

## Lösung zu Aufgabe 16.4

Zum Einlesen der Daten:

Die Aufgabe wird mit dem Statistik-Programm SPSS gelöst. Hierzu muss zunächst der Datensatz *absol95.txt* in den SPSS-Editor eingelesen werden. Laden Sie dazu das Datenfile *absol95.txt* und die Hilfsdatei *readme.txt* aus dem Internet herunter. Die Daten liegen im sog. ASCII-freefield-Format vor.

Öffnen Sie zuerst in einem beliebigen Texteditor die *readme.txt*-Datei. Wählen Sie anschließend im SPSS-Editor in der Menüzeile den Punkt *File® Read ASCII Data® Freefield*. Klicken Sie den *Browse*-Button und arbeiten Sie sich dann durch die Verzeichnisse bis zu *absol95.txt* durch.

Im nächsten Schritt müssen die Variablen definiert werden. Die Variablennamen und -kodierungen können Sie der Datei *readme.txt* entnehmen. Geben Sie dazu den Variablennamen in das Feld *Name* ein, wählen den Variablentyp *Numeric* und klicken den *Add*-Button. Nachdem Sie dies für alle sechs Variablen durchgeführt haben, klicken Sie *OK*. SPSS liest daraufhin die Daten ein.

In SPSS können Sie die Variablen und deren Ausprägungen noch mit den in der *readme.txt*-Datei angegebenen Labels versehen.

Die Lösungen der Teilaufgaben sind folgendermaßen aufgebaut:

Zunächst wird erklärt, wie man sich durch das Menü durchklickt, und welche Angaben gemacht werden müssen, um zu einer Lösung zu gelangen. Danach ist der vom Programm erzeugte Output wiedergegeben.

(a)

Vergleich des mittleren Einkommens in den verschiedenen Beschäftigungsverhältnissen:

Menüleiste *Statistics® Compare Means® Means*:

*Dependent: stlohn*

*Independent: beschver*

Als Output erscheint eine Tabelle, die neben den gesuchten Gruppenmittelwerten auch schon die in Teilaufgabe (b) gefragten Standardabweichungen enthält:

### Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Stundenlohn * Beschäftigungsverh.	182	100.0%	0	.0%	182	100.0%

## Report

### Stundenlohn

unbefristet	Mean	37.2855
	N	106
	Std.	
	Deviation	11.2755
befristet	Mean	30.7725
	N	50
	Std.	
	Deviation	6.9788
selbständig/freiberufl.	Mean	32.3901
	N	26
	Std.	
	Deviation	16.4617
Total	Mean	34.7969
	N	182
	Std.	
	Deviation	11.5456

Die Mittelwerte in den drei Gruppen unterscheiden sich deutlich voneinander. Absolventen in befristeten Beschäftigungsverhältnissen verdienen mit etwa 30.80 DM unterdurchschnittlich. Etwas besser sieht es bei den selbständig oder freiberuflich Tätigen aus. Auch ihr Verdienst liegt zwar unter dem Gesamtmittel von 34.80 DM, sie verdienen aber etwa 1.50 DM in der Stunde mehr als die befristet Angestellten. Der mittlere Stundenlohn von Absolventen in unbefristeten Beschäftigungsverhältnissen ist dagegen mit etwa 37.30 DM deutlich über dem Durchschnitt.

Auffallend sind auch die großen Unterschiede in den Standardabweichungen. Die geringste Variabilität mit einer Standardabweichung von 7 DM tritt in der Gruppe der befristet Beschäftigten auf. Mehr als doppelt so hoch ist die Standardabweichung bei den Selbständigen bzw. freiberuflich Tätigen.

