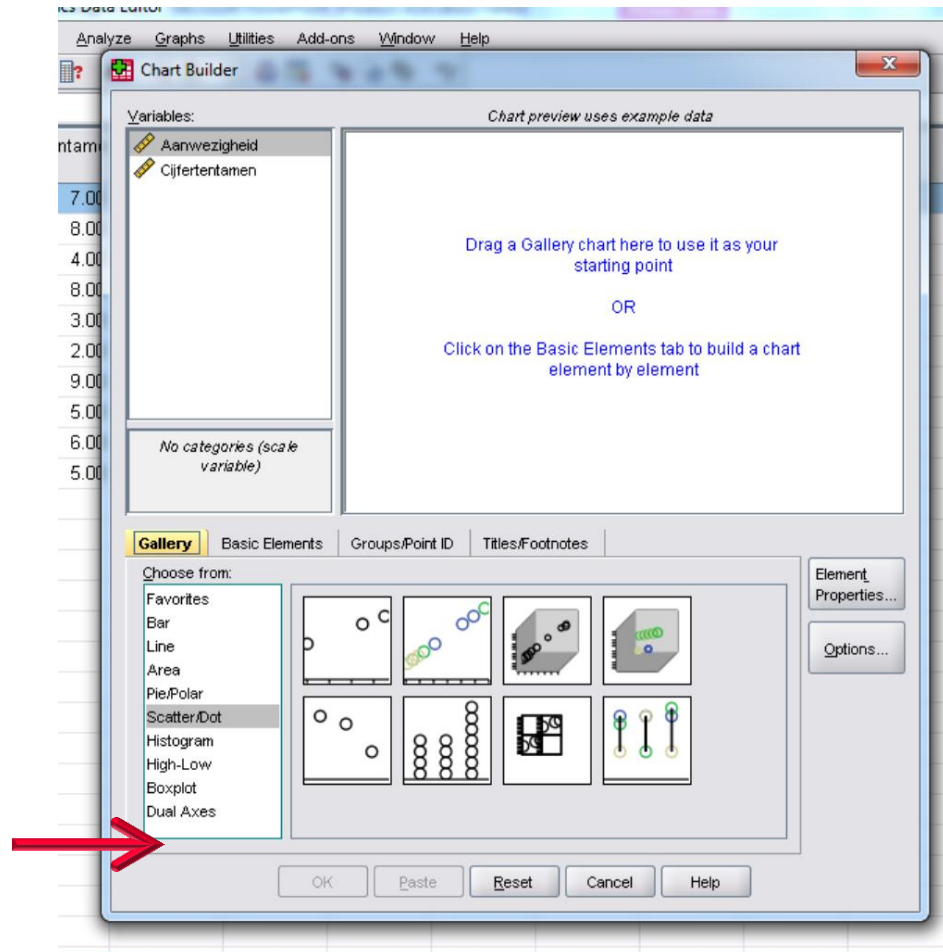
# Correlatie – How to SPSS

## 2. Kies: Graphs > Chart Builder



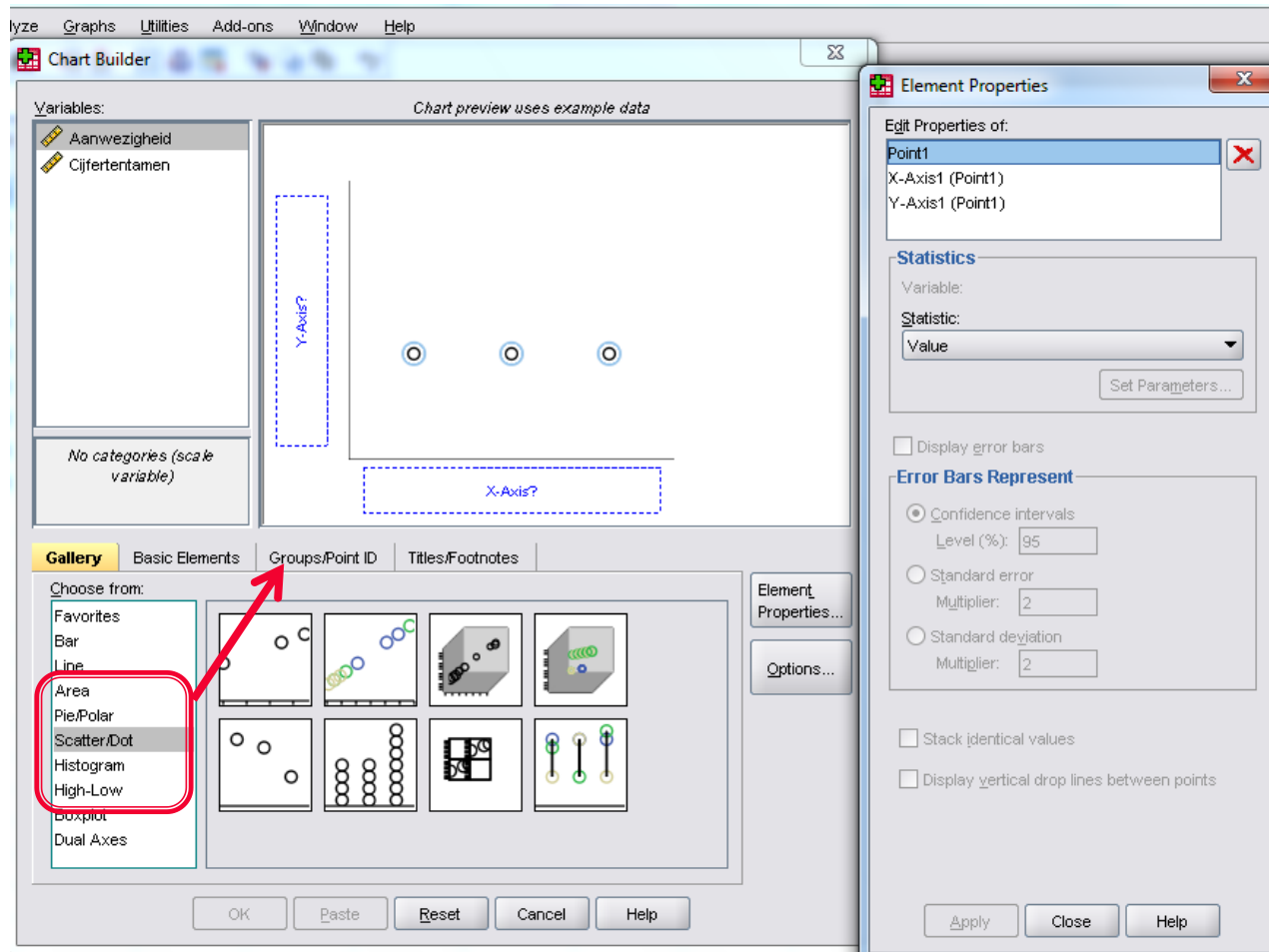
# Correlatie – How to SPSS

- Zo ziet de chart builder eruit.
- Selecteer Scatter/Dot



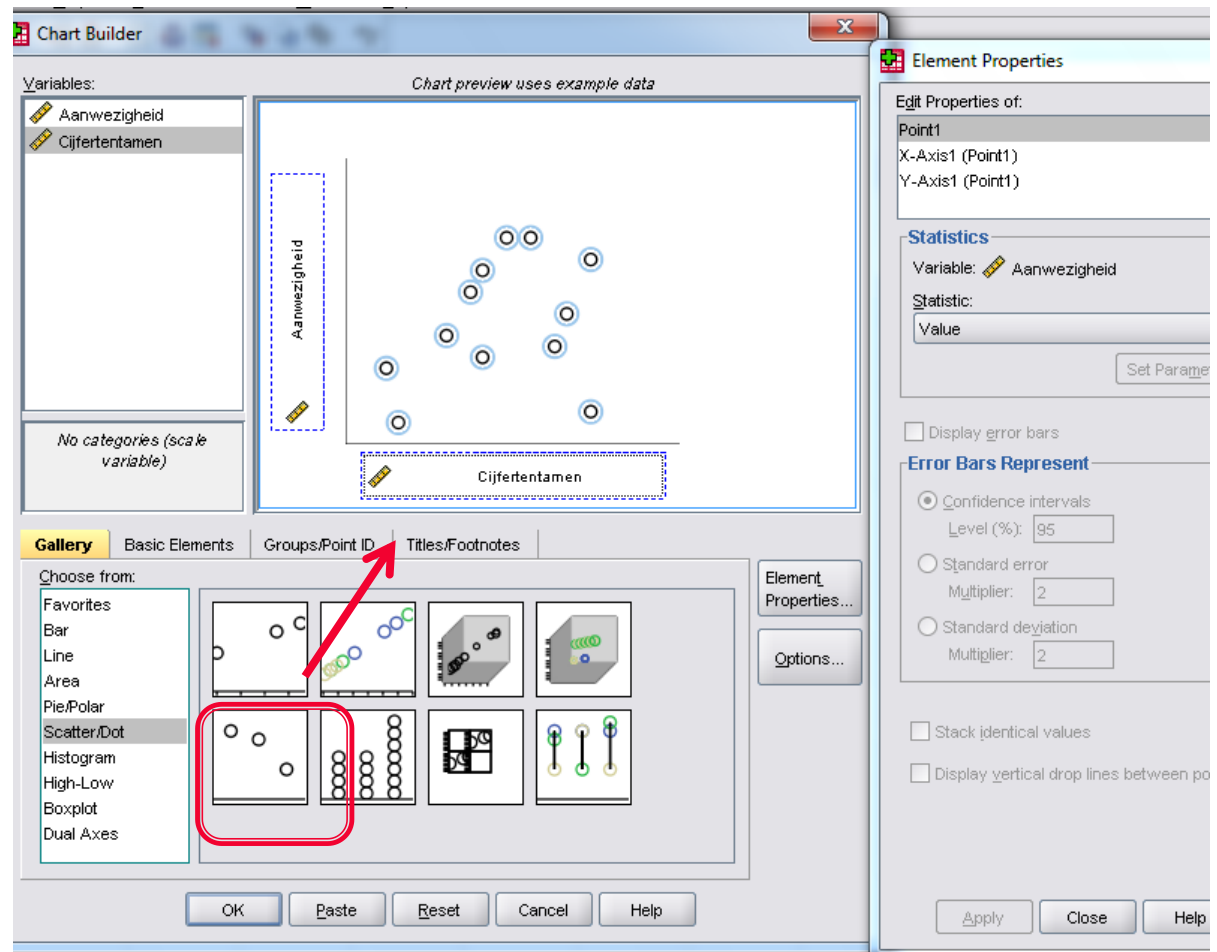
# Correlatie – How to SPSS

- “Sleep” de eerste Scatter/dot optie naar de Chart preview.
