



Hochschule Osnabrück  
University of Applied Sciences

# → SPSS Seminar



Forschungsgruppe Informatik im  
Gesundheitswesen

Zentrum für Multimedia und IT-  
Anwendungen (ZeMIT)

**Jens Hüsters M.A.**

Forschungsgruppe Informatik im  
Gesundheitswesen

Tel.: 0541 / 969-7064

Fax: 0541 / 969-2971

[J.Huesers@hs-osnabrueck.de](mailto:J.Huesers@hs-osnabrueck.de)



# Ablauf des Seminars



Hochschule Osnabrück  
University of Applied Sciences

1. Einführung
2. Dateneingabe und Variablendefinition (ggf. Wiederholung des Variablenbegriffs)
3. Transformationen (Klassenbildung)
4. Transformationen (Klassenbildung)
5. Variablenberechnung (Summenbildung)
6. Fälle auswählen
7. Häufigkeitstabellen
8. Diagramme
9. Maße der zentralen Tendenz und Dispersionsmaße
10. Zusammenhangsmaße
11. Konfidenz-Intervall



**Insgesamt 8 Aufgaben**

## Seminarinformation



<b>Titel:</b>	SPSS - Erstellung von Online-Umfragen und Erstausswertung
<b>Inhalte:</b>	<ul style="list-style-type: none"><li>• Überblick über den Funktionsumfang von SPSS</li><li>• Definieren von Variablen</li><li>• Import / Erfassen von Daten</li><li>• Funktionsumfang deskriptive Statistik – Berechnen von Kennzahlen</li><li>• Erstellen von einfachen Tabellen</li><li>• Erstellen von einfachen Diagrammen</li></ul>
<b>Vorkenntnisse:</b>	Vorkenntnisse in der Benutzung von PCs und Windows sowie Grundkenntnisse deskriptiver Statistik erforderlich. (Es werden keine Statistikkenntnisse vermittelt.)
<b>Veranstalter:</b>	wiconnect e.V. - Alumniverein der Fakultät WiSo Rückfragen an <a href="mailto:alumni@wi.hs-osnabrueck.de">alumni@wi.hs-osnabrueck.de</a>

## ➔ Warum SPSS?

- SPSS gibt es seit über 40 Jahren.
- SPSS ist eins der am häufigsten genutzten Datenanalysesysteme.
- Vergleichsweise intuitiv, Alternativen sind STATA, R oder SAS.
- Wird angewendet in den meisten Institutionen, die sich mit der Auswertung und Darstellung von Daten befassen wie z.B.:

➔ **Universitäten, Hochschulen, allg. Forschungseinrichtungen z.B.:** Soziologie, Politikwissenschaft, Psychologie, Betriebswirtschaftslehre, Medizin, Geographie, Geschichtswissenschaften, Pädagogik, ...

➔ **Private Wirtschaftsunternehmen z.B.:** AUDI, Henkel, Otto-Versand, ...

➔ **Öffentlichen Verwaltungen z.B.:** Bundesagentur für Arbeit, Bundeskriminalamt, ...

# Einführung

## ➔ Anwendung Forschungsprozesses



Hochschule Osnabrück  
University of Applied Sciences

**Problemformulierung, Problembenennung,  
Theorie- und Hypothesenbildung**

### **Konzeptualisierung**

- Operationalisierungsvorgang
- **Konstruktion des Erhebungsinstruments**
- Festlegung der Stichprobe
- Festlegung des Forschungsdesigns
- Pretest

Limesurvey, bzw. andere  
Tools zur Erstellung von  
Onlinefragebögen (z.B.  
Unipark).

**Datenerhebung, -aufbereitung, -analyse**  
**Interpretation und Publikation**

SPSS, bzw. andere  
Datenanalyseprogramme  
(z.B. STATA).