(b1)

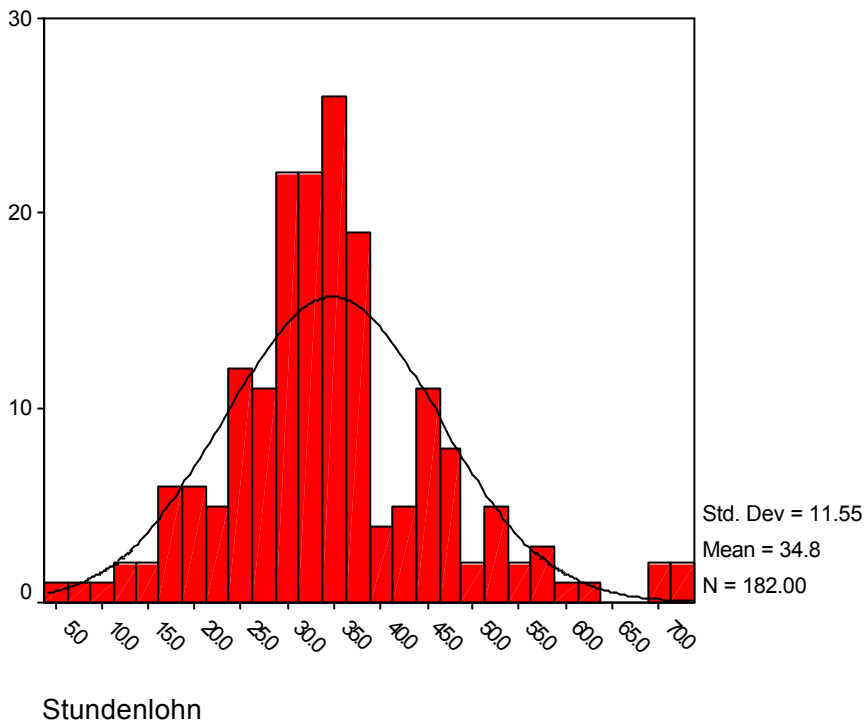
Wie in Teilaufgabe (a) schon bemerkt, unterscheiden sich die Standardabweichungen stark voneinander. Die Daten sprechen also eher gegen das Vorliegen einer Varianzhomogenität.

(b2)

Unter dem Menüpunkt *Graphs* befindet sich der Punkt *Histogram*.

Wählen Sie die Variable *stlohn* und zum Vergleich mit der Normalverteilung auch noch den Punkt *display normal curve*. SPSS erzeugt daraufhin ein Histogramm und eine Normalverteilungskurve mit den empirischen Momenten als Parameter.

## Graph



Die Annahme der Normalverteilung scheint aufgrund des Histogramms eher kritisch zu sein: Die Ränder der Verteilung sind im Gegensatz zur Normalverteilung relativ stark gewichtet. Außerdem weist die Verteilung eine gewisse Unsymmetrie auf.

(b3)

Eine weitere Möglichkeit, die Annahme der Normalverteilung zu beurteilen, bietet der Kolmogorov-Smirnov-Test.

Dieser befindet sich in dem Menüpunkt *Statistics* ® *Nonparametric tests* ® *1-Sample K-S*.

*Test variable: stlohn*

*Test distribution: normal*

## NPar Tests

### One-Sample Kolmogorov-Smirnov Test

		Stundenlohn
N		182
Normal Parameters <sup>a,b</sup>	Mean	34.7969
	Std. Deviation	11.5456
Most Extreme Differences	Absolute	.116
	Positive	.116
	Negative	-.058
Kolmogorov-Smirnov Z		1.568
Asymp. Sig. (2-tailed)		.015

a. Test distribution is Normal.

b. Calculated from data.

Die Testentscheidung wird hier mit Hilfe des  $p$ -Wertes durchgeführt. Aus den Daten erhält man einen  $p$ -Wert von 0.015. Da  $0.015 < 0.05$  kann die Annahme der Normalverteilung zum Signifikanzniveau  $\alpha=0.05$  verworfen werden .

Insgesamt muss also die Annahme der Normalverteilung sehr kritisch beurteilt werden. Außerdem geben die Daten den Hinweis, dass auch nicht selbstverständlich von einer Homogenität der Varianz ausgegangen werden kann.

(c)

Eine einfaktorielle Varianzanalyse kann in SPSS mit der Prozedur *ONEWAY* durchgeführt werden.

Hierzu *Statistics*® *Compare means*® *Oneway ANOVA* .

Wählen Sie *stlohn* als abhängige Variable und *beschver* als Faktor.