- Je ziet nu een voorbeeld Scatter Dots staan.
- Je kunt de variabelen in de grafiek slepen.



# Correlatie – How to SPSS

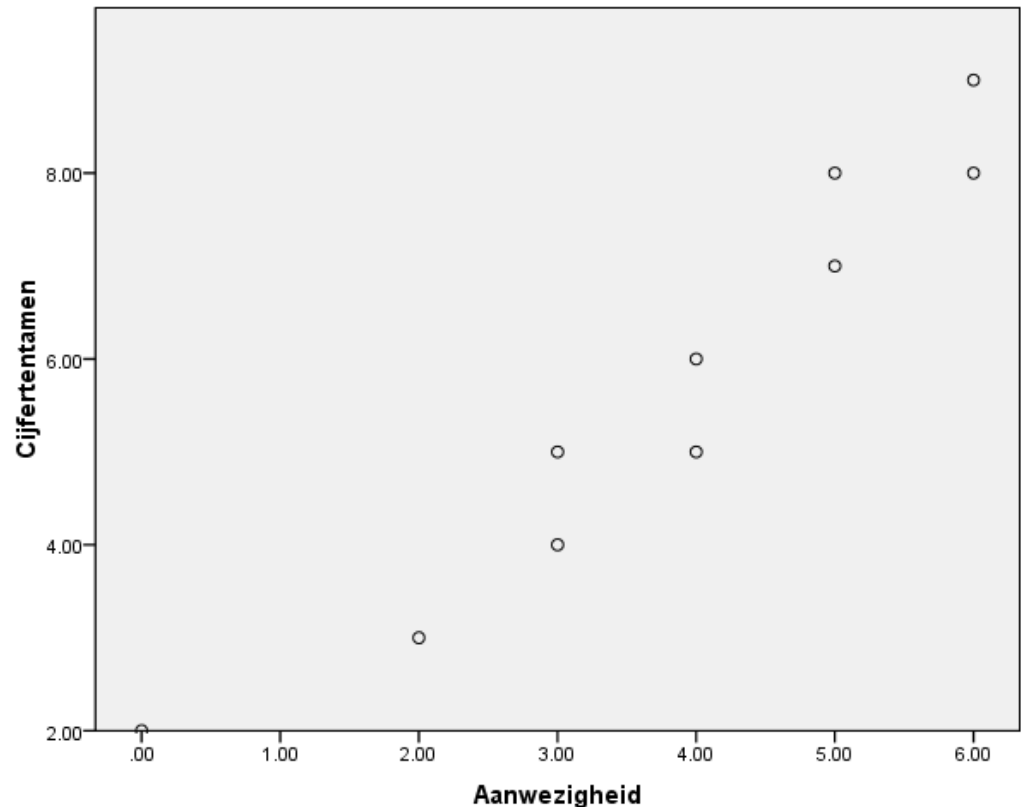
- Wanneer je de variabelen op de x en y as hebt gesleept, zie je scatter dots. Let op, dit is alleen maar een gegenereerd voorbeeld door SPSS!
- Druk nu op Ok.