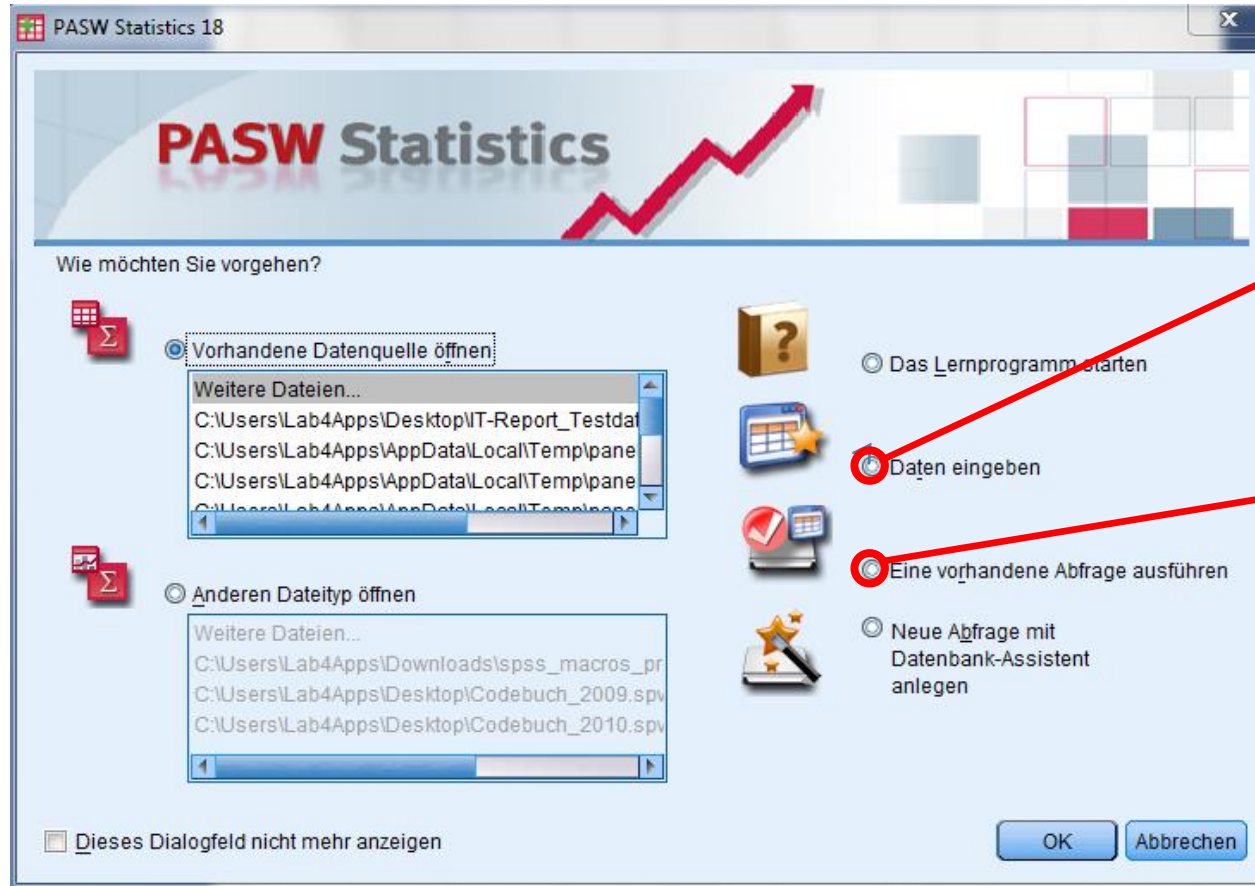
# Grundlagen

## ➔ Öffnen des Programms

➔ Programme ➔ SPSS Statistics



Hochschule Osnabrück  
University of Applied Sciences



Variablendefinition:  
Variablen-Labels,  
Werte-Labels,  
Fehlende Werte

Öffnen eines  
bestehenden  
Datensatzes aus (bspw.  
aus Limesurvey, etc.)

