

INDUCTIEVE STATISTIEK VOOR DE GEDRAGSWETENSCHAPPEN

OPLOSSINGEN BIJ HOOFDSTUK 7

1. Een onderzoeker wil nagaan of de fitheid van jongeren tussen 14 en 18 jaar (laag, matig, hoog) en het geslacht (M, V) een verschil maakt op de mate waarin jongeren zich kunnen concentreren op een concentratietest (score op 100, hoe hoger de score hoe beter men zich kan concentreren). Je mag aannemen dat de concentratiescores normaal verdeeld zijn in de populatie. Welke techniek heb je nodig om een antwoord te geven op de onderzoeksvraag? Breng onderstaande gegevens in SPSS op de correcte wijze en bereken je resultaat.

M	M	M	M	M	M	M	M	M	M	M	M	M	M
La	La	La	La	La	La	La	Mid	Mid	Mid	Mid	Mid	Mid	Mid
65	56	83	68	64	47	59	63	93	65	75	68	68	51

M	M	M	M	M	M	V	V	V	V	V	V	V	V	V
Ho	Ho	Ho	Ho	Ho	Ho	La	La	La	La	La	La	La	La	La
41	63	51	91	88	56	89	80	63	89	61	77	65	91	88

V	V	V	V	V	V	V	V	V	V	V
Mid	Mid	Mid	Mid	Mid	Mid	Ho	Ho	Ho	Ho	Ho
64	65	45	63	64	50	55	79	78	84	85

In deze opgave zijn de onderzoekseenheden zijn de 40 jongeren.

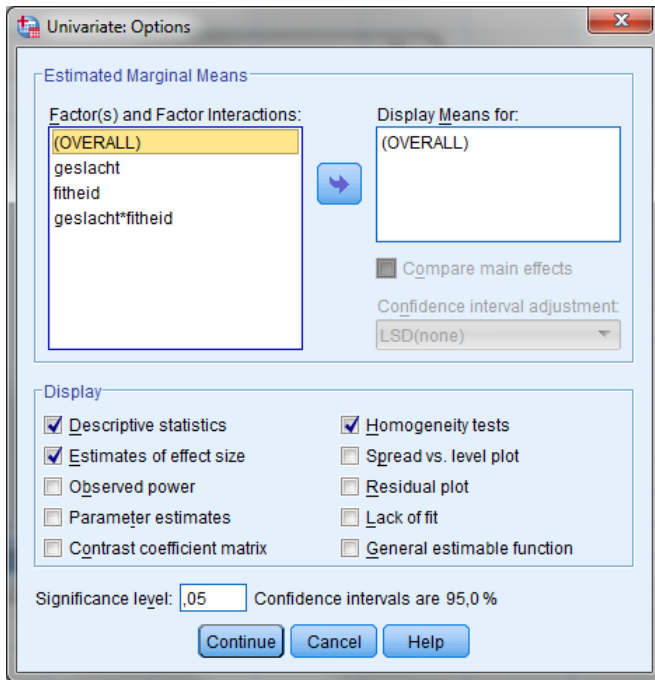
Er zijn drie variabelen : geslacht, fitheid en score. De vraag is: wat is het effect van fitheid (3 niveaus) en geslacht (2 niveaus) op de concentratie van jongeren? We kunnen aannemen dat de score op de concentratietest (afhankelijke variabele) op intervalniveau wordt gemeten. Er zijn twee onafhankelijke variabelen die op ordinaal of nominaal niveau gemeten worden. Daarom wordt een ANOVA gebruikt. Aangezien het gaat om twee OV's, is dit een tweewegs ANOVA.

Via een n-way anova onderzoeken we het mogelijke effect van fitheid en geslacht op de score van concentratie. Bij een n-way anova worden alle onafhankelijke variabelen tegelijkertijd in de analyse betrokken.