## Oneway

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Stundenlohn	Between Groups	1616.855	2	808.427	6.428	.002
	Within Groups	22510.562	179	125.757		
	Total	24127.416	181			

Der  $F$ -Test liefert einen  $p$ -Wert von 0.002. Geht man also davon aus, dass die Voraussetzungen der Varianzanalyse erfüllt sind, kann man zum Signifikanzniveau  $\alpha=0.05$  schließen, dass die Art des Beschäftigungsverhältnisses Einfluss auf den Stundenlohn hat.

Zur Durchführung von Paarvergleichen gibt es in SPSS keine vorgefertigte Prozedur; man muss deshalb den Umweg über sogenannte Filter einschlagen, die über die folgenden Menüpunkte realisiert werden:

*Data* ® *Select Cases* ® „*beschver*“ anklicken® *If condition is satisfied* ® *If* ® „*beschver=1 or beschver=2*“

Auf diese Weise erhält man einen reduzierten Datensatz, der nur noch aus den Fällen besteht, für die die Variable *beschver* die Werte „1“ oder „2“ annimmt. Im Editor sind die ausgesonderten Fälle durchgestrichen.

Führen Sie nun die Varianzanalyse wie oben durch:

*Statistics*® *Compare means*® *Oneway ANOVA*

## Oneway

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Stundenlohn	Between Groups	1441.149	1	1441.149	14.104	.000
	Within Groups	15735.851	154	102.181		
	Total	17177.000	155			

Wiederholen Sie nun dieselben Schritte mit der Filteranweisung „*beschver=1 or beschver=3*“.

## Oneway

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Stundenlohn	Between Groups	500.354	1	500.354	3.232	.075
	Within Groups	20124.051	130	154.800		
	Total	20624.405	131			

Führen Sie die Anweisungsfolge nun ein drittes Mal mit der Filterbedingung „*beschver=2 or beschver=3*“ aus.

## Oneway

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Stundenlohn	Between Groups	44.757	1	44.757	.362	.549
	Within Groups	9161.222	74	123.800		
	Total	9205.979	75			

Zum Signifikanzniveau  $\alpha=0.05$  kann man nur bei dem Vergleich von befristeten mit unbefristeten Beschäftigungsverhältnisse auf einen Einfluss auf den Stundenlohn schließen, da hier der  $p$ -Wert  $< 0.05$  ist. Der bei der Varianzanalyse mit dem kompletten Datensatz entdeckte Zusammenhang zwischen Beschäftigungsverhältnis und Stundenlohn ist also auf einen Unterschied im Stundenlohn zwischen den abhängig unbefristet und befristet Beschäftigten zurückzuführen.

Vor der Bearbeitung der nächsten Aufgaben müssen alle Fälle wieder aktiviert werden, d.h. die Filterung muss aufgehoben werden. Dies realisieren Sie, indem Sie *Data @ Select Cases @ All cases* klicken.

d)

Diese Fragestellung kann mit dem *t*-Test gelöst werden.

*Statistics @ Compare means @ Independent-Samples-T-Test*

*Test Variable:* *stlohn*

*Grouping Variable:* *geschl*

*@Define Groups*

*Group 1:* *0*

*Group 2:* *1*

## T-Test

**Group Statistics**

	Geschlecht	N	Mean	Std. Deviation	Std. Error Mean
Stundenlohn	männlich	81	37.7469	12.4038	1.3782
	weiblich	101	32.4310	10.2707	1.0220

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Mean	
									Lower	Upper
Stundenlohn	Equal variances assumed	3.315	.070	3.163	180	.002	5.3158	1.6808	1.9993	8.6324
	Equal variances not assumed			3.098	154.738	.002	5.3158	1.7158	1.9265	8.7052

Der *p*-Wert (in der Spalte *Sig.*) ist sowohl unter der Annahme gleicher Varianzen, als auch unter Varianzheterogenität mit 0.002 signifikant. Die genannte Vermutung, dass der Stundenlohn von Männern im Schnitt höher ist als der von Frauen, lässt sich also bestätigen.