# Grundlagen

## ➔ Der Dateneditor – Variablen und Datenansicht

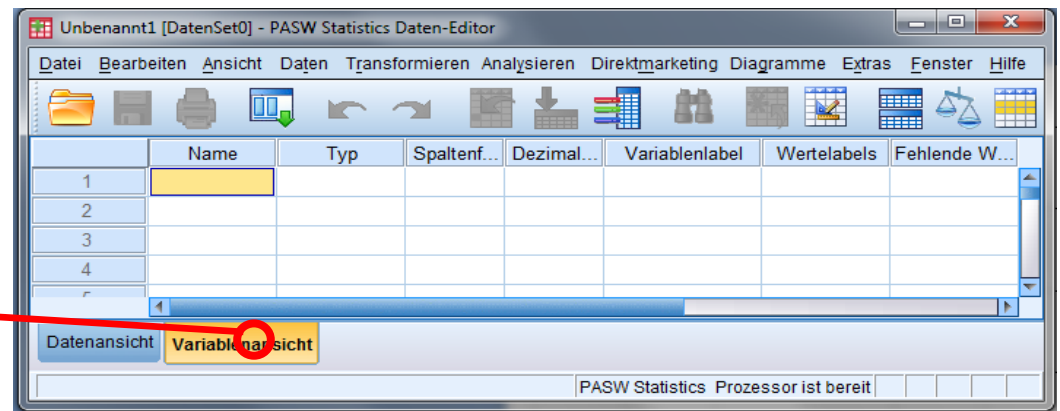
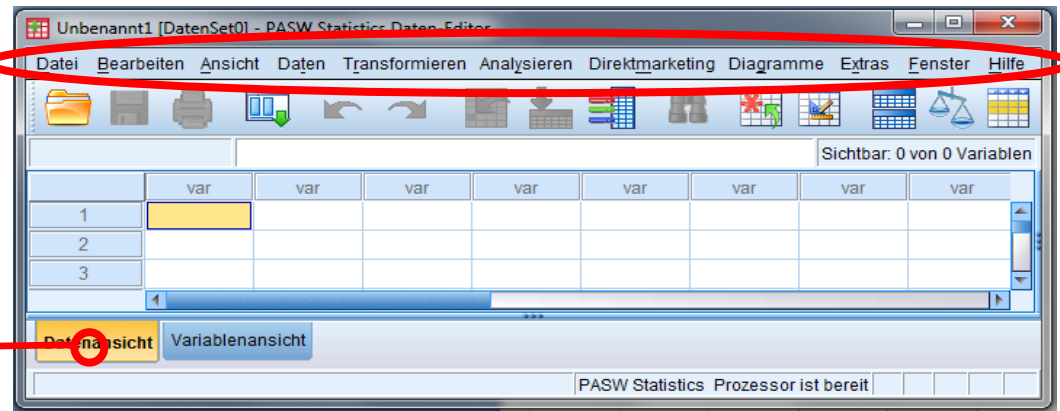


Hochschule Osnabrück  
University of Applied Sciences

Die **Menüpunkte** zum Bearbeiten der Variablen, Daten und zur Erstellung von Statistiken, sind im Dateneditor, Viewer und Syntax-Editor verfügbar!

**Datenansicht:** Darstellung des Datensatzes mit einzelnen Fällen und Variablen.

**Variablenansicht:** Darstellung und Bearbeitungsmenü von Variablenattributen wie Datentyp (Skalenniveau), Variablen- und Wertelabels.



# Grundlagen

## ➔ Der Viewer – Ergebnisausgabe



Hochschule Osnabrück  
University of Applied Sciences

**Zusätzliche Menüpunkte**

**Häufigkeiten**

**Tabellenform der Darstellung mit Ein- und Ausblendfunktionen**

Statistiken

		Zufriedenheit	Ort	Postleitzahl	Alter
N	Gültig	10	10	10	10
	Fehlend	0	0	0	0

- Der Viewer erscheint immer, sobald Sie Variablen umkodieren, den Datensatz speichern oder sich Ergebnisse ausgeben lassen.
- In vielen Fällen können Sie den Viewer ohne abzuspeichern schließen.
- Für die Dokumentation von Variablentransformationen (Umkodierung, etc.) und für die Ergebnisdokumentation sollten Sie den Viewer mit gut nachvollziehbaren Notationen speichern.