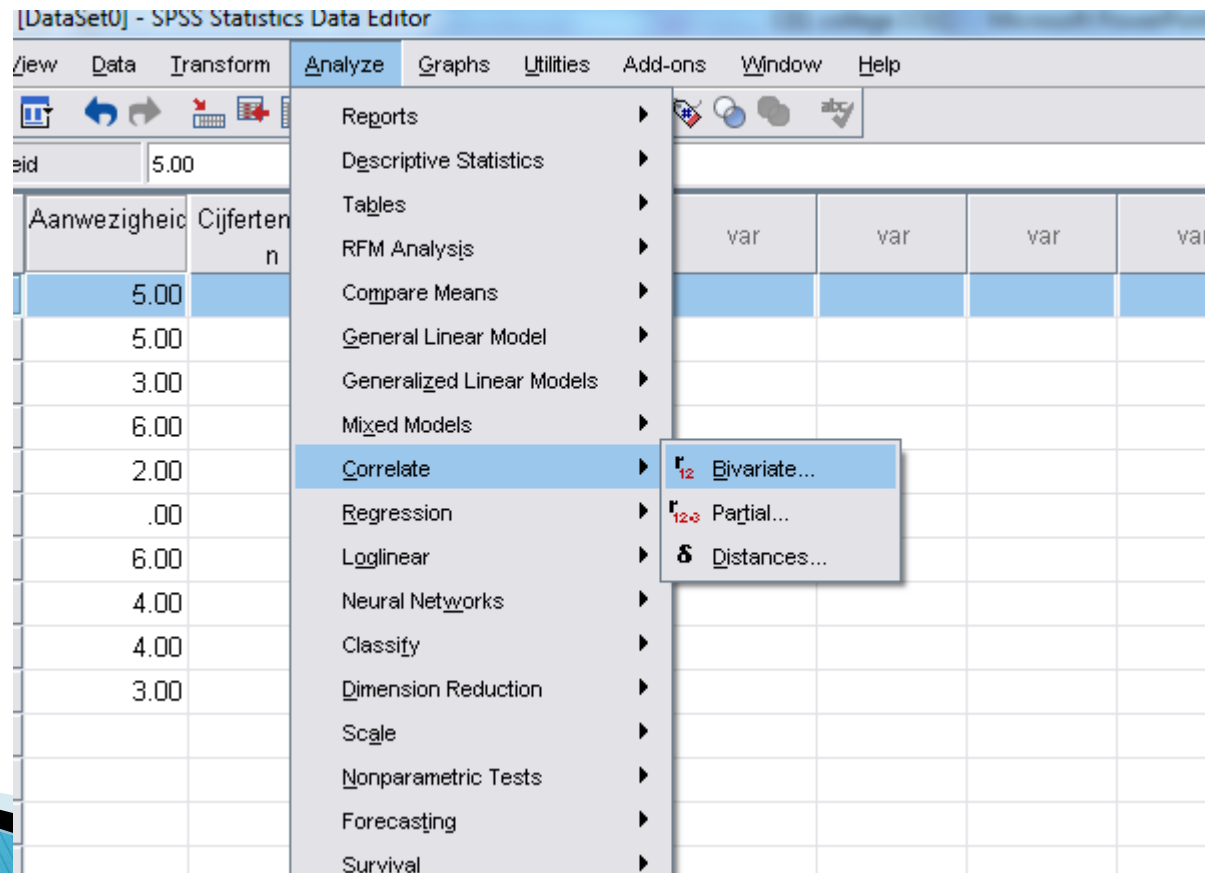
# Correlatie – How to SPSS

- Bekijk de Scatter Dot grafiek, zou hier sprake zijn van een correlatie tussen de aanwezigheid en het tentamencijfer?
- Nu ga je de correlatie toets uitvoeren.



