



Erstellen von statistischen Auswertungen mit Excel in den Sozialwissenschaften

Dr. Viola Vockrodt-Scholz
Telefon: 0172 17 17 791
Email:
Dr.Vockrodt-Scholz@edvdidaktik.de
Internet:
edvdidakik.de

EINFACHE STATISTISCHE AUSWERTUNGEN IN EXCEL	2
FRAGEBOGEN ZUM ERSTELLEN EINER BEISPIEL-STATISTIK IN EXCEL	2
STATISTISCHE KENNWERTE	3
Der Bereichsoperator	3
Der Funktionsassistent	3
HÄUFIGKEITSANALYSEN MIT PIVOTTABELLEN	4
Eine Matrix transponieren	10
Zahlen in einer Word-Tabelle mit dem Dezimaltabulator ausrichten	12
EIN WEITERES BEISPIEL FÜR EINE PIVOTTABELLE	13
KORRELATION	14
ANMERKUNGEN ZU STATISTIK MIT EXCEL	16

Einfache statistische Auswertungen in Excel

Einfache statistische Auswertungen lassen sich relativ einfach und rasch in Excel erstellen. Zur Grundlage der hier vorliegenden Berechnungen wurde ein Beispielfragebogen an einer Stichprobe mit 12 Personen erhoben.

Fragebogen zum Erstellen einer Beispiel-Statistik in Excel

Um eine Fragebogenauswertung in Excel zu simulieren, können wir einige Fragen beantworten. Die Ergebnisse werden wir in Excel eingeben und statistisch auswerten (Prozentuale Verteilung und Häufigkeit, ggf. Korrelationen). Die Ergebnisse werden mit einem Diagramm dargestellt.

Wie ist die Verteilung der Fachrichtungen im Seminar?

Wie ist die Verteilung des Geschlechts im Seminar?

Die Arbeitshypothese lautet:

Gibt es einen Zusammenhang zwischen der Selbsteinschätzung als Computernutzer und der Selbsteinschätzung im Umgang mit technischen Geräten?

Weitere Fragen:

Gibt es einen Zusammenhang zwischen der Dauer der Arbeit mit dem Computer und der Selbsteinschätzung als Computernutzer?

Gibt es einen Zusammenhang zwischen der Dauer der Arbeit mit dem Computer und dem Alter?

Fragebogen im Intensivseminar EDV II zur statistischen Auswertung mit der Tabellenkalkulation Excel

Wie viele Jahre arbeiten Sie mit einem Computer?
01

Jahre

Wie schätzen Sie sich selbst als
Computernutzer ein?
02

Als
Einsteiger

als
Profi

1 2 3 4 5

Fällt es Ihnen generell leicht, mit Technik umzugehen (HIFI, Radio,
Video, TV etc.)?
03

Stimmt
genau

Stimmt
gar nicht

1 2 3 4 5

Angaben zur Ihrer Person

Wie alt sind Sie?
04

___ Jahre

Ihr Geschlecht
05

männlich

weiblich

Ihr Studienfach / Berufliche Richtung
06

Statistische Kennwerte

Die statistischen Kennwerte lassen sich mit einigen Formeln berechnen. Notwendig sind hierzu Kenntnisse über den Bereichsoperator und über den Funktionsassistenten.

Der Bereichsoperator

Der Bereichsoperator besteht aus einer Formel, die einen Zahlenbereich berechnet. Der Zahlenbereich wird durch die Grenzzellen und einen Doppelpunkt bezeichnet.

Der Bereichsoperator

=SUMME(A2:A13)

berechnet die Summe der Zahlen in den Zellen A2 bis A13.

Ein Zahlenbereich ist auch der Bereich von A2 bis C15 etc.

Tabelle 1: Bereichsoperatoren für statistische Kennwerte in Excel.

	A	B	C
1		FB-Nr.	CompJahre
2		1	2
3		2	1
4		3	0
5		4	5
6		5	2
7		6	4
8		7	6
9		8	8
10		9	10
11		10	2
12		11	3
13		12	1
14			
15	Anzahl		=ANZAHL2(C2:C13)
16	Mittelwert		=MITTELWERT(C2:C13)
17	Minimum		=MIN(C2:C13)
18	Maximum		=MAX(C2:C13)
19	Median		=MEDIAN(C2:C13)
20	Modalwert		=MODALWERT(C2:C13)
21	Standardabweichung		=STABW(C2:C13)

Weitere Bereichsoperatoren für statistische Kennwerte finden Sie in Tabelle 1.