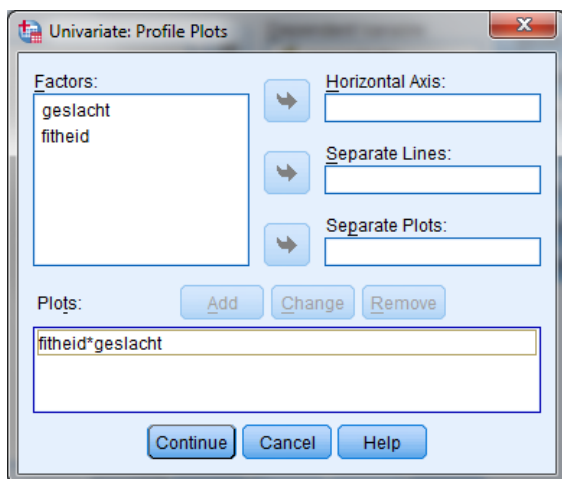
Kies voor Analyse > General Linear model > univariate

Nadien kies je in het keuze venster als dependent variabele de score en als fixed factors het geslacht en de mate van fitheid.

Om de assumpties voor een parametrische toets verder te toetsen, stellen we de vraag of de varianties in de subgroepen gelijklopend zijn. Vandaar dat we de Levene test doen door in het dialoogvenster onder Options te kiezen voor Homogeneity tests. In datzelfde venster kiezen we meteen ook de Descriptive Statistics en de Estimates of effect size.



Onder Post Hoc kiezen we nog voor een post hoc toets voor de OV Fitheid en onder Plots geven we op deze manier nog aan dat we een grafiekje willen van de resultaten:



Uit de Levene's test blijkt dat we de nulhypothese dat de varianties gelijklopend zijn, niet kunnen verwerpen. We kunnen dus een anova analyse uitvoeren, en als posthoc test (eventueel na significantie) kunnen we gebruik maken van de Bonferroni.

Levene's Test of Equality of Error Variances^a

Dependent Variable:concentratie

F	df1	df2	Sig.
1,295	5	34	,289

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + geslacht + fitheid + geslacht * fitheid

De 2-wegs ANOVA levert volgende resultaat op

Tests of Between-Subjects Effects

Dependent Variable:concentratie

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2001,454 ^a	5	400,291	2,277	,069	,251
Intercept	180681,265	1	180681,265	1027,621	,000	,968
geslacht	263,927	1	263,927	1,501	,229	,042
fitheid	410,315	2	205,157	1,167	,324	,064
geslacht * fitheid	1268,544	2	634,272	3,607	,038	,175
Error	5978,046	34	175,825			
Total	197042,000	40				
Corrected Total	7979,500	39				

a. R Squared = ,251 (Adjusted R Squared = ,141)

We onderzoeken hierbij of er twee hoofdeffecten bestaan (nl. van fitheid en van geslacht) en of er een interactie-effect is van fitheid en geslacht op de score.

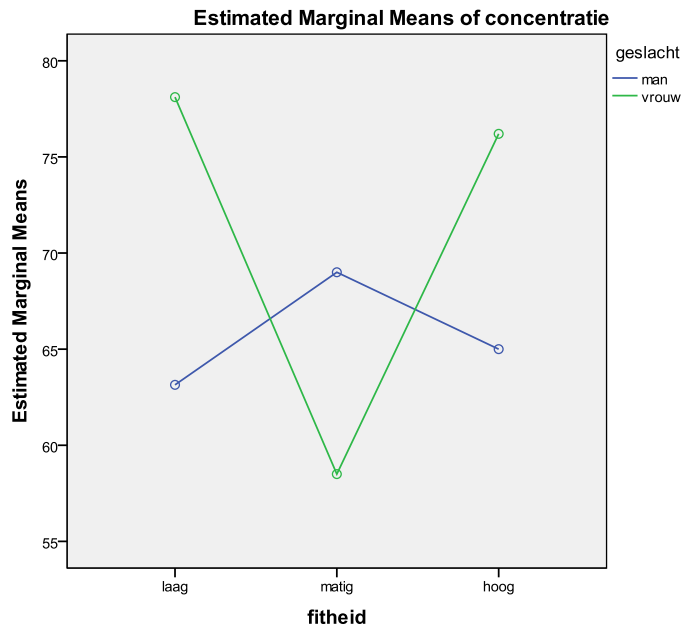
Voor de interpretatie van de output kijken we eerst naar de mogelijke hoofdeffecten.