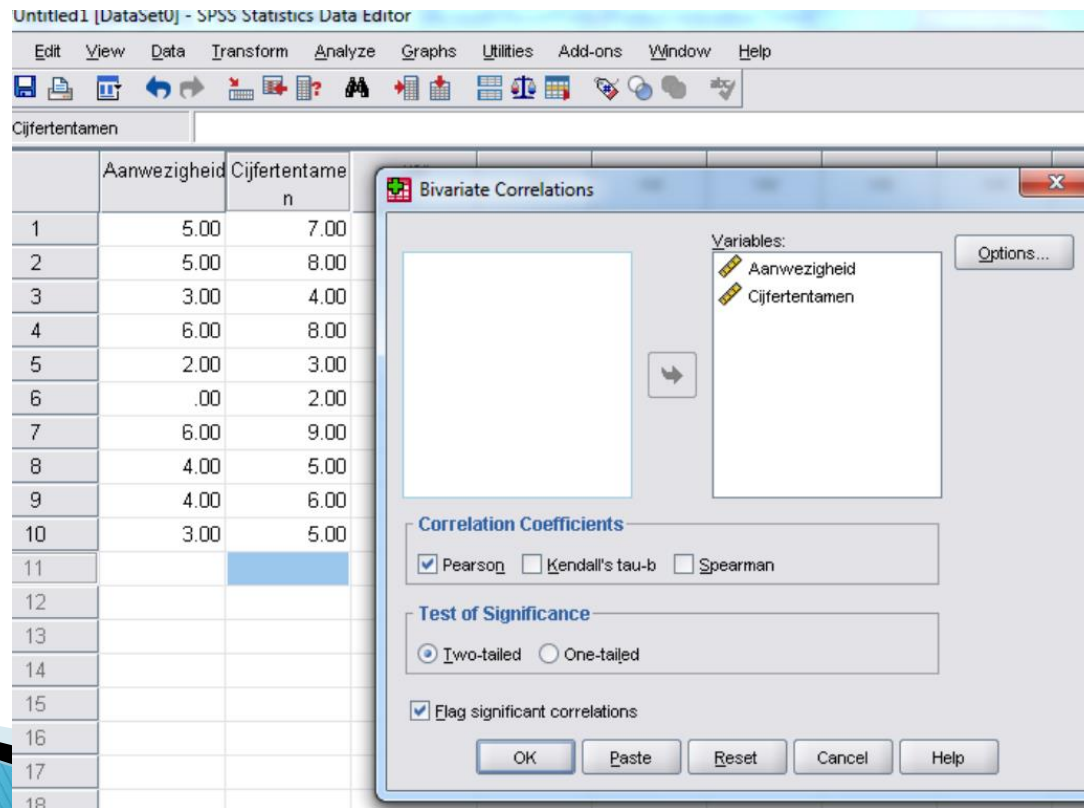
# Correlatie – How to SPSS

3. Kies: Analyze > correlate > bivariate.



# Correlatie – How to SPSS

- Selecteer de variabelen waarover je de correlatie wil laten uitvoeren en zet ze in het rechtste vakje "variables"
- Druk op OK.



# Opdracht 6

- **Opdracht:** Interpreteer de output en formuleer een conclusie.
- **Hoe:** individueel → **Correlations**
- **Tijd:** 5 minuten

[DataSet0]

**Descriptive Statistics**

	Mean	Std. Deviation	N
Aanwezigheid	3.8000	1.87380	10
Cijfertentamen	5.7000	2.31181	10

**Correlations**

		Aanwezigheid	Cijfertentamen
Aanwezigheid	Pearson Correlation	1	.959**
	Sig. (2-tailed)		.000
	N	10	10
Cijfertentamen	Pearson Correlation	.959**	1
	Sig. (2-tailed)	.000	
	N	10	10

\*\* . Correlation is significant at the 0.01 level (2-tailed).

# Opdracht 6: uitwerking

- **Opdracht:** Interpreteer de output en formuleer een conclusie.
- **Hoe:** individueel → **Correlations**
- **Tijd:** 5 minuten

[DataSet0]

Descriptive Statistics

	Mean	Std. Deviation	N
Aanwezigheid	3.8000	1.87380	10
Cijfertentamen	5.7000	2.31181	10

Correlations

		Aanwezigheid	Cijfertentamen
Aanwezigheid	Pearson Correlation	1	.959**
	Sig. (2-tailed)		.000
	N	10	10
Cijfertentamen	Pearson Correlation	.959**	1
	Sig. (2-tailed)	.000	
	N	10	10

\*\*. Correlation is significant at the 0.01 level (2-tailed).



# Opdracht 7

- Open het bestand: geld en geluk met geld en geluktotaal (opgeslagen of portal)
- Voer een correlatie analyse uit waarbij je onderzoekt of het hebben van geld samenhangt met het gelukkig zijn?

# Opdracht 7: uitwerking

**Correlations**

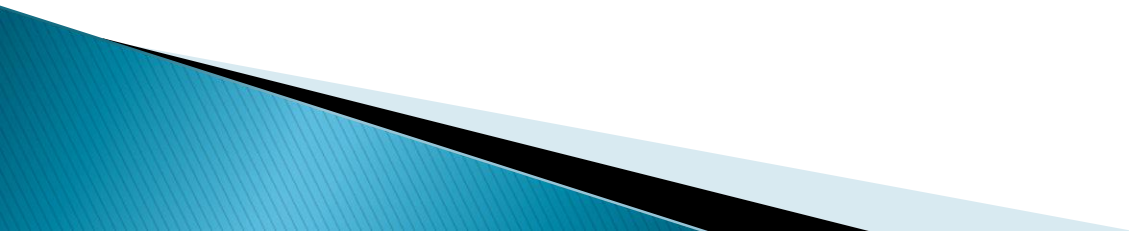
		geldtotaal	geluktotaal
geldtotaal	Pearson Correlation	1	,655**
	Sig. (2-tailed)		,000
	N	1000	990
geluktotaal	Pearson Correlation	,655**	1
	Sig. (2-tailed)	,000	
	N	990	990

\*\* . Correlation is significant at the 0.01 level (2-tailed).

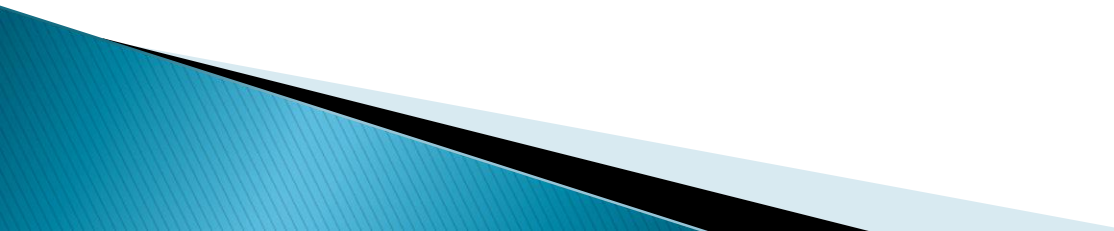
# Voor de pauze

- ▶ **Leerdoel 1:** studenten kunnen het verschil tussen een verband (correlatie) en een causale relatie uitleggen.
- ▶ **Leerdoel 2:** studenten kunnen een correlatie analyse uitvoeren en interpreteren.
- ▶ **Leerdoel 3:** Studenten kunnen een enkelvoudige en meervoudige regressie analyse uitvoeren en interpreteren.