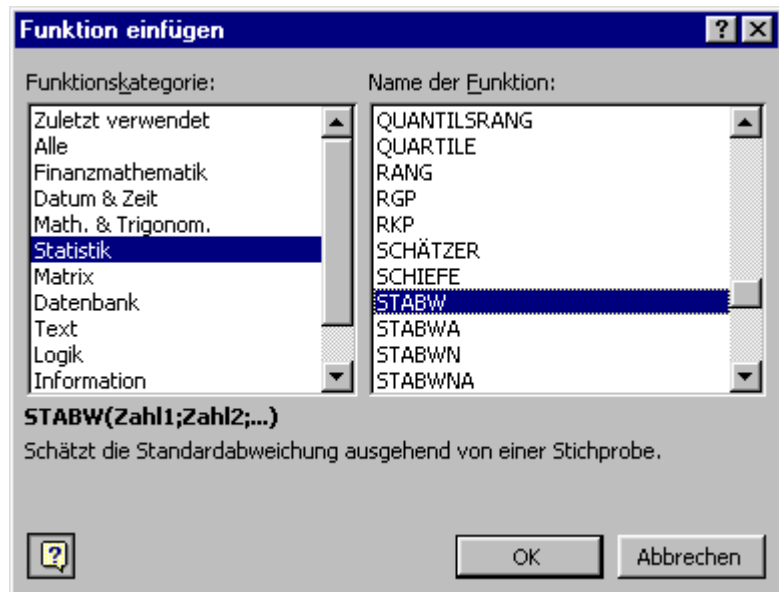
Der Funktionsassistent

Für die statistischen Berechnungen verwenden Sie am einfachsten den Funktionsassistenten.



Sie finden hier eine Kategorie „Statistik“. Im rechten Fenster des Assistenten finden Sie alle statistischen Funktionen, die Excel standardmäßig zur Verfügung stellt. Die Formeln werden

direkt im Assistenten erläutert. Wenn Sie auf das Fragezeichen klicken, erhalten Sie Erklärungen zur Funktion und einige Anwendungsbeispiele.



Häufigkeitsanalysen mit Pivottabellen

Pivottabellen sind in Excel die flexibelste Möglichkeit, um Kreuztabellen und/oder Häufigkeitsauszählungen zu erstellen. Alternativ können mit der Funktion „Teilergebnisse“ Häufigkeiten einzelner Variablen ausgezählt werden. Sie finden weitere Informationen zu beiden Funktionen in den Unterlagen zu Excel 97/2000.

Zusätzlich ist das Erstellen eines **Pivotdiagramms** vereinfacht worden. Das Pivotdiagramm ist zur Analyse der Daten in Excel sehr gut geeignet, weniger jedoch, um Diagramme in einem Word-Bericht zu verwenden. Um Ihre Daten in einem Diagramm darzustellen, verwenden Sie den Diagrammassistenten, wie Sie es in den weiteren Excel-Unterlagen ausführlich beschrieben finden.

Die Pivottabellen in Excel 2000 sind etwas umgestaltet worden, so dass die Handhabung etwas leichter geworden ist. Inhaltlich hat sich in der einfachen Anwendung nichts verändert. Zum Erstellen einer Pivottabelle ist die Fragestellung wichtig.

Wir analysieren die Häufigkeit der Verteilung der Geschlechter in der Stichprobe.

1. Markieren der Daten

Sie markieren die Daten durch Klick mit der Maus innerhalb der Daten. Es reicht, wenn ein Feld innerhalb der Datenmatrix markiert ist.