We stellen vast dat er op 5% significantieniveau geen hoofdeffect is. Er is wel een significant interactie-effect, $p = .039$.

Aangezien er geen hoofdeffect is voor Fitheid, zullen we de gevraagde Post-Hoc tabel dan ook niet verder interpreteren. We kijken wel nog even verder naar het interactie-effect.

Het effect van fitheid is blijkbaar anders voor jongens dan voor meisjes. Of anders gezegd: het verschil tussen jongens en meisjes is anders naargelang de fitheid (bv. bij lage en hoge fitheid: jongens < meisjes; bij middelmatige fitheid meisjes < jongens).

Via plots kunnen we dit interactie-effect illustreren.



We kunnen verder nagaan hoe het interactie-effect zich precies voordoet door een syntax-opdracht te geven:

```
GLM concentratie by geslacht fitheid
/emmeans = tables(fitheid*geslacht)compare(fitheid).
```

Dat levert deze tabel op:

Pairwise Comparisons

Dependent Variable: concentratie

geslacht	(I) fitheid	(J) fitheid	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
man	laag	matig	-5,857	7,088	,414	-20,261	8,547
		hoog	-1,857	7,377	,803	-16,849	13,135
	matig	laag	5,857	7,088	,414	-8,547	20,261
		hoog	4,000	7,377	,591	-10,992	18,992
	hoog	laag	1,857	7,377	,803	-13,135	16,849
		matig	-4,000	7,377	,591	-18,992	10,992
vrouw	laag	matig	19,611 [*]	6,989	,008	5,409	33,814
		hoog	1,911	7,396	,798	-13,119	16,942
	matig	laag	-19,611 [*]	6,989	,008	-33,814	-5,409
		hoog	-17,700 [*]	8,029	,034	-34,017	-1,383
	hoog	laag	-1,911	7,396	,798	-16,942	13,119
		matig	17,700 [*]	8,029	,034	1,383	34,017

Daarin stellen we vast dat er bij de vrouwen significante verschillen zijn tussen de lage en matige niveaus van fitheid enerzijds en de hoge en matige niveaus van fitheid anderzijds.

Rapportering:

Er werd een 2 (Geslacht) X 3 (Fitheid) - ANOVA uitgevoerd om na te gaan of de concentratie wordt beïnvloed door de onafhankelijke variabelen. De concentratiescores waren hoger voor vrouwen ($M = 71.75$, $SD = 13.99$) dan voor mannen ($M = 65.75$, $SD = 14.38$), maar dit hoofdeffect was niet significant, $F = 1.50$, $p = .23$, $r = .18$). De concentratiescores verschilden ook naargelang de jongeren een lage ($M = 71.56$, $SD = 13.74$), matige ($M = 64.15$, $SD = 11.99$) of hoge fitheid ($M = 70.09$, $SD = 17.33$) hadden, maar ook dit hoofdeffect was niet significant, $F = 1.17$, $p = .32$, $r = .22$).

Er werd wel een significant interactie-effect tussen geslacht en fitheid gevonden ($F = 3.61$, $p = .038$, $r = .40$). Het effect van fitheid is blijkbaar anders voor jongens dan voor meisjes: bij lage ($M = 78.11$, $SD = 12.24$) en hoge ($M = 76.20$, $SD = 12.24$) fitheid scoren meisjes significant hoger dan bij matige fitheid ($M = 58.50$, $SD = 8.69$). Bij jongens is er geen significant verschil tussen de niveaus van fitheid (laag: $M = 63.14$, $SD = 11.19$; matig: $M = 69$, $SD = 12.85$; hoog: $M = 65$, $SD = 20.31$).

2. Een HR-manager wil nagaan of de 3 interviewers van de HR-dienst dezelfde maatstaven gebruiken. Daarvoor verzamelt hij een aantal totaalscores die de 3 interviewers gaven aan sollicitanten. Hij wil weten of de 3 interviewers gemiddeld gezien ongeveer dezelfde scores gaven. In onderstaand schema vind je de scores van de interviewers. Deze scores zijn niet normaal verdeeld in de populatie. Ga na of er een verschil is tussen de interviewers.