PAUZE



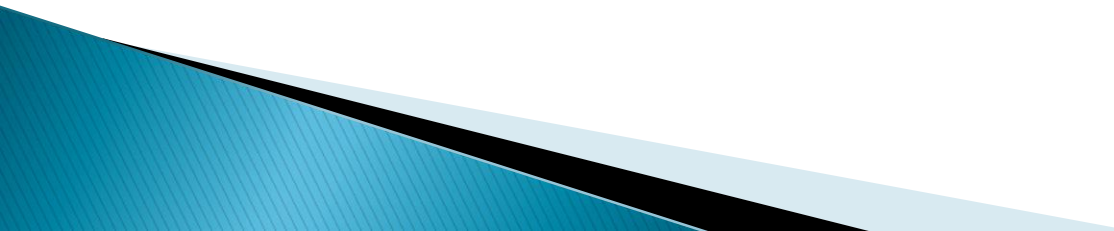
# Na de pauze

- ▶ **Leerdoel 1:** studenten kunnen het verschil tussen een verband (correlatie) en een causale relatie uitleggen.
  - ▶ **Leerdoel 2:** studenten kunnen een correlatie analyse uitvoeren en interpreteren.
  - ▶ **Leerdoel 3:** Studenten kunnen een enkelvoudige en meervoudige regressie analyse uitvoeren en interpreteren.
- 

# REGRESSIE



# Regressie analyse

- ▶ Parametrische test;
  - ▶ Voorspellend verband (invloed) tussen een (of meer) onafhankelijke variabelen en één afhankelijke variabele.
  - ▶ Gaat uit van een lineair verband tussen onafhankelijke en afhankelijke variabele.
  - ▶ Let op: je weet nooit of er nog andere variabelen een rol spelen. Een regressie analyse zegt niet alles.
- 



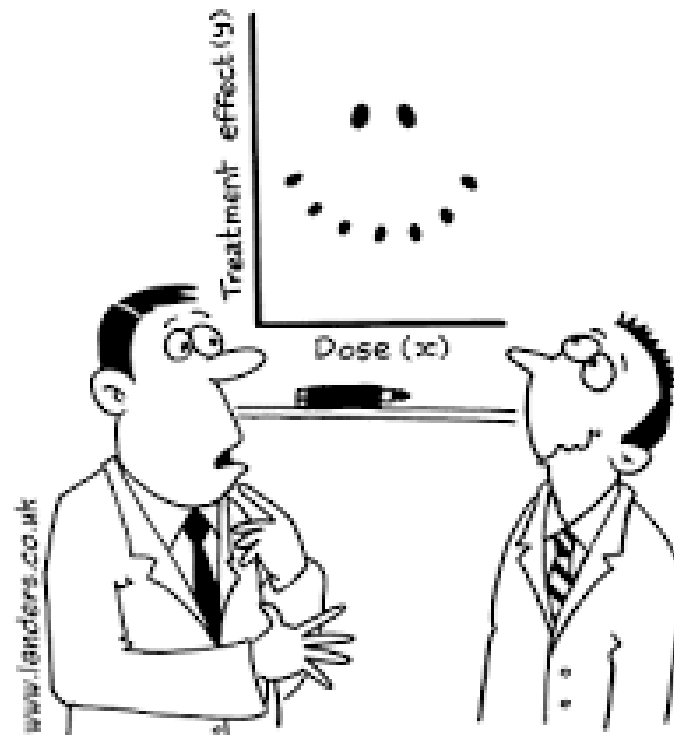
# Opdracht 8

- ▶ Stel dat je onderzoek doet in de muziekbranche. Je hebt in kaart hoe veel minuten mensen gemiddeld per week naar muziek luisteren. Je weet ook hoe vaak mensen per jaar een festival bezoeken. Je bent benieuwd of de frequentie van muziek luisteren invloed heeft op festival bezoekfrequentie.
- ▶ Hoe luiden  $H_0$  en  $H_1$ ?

# Opdracht 8: uitwerking

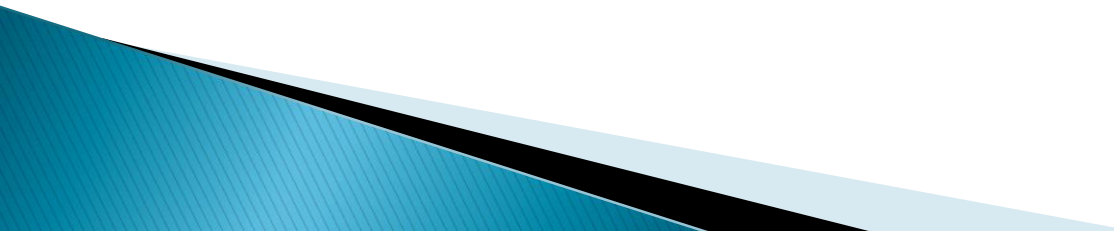
- ▶ Stel dat je onderzoek doet in de muziekbranche. Je hebt in kaart hoe veel minuten mensen gemiddeld per week naar muziek luisteren. Je weet ook hoe vaak mensen per jaar een festival bezoeken. Je bent benieuwd of de frequentie van muziek luisteren invloed heeft op festival bezoekfrequentie.
  - ▶  $H_0$ : Er is geen invloed van de frequentie van muziek luisteren op festival bezoekfrequentie.
  - ▶  $H_1$ : Er is invloed van de frequentie van muziek luisteren op festival bezoekfrequentie.
- 