2. Aufrufen des Assistenten zum Erstellen einer Pivottabelle

Sie wählen den Befehl

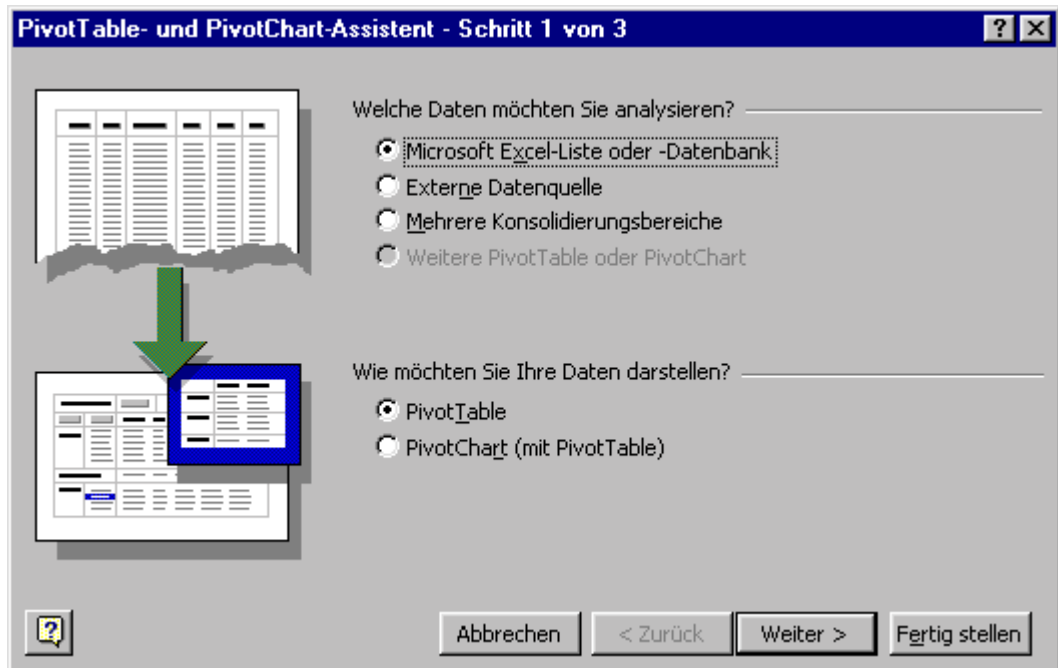
Daten > Pivottable- oder Pivotchart-Bericht

Alternativ klicken Sie in der Standard-Symbolleiste „Schaltflächen hinzufügen oder entfernen“. Sie finden dort eine Schaltfläche „Pivot Table- oder Pivot Chart-Bericht“.



3. Erstellen einer Pivottabelle

Nach Aktivierung des Befehls „Pivot Table- oder Pivot Chart-Bericht“ startet Excel einen Assistenten.



Im ersten Schritt des Assistenten werden Sie nach der Datenquelle für die Pivottabelle gefragt und danach, ob Sie eine Tabelle oder ein Diagramm erstellen wollen.

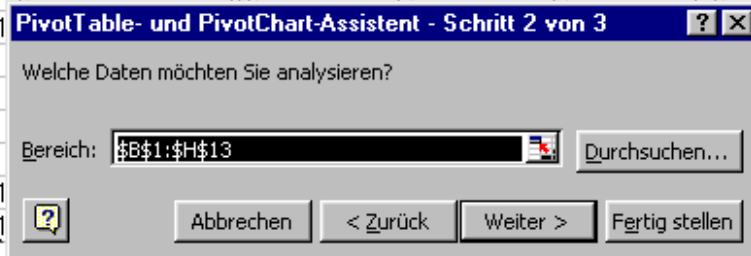
Wir wählen „Microsoft Excel-Liste oder –Datenbank“ und „PivotTable“.

Sie klicken auf „Weiter“.

4. Schritt

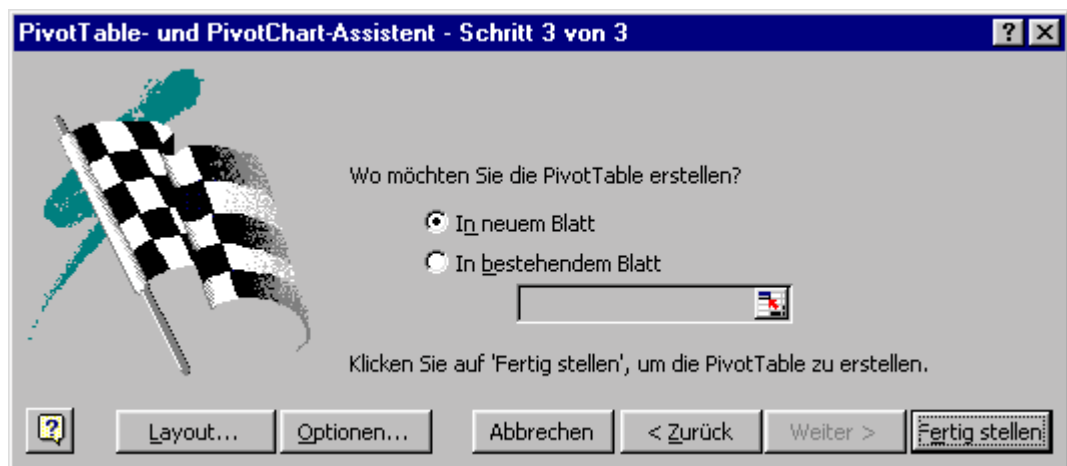
Excel markiert die Daten, die Sie vorher durch die Position des Cursors markiert haben und zeigt den Bereich im Assistenten. Falls Sie hier Abweichungen zu Ihrer gewünschten Datenmatrix finden, können Sie durch Klick auf den roten Pfeil rechts am Bereichseintrag die Markierung der Matrix verändern. Wenn die Markierung richtig ist, klicken Sie auf „Weiter“.