Interviewer 1:

56	65	98	99	69	96	56	99	91	58	53	73	92	97	72	65	51	67	88	87	60	80
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Interviewer 2:

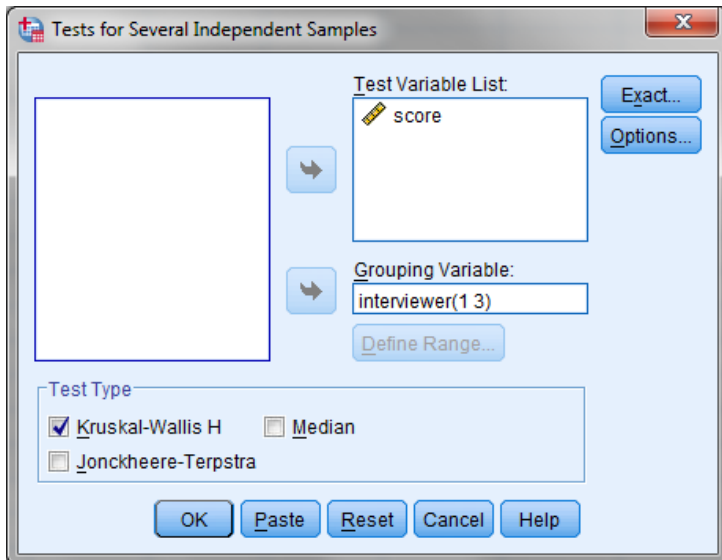
55	77	70	72	54	56	82	62	57	64	80	64	40	65	57	60	43	47	71	47	54
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Interviewer 3:

61	62	88	64	69	67	91	89	70	66	53	65	51	92	70	61	88	51	60	76	59	49	66	85
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

In dit onderzoek bestuderen we 2 variabelen: we willen weten of er een effect is van interviewer (nominaal – 3 niveaus) op de sollicitatiescores (interval). Aangezien de OV 3 niveaus bevat komen we terecht bij een one-way ANOVA. Er is echter gegeven dat de scores niet normaal verdeeld zijn in de populatie, waardoor we genooddaakt zijn een nonparametrische toets te kiezen, nl Kruskal-Wallis. (Eventueel kan je ook een one-way ANOVA overwegen, omdat de schending van normaliteit minder erg is als de groepen ongeveer even groot zijn.)

Kies in SPSS Analyze > Nonparametric tests > Legacy Dialogs > K independent samples



Aan de gemiddelde rangen zien we dat interviewer 1 hogere scores geeft dan interviewer 3 en deze op zijn beurt hogere scores geeft dan interviewer 2.

	interviewer	N	Mean Rank
score	1	22	42,20
	2	21	24,67
	3	24	34,65
	Total	67	

In de volgende tabel kunnen we vaststellen dat dit effect van interviewer ook significant is:

	score
Chi-Square	8,754
df	2
Asymp. Sig.	,013

a. Kruskal Wallis Test

b. Grouping Variable:
interviewer

Ten slotte gaan we na tussen welke interviewers dit effect zich voordoet, aan de hand van enkele Mann-Whitney toetsen. Aangezien we 3 toetsen nodig hebben, passen we het alfaniveau aan tot $.05/3 = .017$.

Uit deze toetsen blijkt dat er een significant verschil is tussen interviewer 1 en 2 ($p = .005$), maar niet tussen de andere combinaties. Het verschil tussen interviewer 2 en 3 benadert wel significantie met $p = .067$.