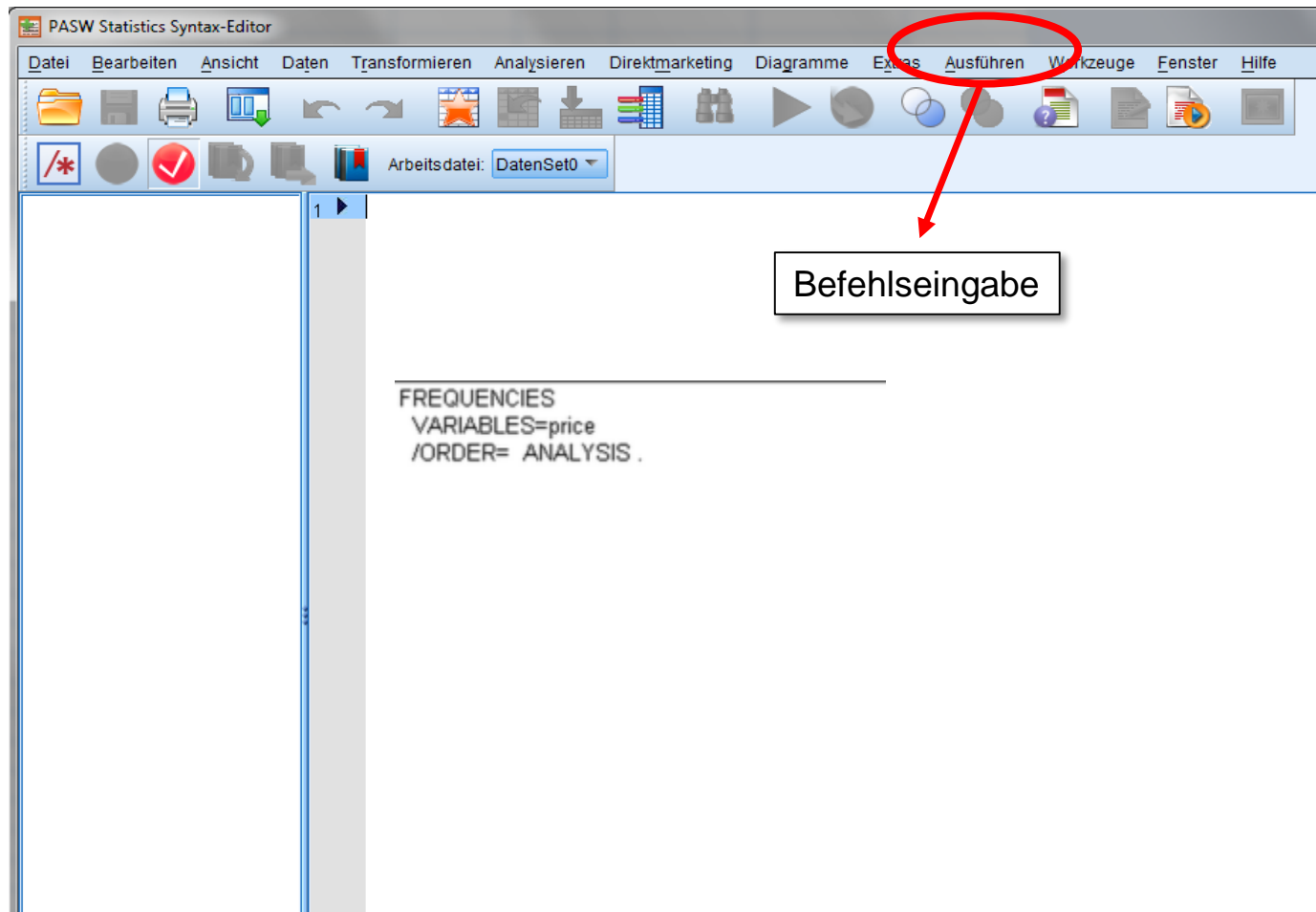
# Grundlagen

## ➔ Syntax

Öffne über: → Datei → Neu → Syntax



Hochschule Osnabrück  
University of Applied Sciences



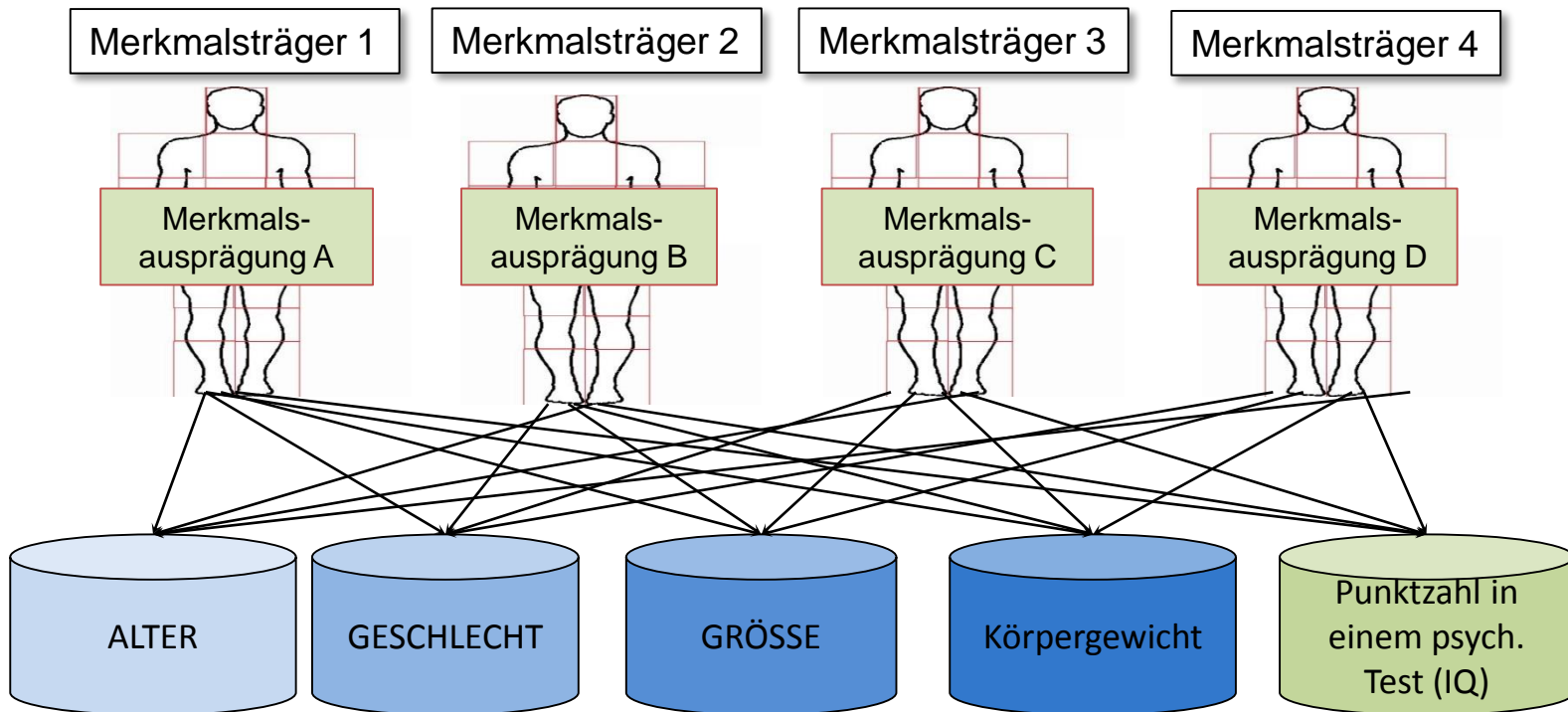


# Was sind Variablen und welche Unterscheidungsmöglichkeiten gibt es?



➔ „Ein Variable bezeichnet ein Merkmal oder eine Eigenschaft“

1. **Merkmalsträgern** → bspw. der Kunde, Krankenhaus
2. **Variablen** (bzw. Merkmal oder Merkmalsdimensionen) → bspw. das Alter (des Kunden)
3. **Ausprägungen** von Variablen (bzw. Kategorien, Merkmalsausprägungen) → 5 Jahre



# Grundsätzlich...



Hochschule Osnabrück  
University of Applied Sciences

*...das Beherrschen der Variablendefinition  
und der Dateneingabe ist die Basis für alle  
Arbeiten mit SPSS!!!*

# Variablendefinition und Dateneingabe



Name

Jede Variable benötigt einen eindeutigen Namen

SPSS vergibt automatisch Namen und ergänzt mit fortlaufender, fünfstelliger Nummer (VAR00001, VAR00002, ...)

## Regeln:

- Maximal 64 Zeichen
- Beginnen mit A-Z, @, #, \$
- Dürfen nicht mit Leerzeichen oder \_ enden
- #\_xxx sind Hilfsvariablen
- \$\_xxx sind Systemvariablen



Hochschule Osnabrück  
University of Applied Sciences

	Name	Typ	Spalten...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W...	Spalten	Ausrichtung	Messniveau
1	V_1_1	String	50	0	Ort	Keine	999	16	Links	Nominal
2	V_1_2	Numerisch	8	0	Postleitzahl	Keine	999	8	Rechts	Nominal
3	V_1_3	Numerisch	8	0	Alter	Keine	999	8	Rechts	Skala
4	V_2_1	Numerisch	8	0	Zufriedenheit	{1, Sehr zufrieden}...	999	8	Rechts	Ordinal
5										
6										
7										
8										
9										
10										
11										