FB-Nr.	CompJahre	CompEin	Technik	Alter	Geschlecht	Studienfach
1	2	1	1	19	1	Geisteswissens
2	1	2	1	20	1	Pädagogik
7	6	2	3	53	1	Pädagogik
8	8	2	4	23	1	Geisteswissens
9	10	2	2	23	1	Pädagogik
1					1	Geisteswissens
					2	Pädagogik
					2	Geisteswiss
					2	Pädagogik
					2	Geisteswiss
					2	Pädagogik
					2	Geisteswissens



5. Schritt

Der Assistent fragt in seinem letzten Schritt nach dem Ort für die Pivottabelle. Der Vorschlag „In neuem Blatt“ ist dann gut, wenn Sie übersichtlich arbeiten wollen.



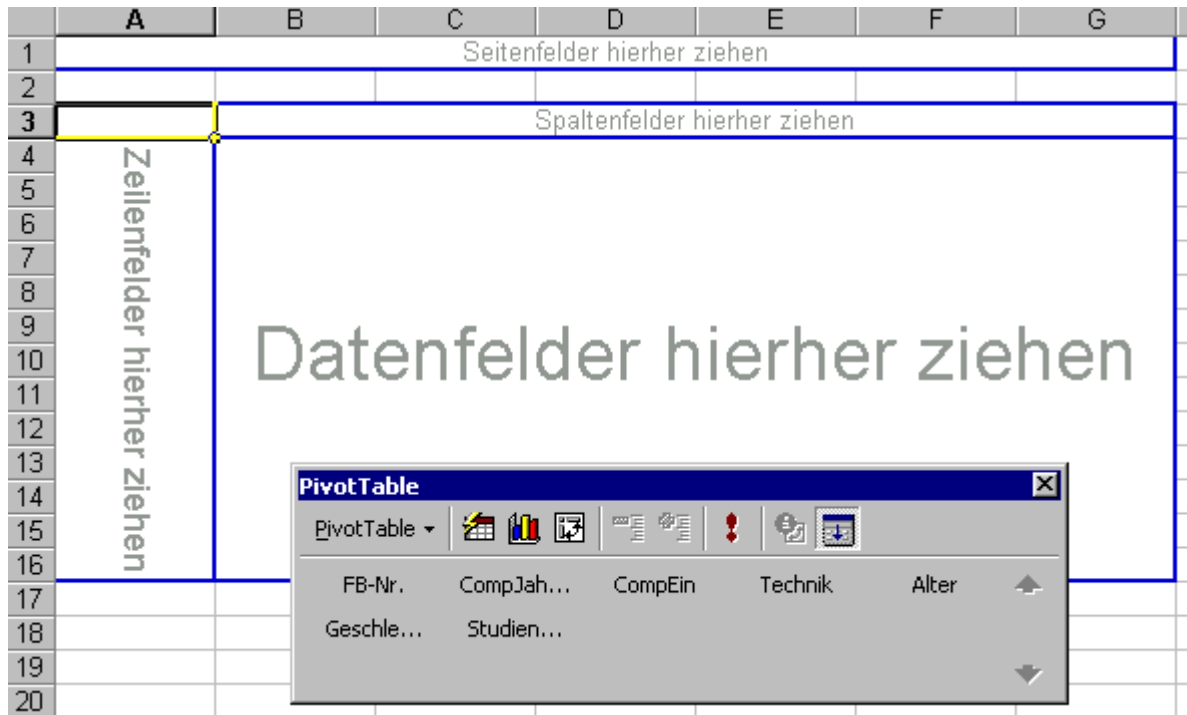
Sie können mit der Schaltfläche „Layout“ in die Ansicht für Pivottabellen wechseln, wie sie in Excel 97 eingerichtet waren. Sie finden hierfür die Beschreibung in Ihrem Skript Excel 97/2000.

In den Optionen für Pivottabellen können Sie Grundeinstellungen verändern. Wir behalten vorerst die voreingestellten Optionen bei.