# Enkelvoudige regressie



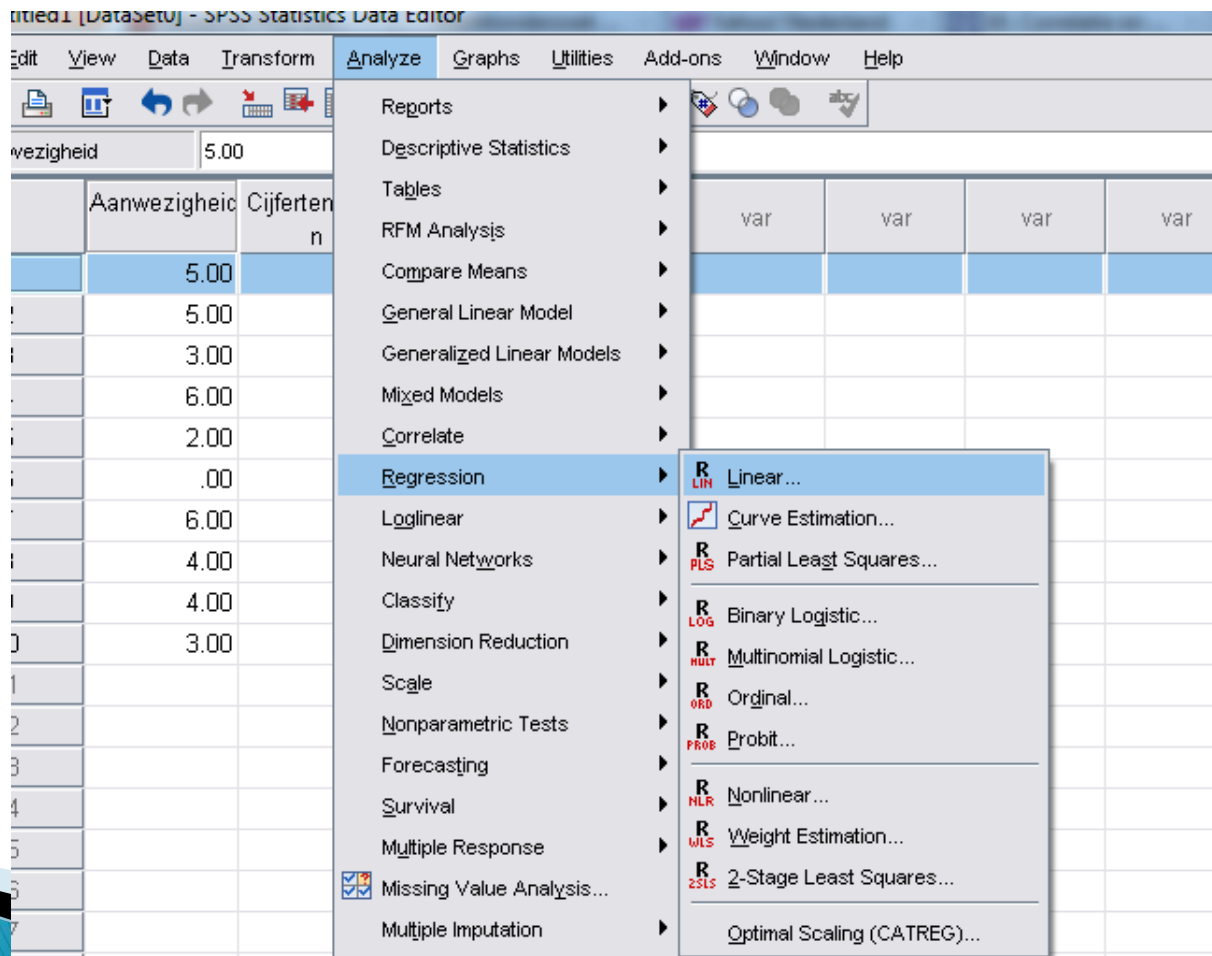
"It's a non-linear pattern with outliers.....but for some reason I'm very happy with the data."

# Enkelvoudige regressie – stappen

1. Formuleer de  $H_0$  en de  $H_1$  hypothese
  2. Voer de regressie toets uit in SPSS
  3. Interpreteer de uitkomsten
  4. Formuleer de conclusie
- 

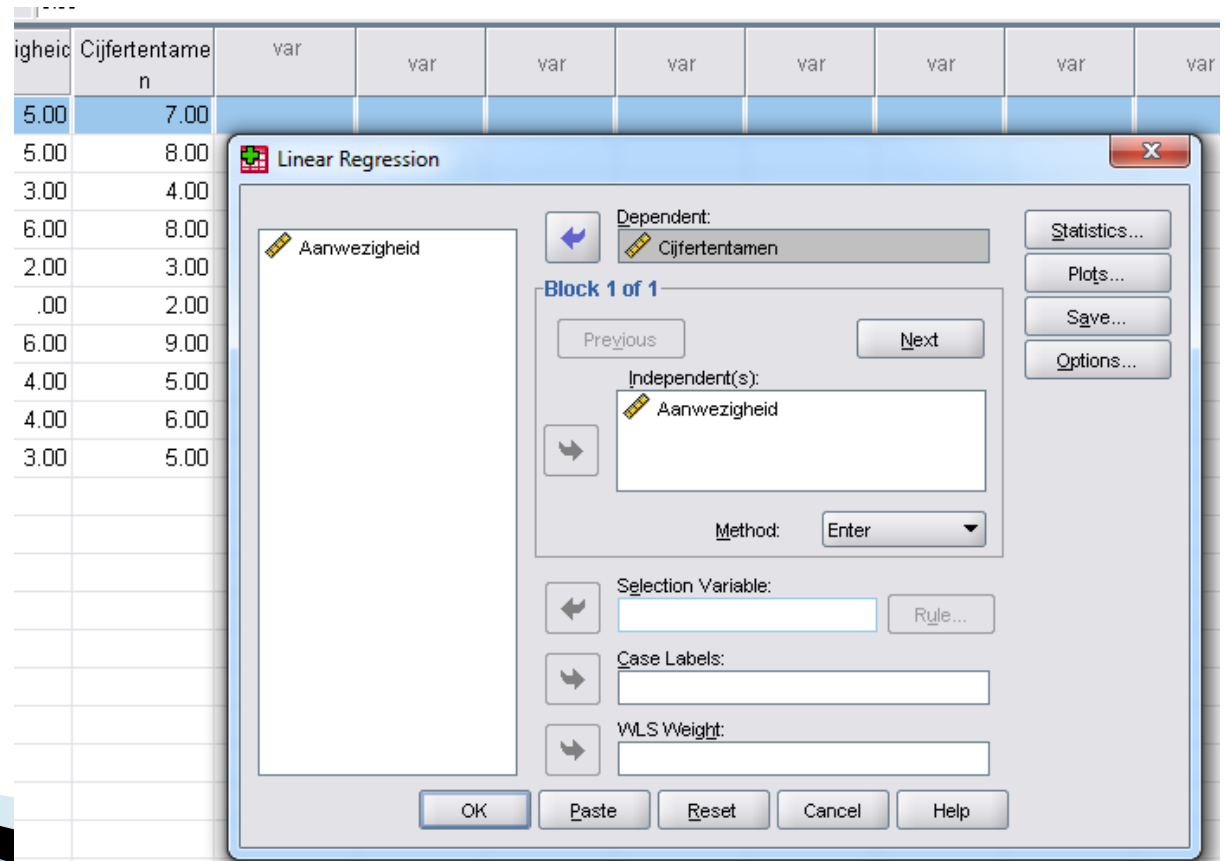
# Enkelvoudige regressie – How to SPSS

Kies: Analyze > Regression > Linear



# Enkelvoudige regressie – How to SPSS

Selecteer de afhankelijke variabele bovenin "dependent", selecteer de onafhankelijke variabele(en) eronder bij "independent". Wanneer je een meervoudige regressie wilt uitvoeren kan je dus de overige onafhankelijke variabelen hier selecteren. Druk op OK.