Rapportering:

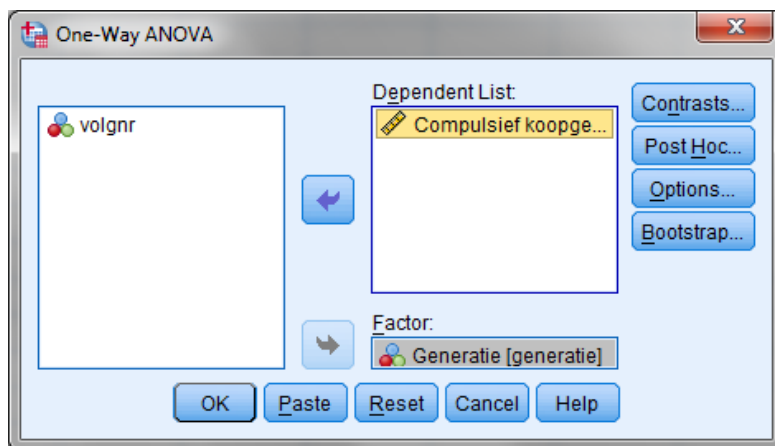
Een Kruskal-Wallis toets werd uitgevoerd om het effect van interviewer op de sollicitatiescores na te gaan. Dit effect bleek inderdaad significant, $H = 8.75$, $p = .013$. Bijkomend werden interviewer 1 (mean

rank = 42.20), interviewer 2 (mean rank = 24.67) en interviewer 3 (mean rank = 34.65) onderling vergeleken door middel van een Wilcoxon rank-sum toets, waarbij een gecorrigeerd significantieniveau van $\alpha = .017$ werd gehanteerd. Hieruit bleek dat er enkel een significant verschil was tussen interviewers 1 en 2 ($W_s = 346.5$, $z = -2.8$, $p = .005$, $r = -.43$). Het verschil tussen interviewers 1 en 3 ($W_s = 499$, $z = -1.43$, $p = .15$, $r = -.21$) was niet significant. Het verschil tussen interviewers 2 en 3 benaderde slechts significantie ($W_s = 402.5$, $z = -1.833$, $p = .067$, $r = -.27$).

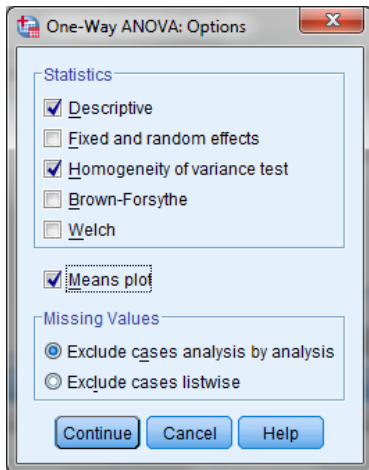
3. We onderzoeken of de leeftijd een invloed heeft op het compulsief koopgedrag. De proefpersonen werden verdeeld in generatie Y (twintigplussers), generatie X (dertigplussers) en babyboomers (40 a 60 jarigen). Alle proefpersonen kregen een score op de Faber en O'Guinn test voor compulsief koopgedrag (intervalschaal, normaal verdeeld). Je vindt de scores in de onderstaande tabel. Plaats de gegevens in SPSS en ga na of de leeftijd effectief een invloed heeft op dit gedrag.

Gen Y	Gen. X	babyboomers
65.00	42.00	66.00
72.00	50.00	56.00
55.00	44.00	22.00
63.00	66.00	33.00
69.00	22.00	35.00
55.00	30.00	40.00
	45.00	28.00
	45.00	39.00

We bestuderen twee variabelen, generatie (nominaal, 3 niveaus) en koopgedrag (interval). De drie niveaus van de OV impliceren 3 populaties, dus we overwegen een one-way ANOVA. Er is voldaan aan de assumpties voor parametrisch toetsen, dus we kiezen in SPSS voor Analyze > Compare Means > one-way ANOVA. (Dit kan ook via General Linear Models – de dialogvensters zien er dan wel anders uit).



Onder Post Hoc kiezen we nog Bonferroni en onder Options duiden we nog aan:



De Levene statistic is niet significant, dus we mogen aannemen dat de varianties homogeen zijn in de populatie.

In de tabel met beschrijvende gegevens zien we dat er meer compulsief koopgedrag voorkomt bij Generatie Y dan bij Generatie X, en meer bij Generatie X dan bij de Babyboomers.