Wir klicken auf „Fertig stellen“.

6. Schritt

Wir finden Eintragungen in der Excel-Tabelle und eine Symbolleiste, auf der neben einigen neuen Symbolen die Variablenamen aus der Datenmatrix eingetragen sind.



Die Aufforderungen Daten-, Zeilen-, Spalten- und Seitenfelder zu ziehen ist wörtlich zu nehmen. Wir können direkt aus der Symbolleiste die Variablen in die Pivottable ziehen. Dabei können die Variablennamen mehrfach eingesetzt werden.

Für die Auszählung des Geschlechts wird die Variable „Geschlecht“ in das Datenfeld gezogen.

A3		= Summe - Geschlecht	
	A	B	C
1	Seitenfelder hierher ziehen		
2			
3	Summe - Geschlecht	Ergebnis	
4	Ergebnis		18

Die Pivottable liegt oben links in dem neuen Tabellenblatt. Die Formel heißt „Summe Geschlecht“ und zeigt als Ergebnis 18.

Wir haben die Information, dass die Summe der Werte für Geschlecht 18 ist.

Wir ziehen den Variablennamen „Geschlecht“ in den Bereich „Spaltenfelder“.

A3		= Summe - Geschlecht		
	A	B	C	D
1	Seitenfelder hierher ziehen			
2				
3	Summe - Geschlecht	Geschlecht		
4		1	2	Gesamtergebnis
5	Ergebnis	6	12	18
6				

Die Formel heißt weiterhin „Summe – Geschlecht“ und in der Pivottabelle werden die verwendeten Werte für „Geschlecht“ aufsummiert. Da diese Kodierung keine numerischen Werte sind, sondern zwei Kategorien bezeichnen, wählen wir die Formel „Anzahl“.

Sie klicken hierfür doppelt auf die graue Fläche „Summe-Geschlecht“. Sie finden eine Liste der zur Verfügung stehenden Statistiken. Sie wählen „Anzahl-Geschlecht“.



Sie bestätigen mit „OK“. Die Häufigkeitsauszählung ist perfekt.

	A	B	C	D
1	Seitenfelder hierher ziehen			
2				
3	Anzahl - Geschlecht	Geschlecht		
4		1	2	Gesamtergebnis
5	Ergebnis	6	6	12

Sie können aus der Tabelle ablesen, dass 6 Männer (Wert = 1) und 6 Frauen (Wert = 2) und insgesamt 12 Personen in der Stichprobe sind.

7. Schritt

Falls Sie eine Häufigkeitstabelle für die Variable „Geschlecht“ wollen, ziehen Sie die Variable „Fragebogennummer“ in den Zeilenbereich. Sie erhalten eine vollständige Häufigkeitstabelle.

A3		= Anzahl - Geschlecht		
	A	B	C	D
1	Seitenfelder hierher ziehen			
2				
3	Anzahl - Geschlecht	Geschlecht		
4	FB-Nr.	1	2	Gesamtergebnis
5	1	1		1
6	2	1		1
7	3		1	1
8	4		1	1
9	5		1	1
10	6		1	1
11	7	1		1
12	8	1		1
13	9	1		1
14	10	1		1
15	11		1	1
16	12		1	1
17	Gesamtergebnis	6	6	12

8. Schritt – Berechnen der prozentualen Anteile

Zum Berechnen der prozentualen Anteile ziehen wir vorerst das Feld „FB-Nr.“ wieder aus der Pivottabelle heraus. Es wird gelöscht und die kleine Tabelle ist wieder hergestellt.

	A	B	C	D
1				
2				
3	Anzahl - Geschlecht	Geschlecht		
4		1	2	Gesamtergebnis
5	Ergebnis	6	6	12
6				

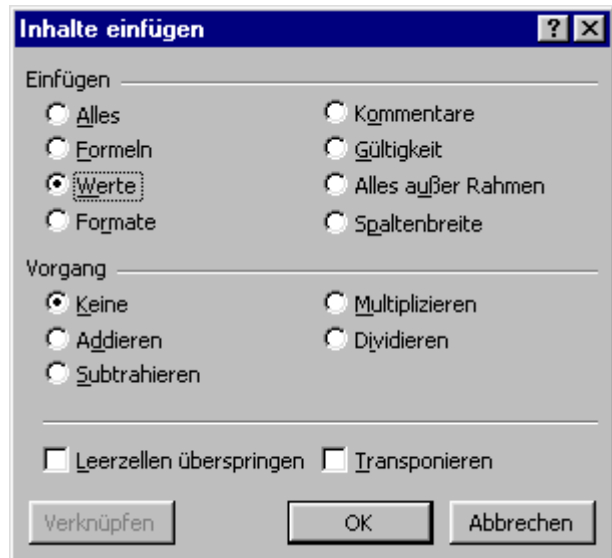
Zum Berechnen der prozentualen Anteile gibt es zwei Wege.