	V_1_1	V_1_2	V_1_3	V_2_1	var	var	var	var	var	var
1	Osnabrück	49076	22	Zufrieden						
2	Hameln	31789	24	Zufrieden						
3	Rinteln	31737	19	Sehr unzuf...						
4	Blomberg	32825	19	neutral						
5	Hannover	30453	27	Unzufrieden						
6	Berlin	12524	23	Sehr zufrie...						
7	Osnabrück	49086	22	Sehr zufrie...						
8	Osnabrück	49090	28	Zufrieden						
9	Osnabrück	49090	30	Zufrieden						
10	Minden	42423	25	Sehr zufrie...						

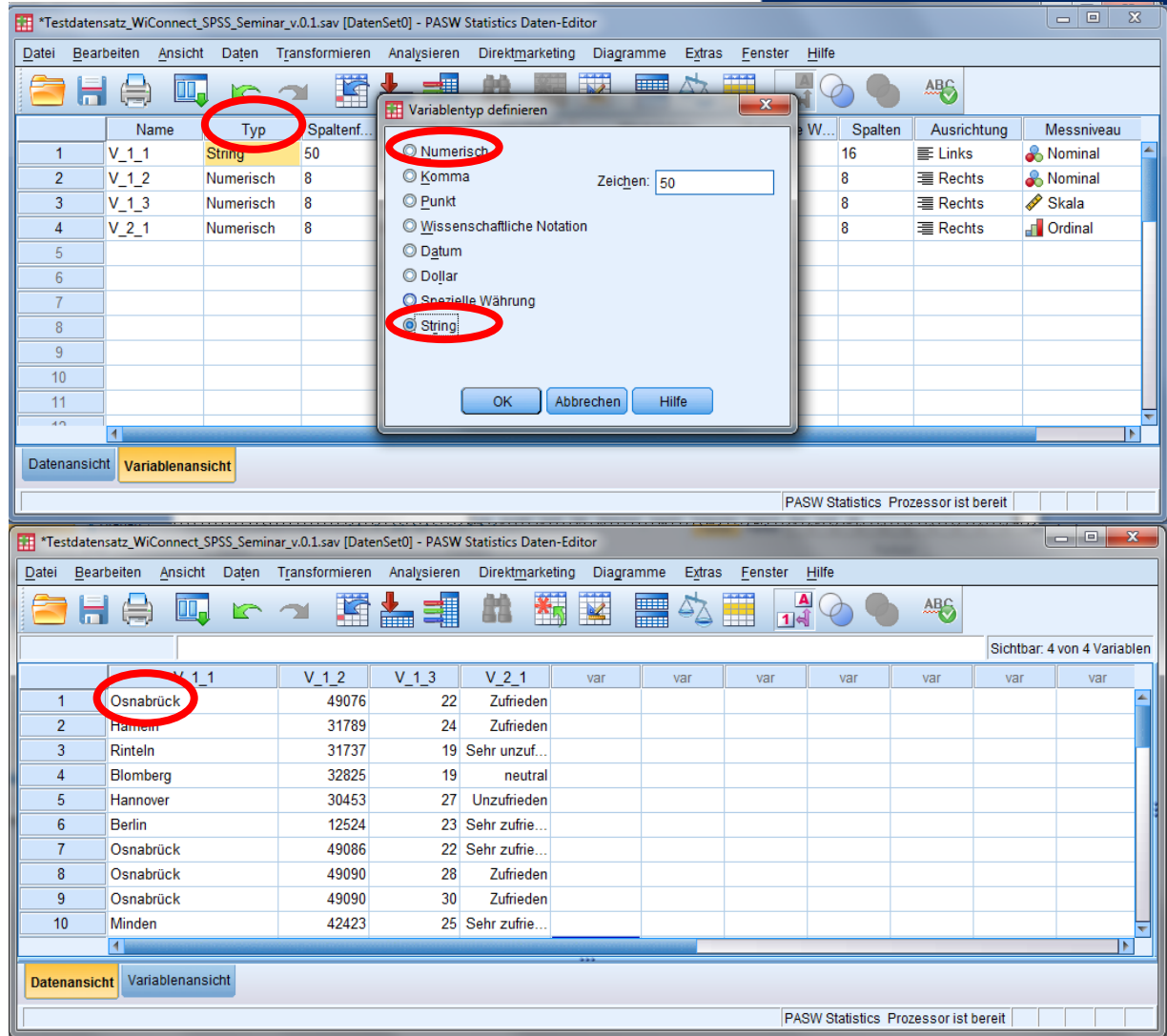
# Variablendefinition und Dateneingabe



Komma, Punkt, wiss. Notation, Datum, Dollar und spezielle Währung sind Varianten numerischer Variablen.

**Typ String:** immer dann wenn es sich um eine Texteingabe nicht standardisierter Daten handelt (z.B. PLZ, Telefonnummer) Anzahl der Zeichen festlegen

„Numerisch“ und „String“ sind die am häufigsten vorkommenden Typen.



The top screenshot shows the 'Variablentyp definieren' (Define Variable Type) dialog box in SPSS. The 'String' option is selected, and the 'Zeichen' (Characters) field is set to 50. The background shows a data editor with columns V\_1\_1, V\_1\_2, V\_1\_3, and V\_2\_1.

The bottom screenshot shows the 'Datenansicht' (Data View) of the same dataset. The first column, V\_1\_1, contains the text 'Osnabrück' in the first row, which is circled in red. The other columns contain numerical values and categorical labels.

	V_1_1	V_1_2	V_1_3	V_2_1
1	Osnabrück	49076	22	Zufrieden
2	Hannover	31789	24	Zufrieden
3	Rinteln	31737	19	Sehr unzuf...
4	Blomberg	32825	19	neutral
5	Hannover	30453	27	Unzufrieden
6	Berlin	12524	23	Sehr zufrie...
7	Osnabrück	49086	22	Sehr zufrie...
8	Osnabrück	49090	28	Zufrieden
9	Osnabrück	49090	30	Zufrieden
10	Minden	42423	25	Sehr zufrie...