Descriptives

Compulsief koopgedrag

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Generatie Y	6	63,17	7,055	2,880	55,76	70,57	55	72
Generatie X	8	43,00	13,082	4,625	32,06	53,94	22	66
Babyboomers	8	39,88	14,515	5,132	27,74	52,01	22	66
Total	22	47,36	15,463	3,297	40,51	54,22	22	72

De ANOVA-tabel toont dat dit effect ook significant is:

ANOVA

Compulsief koopgedrag

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2099,383	2	1049,691	6,826	,006
Within Groups	2921,708	19	153,774		
Total	5021,091	21			

In de post hoc tabel kunnen we vaststellen dat de significante verschillen zich voordoen tussen Generatie Y en Generatie X, en ook tussen Generatie Y en de Babyboomers.

Multiple Comparisons

Dependent Variable: Compulsief koopgedrag

Bonferroni

(I) Generatie	(J) Generatie	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Generatie Y	Generatie X	20,167*	6,697	,022	2,59	37,75
	Babyboomers	23,292*	6,697	,008	5,71	40,87
Generatie X	Generatie Y	-20,167*	6,697	,022	-37,75	-2,59
	Babyboomers	3,125	6,200	1,000	-13,15	19,40
Babyboomers	Generatie Y	-23,292*	6,697	,008	-40,87	-5,71
	Generatie X	-3,125	6,200	1,000	-19,40	13,15

*. The mean difference is significant at the 0.05 level.

Rapportering:

Er was een significant verschil tussen de generaties wat betreft compulsief koopgedrag, $F(2, 19) = 6.83$, $p = .006$, $r = .65$. Generatie Y vertoonde meer compulsief koopgedrag ($M = 63.17$, $SD = 7.06$) dan Generatie X ($M = 43$, $SD = 13.01$) en de Babyboomers ($M = 39.88$, $SD = 14.52$). Generatie X verschilde niet significant van de Babyboomers.

4. Uit De Standaard, 27 mei 2011: *Autorijden in een wolkje pepermunt: het maakt u meteen rustiger. Dat beweren drie Nederlandse studenten uit de provincie Gelderland. Inspiratie haalden ze bij de Amerikaanse onderzoeker Bryan Raudenbush, die in 2001 en 2002 verschillende geuren vernevelde over atleten die een stresstest op een loopband uitvoerden. De atleten die pepermunt roken waren duidelijk alterter en minder gefrustreerd.*

'Als het op een loopband werkt, dan ook achter het stuur', dachten de studenten. Raudenbush zelf stelde dat al vast in een rijnsimulator, de Nederlandse studenten –stagiairs bij het Regionaal orgaan verkeersveiligheid van Gelderland– stuurden hun proefpersonen écht de weg op en kwamen tot dezelfde conclusie.

Munt, babyolie of niets

De proefpersonen –27 mannen en 49 vrouwen– werd wijsgemaakt dat ze een nieuw soort navigatietoestel uittesten. Een camera legde hun gedrag tijdens het rijden vast. Bij één op de drie proefpersonen werd een pepermuntgeur in de auto verspreid, bij een ander derde de geur van babyolie en de rest moest het zonder geurtje stellen. De studenten berekenden de gemiddelde snelheid van de proefpersonen op 16 wegstukken met een maximumsnelheid tussen de 30 en de 80 km per uur.

De data van dit onderzoek vindt je in het bestand *Opdr_Pepermunt.sav*. Ga na of de deelnemers inderdaad trager reden in de pepermuntconditie. Wat concludeer je?

De variabelen in dit onderzoek zijn geslacht, geurconditie (3 niveaus, nominaal) en gemiddelde snelheid (interval). De onderzoeksvraag gaat echter enkel over het verschil tussen de geurcondities, dus we krijgen te maken met een one-way ANOVA, op voorwaarde dat voldaan is aan de assumpties voor parametrisch toetsen.