- Sie können direkt in der Tabelle rechnen.
- Wollen Sie jedoch die grauen Felder nicht weiter verwenden, markieren Sie die Tabelle und kopieren sie in den Zwischenspeicher



- Danach wechseln Sie zu einem neuen Tabellenblatt und verwenden Sie den Befehl

Bearbeiten > Inhalte einfügen > Werte



Danach ist die Pivottabelle eine ganz normale Excel-Tabelle geworden. Sie können hier nunmehr Prozentwerte beliebig eintragen und die Tabelle gestalten, um die Tabelle danach in Word weiter zu verwenden.

	Männlich	Weiblich	Gesamt
Häufigkeit	6	6	12
Prozent	50%	50%	100%

Die Formel zur Berechnung der Prozentwerte in der Beispieltabelle:

	A	B	C	D
1		Männlich	Weiblich	Gesamt
2	Häufigkeit	6	6	12
3	Prozent	=B2/\$D\$2	=C2/\$D\$2	=D2/\$D\$2

Eine Matrix transponieren

Falls Sie die Tabelle transponieren (Zeilen und Spalten austauschen) wollen, markieren Sie die Tabelle und kopieren sie in den Zwischenspeicher.

Sie markieren die Zielzelle, in deren Bereich die transponierte Matrix eingefügt werden soll. Sie wählen

Bearbeiten > Inhalte einfügen > Transponieren.

Sie bestätigen mit „OK“.

	A	B	C	D
1		Männlich	Weiblich	Gesamt
2	Häufigkeit	6	6	12
3	Prozent	50%	50%	100%
4				
5		Häufigkeit	Prozent	
6	Männlich	6	50%	
7	Weiblich	6	50%	
8	Gesamt	12	100%	
9				

9. Schritt – Die Tabelle in Word übernehmen

Für Ihren Bericht in Word markieren Sie die Excel-Tabelle und kopieren sie in den Zwischenspeicher.



Sie wechseln zu Word und wählen in Word „Einfügen“



Die Tabelle ist in Word eingefügt.

	Häufigkeit	Prozent
Männlich	6	50%
Weiblich	6	50%
Gesamt	12	100%

Sie können die Tabelle mit dem Befehl

Tabelle > Tabelle Autoformat

einfach und schnell formatieren.

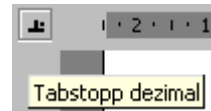


Zahlen in einer Word-Tabelle mit dem Dezimaltabulator ausrichten

Um die Zahlen schöner auszurichten, wählen Sie den Dezimaltabulator.

	Häufigkeit	Prozent
Männlich	6	50%
Weiblich	6	50%
Gesamt	12	100%

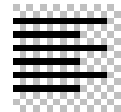
Sie wählen den Dezimaltabulator im Lineal.



Sie markieren die Spalten mit den Zahlen.

	Häufigkeit	Prozent
Männlich	6	50%
Weiblich	6	50%

Sie klicken auf „Linksbündig“.



Sie klicken im Lineal oberhalb der ersten markierten Spalte an die Stelle, an der die Zahlen ausgerichtet werden sollen.

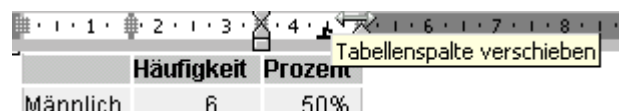


Der Dezimaltabulator steht jetzt über jeder Tabellenspalte im Lineal. Sie sehen ihn, wenn Sie in die jeweilige Tabellenspalte klicken.

Sie können die Position der Zahlen ändern, wenn Sie die Spalte markieren und den Dezimaltabulator verschieben.



Sie können die Spaltenbreiten in der Tabelle ändern, wenn Sie in die Spalte klicken und im Lineal an den Tabellenbegrenzungen mit dem waagerechten Doppelpfeil der Maus ziehen.



Ein weiteres Beispiel für eine Pivottabelle

Berechnung von Mittelwert und Standardabweichung einer Variablen insgesamt und der Geschlechter.

Wir wollen die Mittelwerte und die Standardabweichung der Variablen „Einschätzung als Computernutzer“ für die Geschlechter und insgesamt ermitteln.