# Enkelvoudige regressie – How to SPSS

92% van de variatie op cijfertentamen wordt verklaard door aanwezigheid.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.959 <sup>a</sup>	.920	.910	.69241

a. Predictors: (Constant), Aanwezigheid

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.265	1	44.265	92.327	.000 <sup>a</sup>
	Residual	3.835	8	.479		
	Total	48.100	9			

a. Predictors: (Constant), Aanwezigheid

b. Dependent Variable: Cijfertentamen

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.203	.517		2.327	.048
	Aanwezigheid	1.184	.123	.959	9.609	.000

a. Dependent Variable: Cijfertentamen

Het model is significant. We weten nu voor minimaal 95% zeker dat het model 'iets' verklaart.

De invloed van aanwezigheid op cijfertentamen is significant. We weten nu voor minimaal 95% zeker dat de relatie bestaat. H<sub>0</sub> verwerpen.

De relatie is positief en geeft de grootte weer van de van deze relatie.



# Enkelvoudige regressie – How to SPSS

- ▶ We kunnen nu een voorspellingsformule maken:
- ▶  $Y = 1,20 + 1,18 * X$
- ▶ Wat voor tentamencijfer kan er voorspeld worden voor een student die 5 maal aanwezig was?

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.959 <sup>a</sup>	.920	.910	.69241

a. Predictors: (Constant), Aanwezigheid

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.265	1	44.265	92.327	.000 <sup>a</sup>
	Residual	3.835	8	.479		
	Total	48.100	9			

a. Predictors: (Constant), Aanwezigheid

b. Dependent Variable: Cijfertentamen

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.203	.517		2.327	.048
	Aanwezigheid	1.184	.123	.959	9.609	.000

a. Dependent Variable: Cijfertentamen

# Enkelvoudige regressie – How to SPSS

- ▶ We kunnen nu een voorspellingsformule maken:
- ▶  $Y = 1,20 + 1,18 * X$
- ▶ Wat voor tentamencijfer kan er voorspeld worden voor een student die 5 maal aanwezig was?

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.959 <sup>a</sup>	.920	.910	.69241

a. Predictors: (Constant), Aanwezigheid

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.265	1	44.265	92.327	.000 <sup>a</sup>
	Residual	3.835	8	.479		
	Total	48.100	9			

a. Predictors: (Constant), Aanwezigheid

b. Dependent Variable: Cijfertentamen

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.203	.517		2.327	.048
	Aanwezigheid	1.184	.123	.959	9.609	.000

a. Dependent Variable: Cijfertentamen

Antwoord:  $1,20 + 1,18 * 5 = 7,1$

# Regressie B, Beta en $R^2$

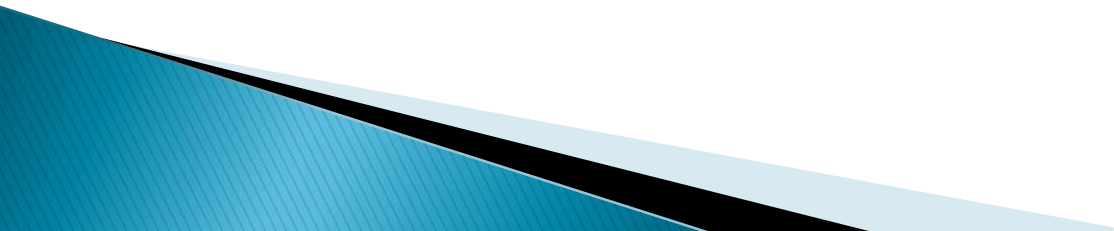
Twee coëfficiënten van belang:

- ▶ de *ongestandaardiseerde regressiecoëfficiënt* (**B**), oftewel de **absolute toe- of afname van de afhankelijke variabele** ( $y$ ) bij één eenheid toename van de onafhankelijke variabele. Bijvoorbeeld: "wanneer iemand één jaar ouder is, neemt het aantal uur per week kijken naar Net5 gemiddeld met 0,5 toe". Anders gezegd: hoe hoger de leeftijd, hoe meer naar Net5 wordt gekeken.
- ▶ de *gestandaardiseerde regressiecoëfficiënt* (**Beta**), de kracht van de toename waarmee je verschillende onafhankelijke variabelen kan vergelijken (welke draagt het meeste bij).
- ▶ de *proportie verklaarde variantie* ( $R^2$ ) kan geïnterpreteerd worden als het **percentage verklaarde variantie** in de afhankelijke variabele op grond van de onafhankelijke variabele. Bijvoorbeeld: "de variantie in het aantal uur Net5 kijken kan voor 67% verklaard worden door de variantie in leeftijd".

# Meervoudige regressie

- ▶ Een meervoudige regressie wordt hetzelfde uitgevoerd als de enkelvoudige regressie maar bij independent variabele worden meerdere variabelen geplaatst.

# Opdracht 9

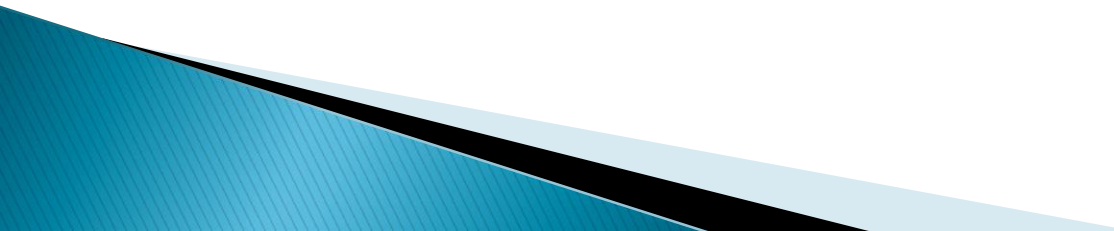
- ▶ Open het bestand van John Legend.Sav
  - ▶ Onderzoek of attitudeoptreden en attitudeinterview een invloed hebben op de bezoekingintentie.
  - ▶ Formuleer (2 sets) hypothesen.
  - ▶ Teken een conceptueel model van deze hypothesen.
  - ▶ Toets deze met een meervoudige regressie in SPSS.
  - ▶ Interpreteer de uitkomst.
- 