# Variablendefinition und Dateneingabe

## ➔ Variablen- und Wertelabel



Hochschule Osnabrück  
University of Applied Sciences

Ein **Variablenlabel** beschreibt die inhaltliche Bedeutung der Variable (max. 255 Zeichen) und kommt insb. in der Ergebnisausgabe als Tabellen- bzw. Abbildungsüberschrift (im Viewer) zur Geltung.

**Wertelabels:** „Etikettierung“ der Merkmalsausprägung nominal oder ordinal skaliert Werte.

1 = JA;  
2 = NEIN

1 = zufrieden  
2 = neutral  
3 = unzufrieden

Button zum Wechseln von Werten und Label in der Datenansicht

The screenshot shows the PASW Statistics Daten-Editor interface. The top window displays the 'Variablenlabel' and 'Wertelabels' columns in the variable list, which are circled in red. The bottom window shows the 'Datenansicht' (Data View) with a table of data. A red arrow points from the 'Datenansicht' button in the bottom left to the '1' icon in the toolbar, which is also circled in red.

Name	Typ	Spaltenf...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W...	Spalten	Ausrichtung	Messniveau
V_1_1	String	50	0	Ort	Keine	999	16	Links	Nominal
V_1_2	Numerisch	8	0	Postleitzahl	Keine	999	8	Rechts	Nominal
V_1_3	Numerisch	8	0	Alter	Keine	999	8	Rechts	Skala
V_2_1	Numerisch	8	0	Zufriedenheit	(1, Sehr zufrieden)...	999	8	Rechts	Ordinal

	V_1_1	V_1_2	V_1_3	V_2_1	var	var	var	var	var	var
1	Osnabrück	49076	22	Zufrieden						
2	Hameln	31789	21	Zufrieden						
3	Rinteln	31737	19	Sehr unzuf...						
4	Blomberg	32825	19	neutral						
5	Hannover	30453	27	Unzufrieden						
6	Berlin	12524	23	Sehr zufrie...						
7	Osnabrück	49086	22	Sehr zufrie...						
8	Osnabrück	49090	28	Zufrieden						
9	Osnabrück	49090	30	Zufrieden						
10	Minden	42423	25	Sehr zufrie...						