1. Schritt

Sie löschen die Felder der vorherigen Pivottabelle. Sie ziehen hierfür die grauen Felder einfach aus dem Bereich der Pivottabelle.

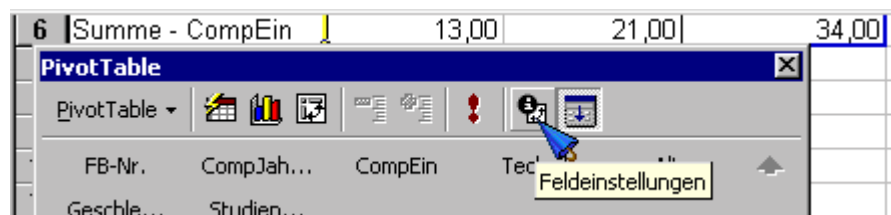
Anzahl - Geschlecht	Geschlecht		
	1	2	Gesamtergebnis
Ergebnis	6,00	6,00	12,00

2. Schritt – Berechnen des Mittelwertes und der Standardabweichung

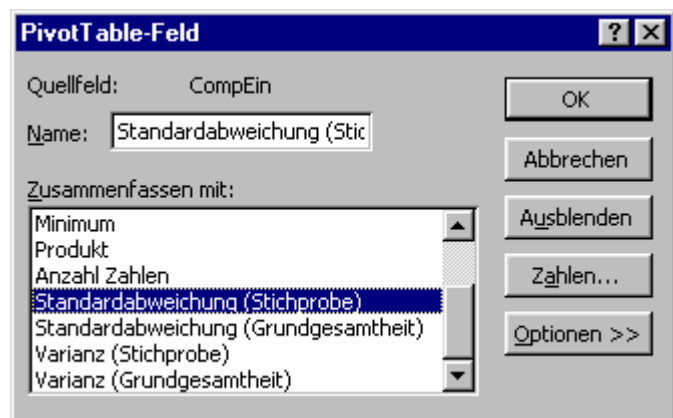
Sie ziehen die Variable „Einschätzung als Computernutzer“, das Feld „CompEin“ in das Datenfeld. Die Variable „Geschlecht“ ziehen Sie in das Spaltenfeld.

Als Formel wählen Sie für „CompEin“ Mittelwert. Sie klicken doppelt auf das Feld „Summe – CompEin“ und wählen „Mittelwert“.

Für die Standardabweichung ziehen Sie das Feld „CompEin“ nochmals in den Datenbereich. Excel berechnet wieder automatisch die Summe. Da Sie jetzt nicht mehr Doppelklicken können, markieren Sie das Feld „Summe – CompEin“ und wählen in der Symbolleiste das Symbol „Feldeinstellungen“.



Sie können hier die Standardabweichung auswählen. Sie bestätigen mit „OK“.



3. Schritt – Ergebnis

Die Pivottabelle ist mit der Berechnung des Mittelwertes und der Standardabweichung der Variablen „Einschätzung als Computernutzer“ fertig.

	A	B	C	D
3		Geschlecht ▾		
4	Daten ▾	1	2	Gesamtergebnis
5	Mittelwert - CompEin	2,17	3,50	2,83
6	Standardabweichung (Stichprobe) - CompEin	0,98	0,84	1,11

4. Schritt – Formatieren der Tabelle und Übertragen in Word

Zum Formatieren der Tabelle und Übertragen in Word gehen Sie vor, wie oben beschrieben.

Korrelation

Eine Zusammenhangsrechnung ist in Excel u. a. mit dem Pearson'schen Korrelationskoeffizienten möglich. Dieser Koeffizient ist bewährt und wird, vor allem in kleineren Stichproben bis zu 100 bis 1.000 Personen gerne eingesetzt. Der Koeffizient wird auf der Basis des künstlich angenommenen metrischen Datenniveaus für die Antworten auf einer Ratingskala eingesetzt. Die Antworten auf einer Ratingskala sind jedoch auf Ordinalniveau, weil die Einschätzung von Personen keine genauen Abstände auf der Skala darstellen.

Die Formel in Excel für den Pearson'schen Korrelationskoeffizienten heißt:

$$= \text{Pearson}(\text{Matrix1}; \text{Matrix2})$$

Die Argumente „Matrix1“ bzw. „Matrix2“ sind die Zahlenbereiche der beiden Variablen, die wir korrelieren wollen.

Wir korrelieren die Variablen „Dauer der Computernutzung“ und „Selbsteinschätzung als Computernutzer“.