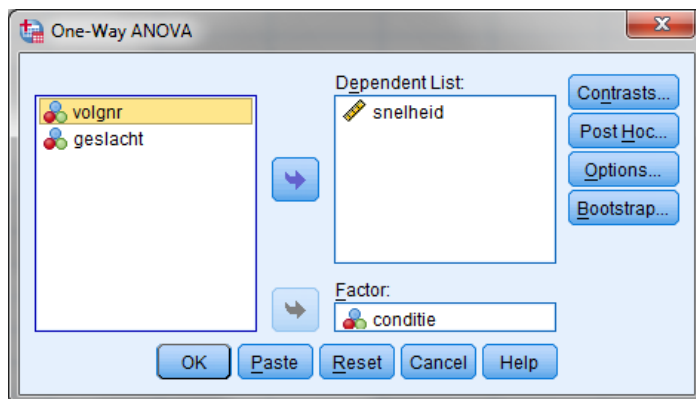
Er zijn minder dan 30 deelnemers per conditie, dus we toetsen de assumptie van normaliteit met behulp van het Explore-commando.

Tests of Normality							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
snelheid	pepermunt	,140	25	,200 [*]	,956	25	,336
	babyolie	,147	25	,169	,967	25	,577
	geen geur	,136	26	,200 [*]	,936	26	,105

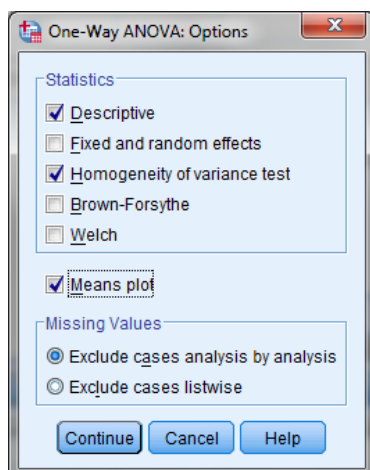
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Hieruit blijkt dat de nulhypothese van normaliteit kan behouden worden, dus we gaan door met de ANOVA, via Analyze > Compare Means > one-way ANOVA (of alternatief via General Linear Model).



Onder Post Hoc kiezen we nog Bonferroni en onder Options duiden we nog aan:



De Levene statistic is niet significant, dus we mogen aannemen dat de varianties homogeen zijn in de populatie.

Aan de beschrijvende statistieken zien we dat er minder snel werd gereden in de pepermint-conditie dan in de andere twee condities. De ANOVA-tabel toont dat dit effect ook significant is.

Descriptives

snelheid

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
pepermunt	25	50,6534	10,08580	2,01716	46,4902	54,8166	33,87	71,17
babyolie	25	58,1079	10,82306	2,16461	53,6404	62,5755	31,65	75,73
geen geur	26	59,3136	13,28386	2,60518	53,9482	64,6791	38,46	80,89
Total	76	56,0683	11,98615	1,37491	53,3293	58,8072	31,65	80,89

ANOVA

snelheid

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1110,880	2	555,440	4,196	,019
Within Groups	9664,210	73	132,386		
Total	10775,089	75			

In de post hoc tabel zien we dat er enkel een significant verschil is tussen de peppermunt-conditie en de conditie zonder geur.

Multiple Comparisons

Dependent Variable: snelheid

Bonferroni

(I) conditie	(J) conditie	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
pepermunt	babyolie	-7,45456	3,25437	,075	-15,4291	,5199
	geen geur	-8,66028*	3,22292	,027	-16,5577	-,7628
babyolie	pepermunt	7,45456	3,25437	,075	-,5199	15,4291
	geen geur	-1,20572	3,22292	1,000	-9,1032	6,6917
geen geur	pepermunt	8,66028*	3,22292	,027	,7628	16,5577
	babyolie	1,20572	3,22292	1,000	-6,6917	9,1032

*. The mean difference is significant at the 0.05 level.

Rapportering:

Om na te gaan of er een effect is van geur op de snelheid waarmee de deelnemers reden werd een one-way ANOVA uitgevoerd. Er was een significant effect van geurconditie op de gemiddelde snelheid van de deelnemers, $F(2, 73) = 4.20, p = .019, r = .32$. De deelnemers in de peppermuntconditie reden significant minder snel ($M = 50.65, SD = 10.09$) dan de deelnemers in de conditie zonder geur ($M = 59.31, SD = 13.28$). De overige verschillen tussen peppermuntconditie, conditie zonder geur en de babyolieconditie ($M = 58.11, SD = 10.82$) waren niet significant.