# Opdracht 9: Antwoord (1)

## ▶ Attitudeoptreden

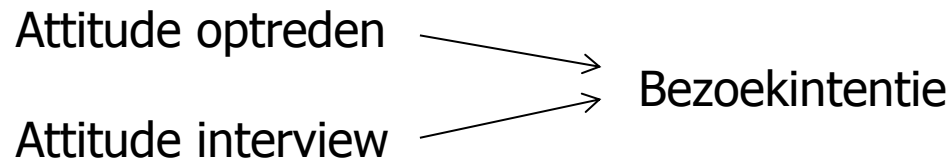
- H0: Attitudeoptreden heeft geen invloed op de bezoekintentie
- H1: Attitudeoptreden heeft wel een invloed op de bezoekintentie

## ▶ Attitudeinterview

- H0: Attitudeinterview heeft geen invloed op de bezoekintentie
  - H1: Attitudeinterview heeft wel een invloed op de bezoekintentie
- 

# Opdracht 9: Antwoord (2)

- ▶ Conceptueel model:





# Opdracht 9: Antwoord (3)

- ▶ De R square is 3%
- ▶ Het model (ANOVA) is niet significant.
- ▶ Attitudeinterview en attitudeoptreden hebben geen significante invloed op bezoekingtentie.
- ▶ Van beide hypothesen mag de H0 aangehouden worden.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,175 <sup>a</sup>	,031	,017	,437

a. Predictors: (Constant), attitudeoptreden, attitudeinterview

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,853	2	,427	2,236	,111 <sup>b</sup>
	Residual	26,897	141	,191		
	Total	27,750	143			

a. Dependent Variable: bezoekingtentie1

b. Predictors: (Constant), attitudeoptreden, attitudeinterview

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,946	,443		8,914	,000
	attitudeinterview	,031	,060	,046	,517	,606
	attitudeoptreden	,166	,098	,153	1,702	,091

a. Dependent Variable: bezoekingtentie1

# Voorwaarden voor regressie

- ▶ VIF-waarde
- ▶ Geeft aan of er sprake kan zijn van multicollineariteit. Dit wil zeggen dat de onafhankelijke variabelen zeer sterk correleren (en dus hetzelfde meten).
  - $VIF < 3$ : waarschijnlijk geen sprake van multicollineariteit.
  - $VIF > 3$ : waarschijnlijk sprake van multicollineariteit maar niet problematisch/
  - $VIF > 10$ : zeker sprake van multicollineariteit en moet opgelost worden.
- Oplossing: kijk in correlatiematrix of er inderdaad hoge correlaties zijn tussen onafhankelijke variabelen en voer een factoranalyse uit. Vervolgens een nieuwe regressie met 'nieuwe' factoren.

# VIF: how to...

John Legend.sav [DataSet3] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Visible: 8 of 8 Variables

	respondent	geslacht	leeftijd	opleiding	attitudeint1	attitudeinterview	attitudeoptreden	bezoekintentie	var	var	var	var	var
1	Respondent 1	2	23	2	3	3,00	5,00	5					
2	Respondent 2						5,00	5					
3	Respondent 3						3,50	5					
4	Respondent 4						5,00	5					
5	Respondent 5						5,00	5					
6	Respondent 6						5,00	5					
7	Respondent 7						5,00	5					
8	Respondent 8						5,00	5					
9	Respondent 9						5,00	4					
10	Respondent 10						5,00	5					
11	Respondent 11						5,00	5					
12	Respondent 12						5,00	5					
13	Respondent 13						5,00	5					
14	Respondent 14						5,00	5					
15	Respondent 15						5,00	5					
16	Respondent 16						3,25	4					
17	Respondent 17						5,00	5					
18	Respondent 18						5,00	5					
19	Respondent 19						5,00	5					
20	Respondent 20						4,25	4					
21	Respondent 21						5,00	5					
22	Respondent 22	2	22	3	5	5,00	5,00	5					
23	Respondent 23	2	24	3	5	5,00	5,00	5					

Linear Regression

Dependent: respondent

Linear Regression: Statistics

Regression Coefficients

- ☐ Estimates
- ☐ Confidence intervals
- Level(%): 95
- ☐ Covariance matrix
- ☐ Model fit
- ☐ R squared change
- ☐ Descriptives
- ☐ Part and partial correlations
- ☒ Collinearity diagnostics

Residuals

- ☐ Durbin-Watson
- ☐ Casewise diagnostics
- ☒ Outliers outside: 3 standard deviations
- ☒ All cases

Continue Cancel Help

OK Paste Reset Cancel Help

Data View Variable View

IBM SPSS Statistics Processor is ready

23:15  
22-3-2015

# VIF: how to

- ▶ In output krijg je een extra tabel.

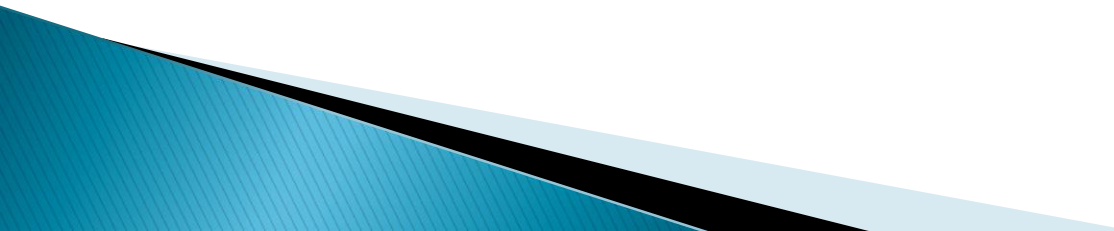
**Coefficients<sup>a</sup>**

Model		Collinearity Statistics	
		Tolerance	VIF
1	attitudeinterview	,857	1,167
	attitudeoptreden	,857	1,167

a. Dependent Variable: bezoekintentie1

- ▶ De VIF-waardes zijn kleiner dan 3, dus er is geen sprake van multicollineariteit en de resultaten van de regressie mogen geïnterpreteerd worden.

# Deze les

- ▶ **Leerdoel 1:** studenten kunnen het verschil tussen een verband (correlatie) en een causale relatie uitleggen.
  - ▶ **Leerdoel 2:** studenten kunnen een correlatie analyse uitvoeren en interpreteren.
  - ▶ **Leerdoel 3:** Studenten kunnen een enkelvoudige en meervoudige regressie analyse uitvoeren en interpreteren.
- 

# Huiswerk

- Maak de laatste bonusopdracht 7. Inleveren paperday (postvak docent).
  - Let op: onderbouw goed je antwoorden!
  - Bundel alle opdrachten in één document.
- Alle stof herhalen; volgende les tentamentraining!