	A	B	C	D
1	FB-Nr.	CompJahre	CompEin	Tecl
2	1	2	1	
3	2	1	2	
4	7	6	2	
5	8	8	2	
6	9	10	2	
7	10	2	4	
8	3	0	3	
9	4	5	5	
10	5	2	4	
11	6	4	3	
12	11	3	3	
13	12	1	3	
14				
15		=PEARSON(B2:B13;C2:C13)		

Das Ergebnis der Korrelation heißt in der Regel „r“.

Im Beispiel ist $r=-0,20$ für den Zusammenhang zwischen der Selbsteinschätzung als Computernutzer und der Dauer der Computernutzung. Diese Korrelation ist gegenläufig und sehr gering.

Die Korrelation zwischen den Variablen „Selbsteinschätzung als Computernutzer“ und „Umgang mit Technik“ ist mit $r=0,63$ wesentlich höher.

5. Berechnung des Bestimmtheitsmaßes oder der aufgeklärten Varianz

Um aus einer Korrelation eine genauere Vorhersage schätzen zu können, kann das Bestimmtheitsmaß errechnet werden. De facto muss die Stichprobengröße hinreichend sein, um eine Vorhersage zu ermöglichen. Die vorliegende Stichprobe mit 12 Personen ist viel zu klein, um eine Vorhersage zu ermöglichen. Die Zufallseffekte in den individuellen Antworten auf die Fragen sind viel zu groß.

Prinzipiell kann die Vorhersage aus der Korrelation (unter der Voraussetzung einer Stichprobe mit mehr als 100 Personen) aus dem Quadrat des Korrelationskoeffizienten geschätzt werden.

Die Formel in Excel wird dem Exponentialzeichen berechnet.

=Feld^2

=PEARSON(B2:B13;C2:C13)	=PEARSON(C2:C13;D2:D13)
=B15^2	=C15^2

Nach Eingeben der Formel formatieren Sie die Zellen im Prozentformat.



Sie erhalten das Bestimmtheitsmaß r^2 oder die Varianzaufklärung als relativ einfache Schätzung der genauen Vorhersage.

Falls die Beispielstichprobe mehr als 100 Personen enthielte und die gleichen Korrelationskoeffizienten berechnet würden, würde das Bestimmtheitsmaß die Vorhersage ermöglichen:

-0,204672997	0,631699696
4%	40%

Für 4% der Personen trifft es zu, dass sie sich trotz längerer Dauer der Computernutzung eher als Einsteigerinnen oder Einsteiger einschätzen.

Für 40% der Personen in unserer kleinen Zufallsstichprobe trifft es zu, dass sie sich sowohl als Computernutzer eher als Fortgeschrittene einschätzen und weniger gut mit Technik umgehen können.

Anmerkungen zu Statistik mit Excel

Sie finden im Funktionsassistenten von Excel eine ganze Reihe statistischer Funktionen. Wenn Sie auf das Fragezeichen klicken, erhalten Sie weitere Hinweise zur Anwendung der Funktionen.

Wenn Sie tatsächlich größere empirische Untersuchungen haben, empfiehlt sich der Einsatz eines Statistikprogramms, z. B. SPSS. Sie können dann differenziertere Rechnungen vornehmen und genauere Rechnungen vornehmen. Z. B. stehen bivariate Korrelationen häufig unter dem Einfluss weiterer Variablen. Diese Einflüsse können Sie mit Excel unter Einsatz der genauen Formeln herausrechnen, jedoch nicht mit bereits fertigen Funktionen. Wenn Sie also fit in der Anwendung von direkten statistischen Formeln sind, können Sie jede Rechnung in einer Tabellenkalkulation vornehmen.

Bei einigen Rechnungen werden Sie in SPSS oder anderen Statistikprogrammen feststellen, dass es diese gar nicht gibt. In dem Fall sind Sie auf die Tabellenkalkulation und die hohe Flexibilität des Programmes angewiesen. Sie können hier einfach Ihre Tabellen mit den entsprechenden Berechnungen aufbauen.

Zusätzlich ist Excel bei der Übertragung der Ergebnisse aus einem Statistikprogramm sehr hilfreich. Sie können die Tabellen in Word schöner formatieren, wenn sie vorher in Excel eingefügt werden (kopieren > einfügen). Zusätzlich sind Excel-Diagramme rasch und leicht erstellt und unterstützen die Aussagekraft Ihrer Ergebnisse in Ihrer empirischen Arbeit.