(e)

Zuerst wird eine neue Variable erzeugt, die die Variable Fachadäquanz in drei Kategorien einteilt. In SPSS benötigt man hierzu den Befehl *Compute*, den man unter dem Menüpunkt *Transform* findet. Als *Target-Variable* wählt man den Namen für die neue Variable, z.B. *kat*, und teilt die Ausprägungen der ursprünglichen Variable Fachadäquanz folgendermaßen den neugebildeten Kategorien zu:

Numeric Expression „1“ ® If ® Include if case... ® adaequat < 9® Continue® OK  
 Numeric Expression „2“ ® If ® Include if case... ®  
 adaequat >8 and adaequat <17® Continue® OK  
 Numeric Expression „3“ ® If ® Include if case... ® adaequat > 16® Continue® OK

Nachdem nun die Variablen in der benötigten Form vorliegen, können die deskriptiven Maße für die Stärke des Zusammenhangs berechnet werden. Dies geschieht mit der Prozedur *Crosstabs*. Zuerst wird der Zusammenhang zwischen *zufried* und *wunsch* betrachtet:

*Statistics* ® *Summarize* ® *Crosstabs*.  
*Rows* zufried  
*Column* wunsch  
*Statistics* Contingency coefficient, Chi-Square

Es wird nun eine Kontingenztafel erstellt, der Kontingenzkoeffizient berechnet und der  $\chi^2$ -Test durchgeführt.

## Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
berufliche Zufriedenheit * WUNSCH	182	100.0%	0	.0%	182	100.0%

berufliche Zufriedenheit \* WUNSCH Crosstabulation

Count

		WUNSCH		Total
		keine Übereinst.	Übereinst.	
berufliche Zufriedenheit	sehr unzufrieden	5	4	9
	unzufrieden	15	11	26
	zufrieden	53	56	109
	sehr zufrieden	15	23	38
Total		88	94	182

### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.298 <sup>a</sup>	3	.513
Likelihood Ratio	2.311	3	.510
Linear-by-Linear Association	2.011	1	.156
N of Valid Cases	182		

a. 2 cells (25.0%) have expected count less than 5.  
The minimum expected count is 4.35.

### Symmetric Measures

	Value	Approx. Sig.
Nominal by Nominal Contingency Coefficient	.112	.513
N of Valid Cases	182	

- a. Not assuming the null hypothesis.  
b. Using the asymptotic standard error assuming the null hypothesis.

Der Kontingenzkoeffizient von 0.112 lässt keinen starken Zusammenhang zwischen der beruflichen Zufriedenheit und der Übereinstimmung mit dem Berufswunsch vermuten. Dies wird durch den hohen  $p$ -Wert des  $\chi^2$ -Tests (0.513) bestätigt.

Da bei der Prozedur *Crosstabs* mit Option *Contingency coefficient* der  $\chi^2$ -Test schon standardmäßig integriert ist, wird im folgenden auf eine getrennte Angabe verzichtet. Die Beurteilung des Zusammenhangs zwischen der Fachadäquanz (bzw. der neu erzeugten Variable *kat*) und der Zufriedenheit (*zufried*) läuft völlig analog:

*Statistics* ® *Summarize* ® *Crosstabs*.  
*Rows* *zufried*  
*Column* *kat*  
*Statistics* *Contingency coefficient*



## Crosstabs

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
berufliche Zufriedenheit * KAT	182	100.0%	0	.0%	182	100.0%

### berufliche Zufriedenheit \* KAT Crosstabulation

Count

		KAT			Total
		1.00	2.00	3.00	
berufliche Zufriedenheit	sehr unzufrieden	6	2	1	9
	unzufrieden	14	10	2	26
	zufrieden	31	65	13	109
	sehr zufrieden	10	20	8	38
Total		61	97	24	182

### Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Contingency Coefficient	.264	.034
N of Valid Cases		182	

- a. Not assuming the null hypothesis.
- b. Using the asymptotic standard error assuming the null hypothesis.

Hier kann zum Signifikanzniveau  $\alpha=0.05$  von einem Zusammenhang zwischen der Fachadäquanz und der beruflichen Zufriedenheit ausgegangen werden, da der  $p$ -Wert des  $\chi^2$ -Tests  $0.034 < 0.05$  ist.