# Variablendefinition und Dateneingabe

## ➔ Fehlende Werte

Falls ein Proband zu einer Frage **keine Angaben** macht oder die Angabe als bei einer Prüfung als **ungültig** identifiziert wurde, muss SPSS mit dem fehlenden Wert in der entsprechende Variable „umgehen“ können.

SPSS kennt systemdefinierte fehlende Werte (system missing values) und benutzerdefinierte Fehlende Werte (user missing values).

**User Missing Value:** Nehmen Sie einen Wert, der in der Empirie als Merkmalsausprägung nicht vorkommt, bspw. -999, -777, ...

**Wertelabels:** Für fehlende Werte können auch Wertelabel vergeben werden.

Bspw.: -999 = ungültig/fehlend



Hochschule Osnabrück  
University of Applied Sciences

\*Testdatensatz\_WiConnect\_SPSS\_Seminar\_v.0.1.sav [DatenSet0] - PASW Statistics Daten-Editor

	Name	Typ	Spalten...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W	Spalten	Ausrichtung	Messniveau
1	V_1_1	String	50	0	Ort	Keine	999	16	Links	Nominal
2	V_1_2	Numerisch	8	0	Postleitzahl	Keine	999	8	Rechts	Nominal
3	V_1_3	Numerisch	8	0	Alter	Keine	999	8	Rechts	Skala
4	V_2_1	Numerisch	8	0	Zufriedenheit	{1, Sehr zufrieden}...	999	8	Rechts	Ordinal
5										
6										
7										
8										
9										
10										
11										

Datenansicht Variablenansicht

PASW Statistics Prozessor ist bereit

\*Testdatensatz\_WiConnect\_SPSS\_Seminar\_v.0.1.sav [DatenSet0] - PASW Statistics Daten-Editor

	V_1_1	V_1_2	V_1_3	V_2_1	var	var	var	var	var	var
1	Osnabrück	49076	22	Zufrieden						
2	Hameln	31789	24	Zufrieden						
3	Rinteln	31737	19	Sehr unzuf...						
4	Blomberg	32825	19	neutral						
5	Hannover	30453	27	Unzufrieden						
6	Berlin	12524	23	Sehr zufriede...						
7	Osnabrück	49086	22	Sehr zufriede...						
8	Osnabrück	49090	28	Zufrieden						
9	Osnabrück	49090	30	Zufrieden						
10	Minden	42423	25	Sehr zufriede...						

Datenansicht Variablenansicht

PASW Statistics Prozessor ist bereit

# Variablendefinition und Dateneingabe

## ➔ Fehlende Werte



Hochschule Osnabrück  
University of Applied Sciences

### Statistiken

Geschlecht

N	Gültig	13
	Fehlend	2

### Geschlecht

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
➔ Gültig	Weiblich	8	53,3	61,5	61,5
	Männlich	5	33,3	38,5	100,0
	Gesamt	13	86,7	100,0	
Fehlend	-999	2	13,3		
Gesamt		15	100,0		



# Variablendefinition und Dateneingabe

## ➔ Messniveau (reine Signalfunktion)



Hochschule Osnabrück  
University of Applied Sciences

**Nominal:** Gleich / Ungleich.

- männlich = 1, weiblich = 2
- rot = 22; gelb = 28, blau = 1212
- Geschlecht, Familienstand, etc.

**Ordinal:** Gleich / Ungleich und Ausdruck einer Rangfolge aber keine Äquidistanz zwischen den Werten.

- Schulnoten
- Wettkampfplatzierungen
- Subjektive Einschätzungen aller Art

**Metrisch:** Gleich / Ungleich und Ausdruck einer Rangfolge und Äquidistanz zwischen den Werten.

- Längen- und Gewichtsmaße,
- Geldeinheiten
- Zeitskalen

	Name	Typ	Spaltenf...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W...	Spalten	Ausrichtung	Messniveau
1	V_1_1	String	50	0	Ort	Keine	999	16	Links	Nominal
2	V_1_2	Numerisch	8	0	Postleitzahl	Keine	999	8	Rechts	Nominal
3	V_1_3	Numerisch	8	0	Alter	Keine	999	8	Rechts	Skala
4	V_2_1	Numerisch	8	0	Zufriedenheit	{1, Sehr zufrieden}...	999	8	Rechts	Ordinal
5										
6										
7										
8										
9										
10										
11										

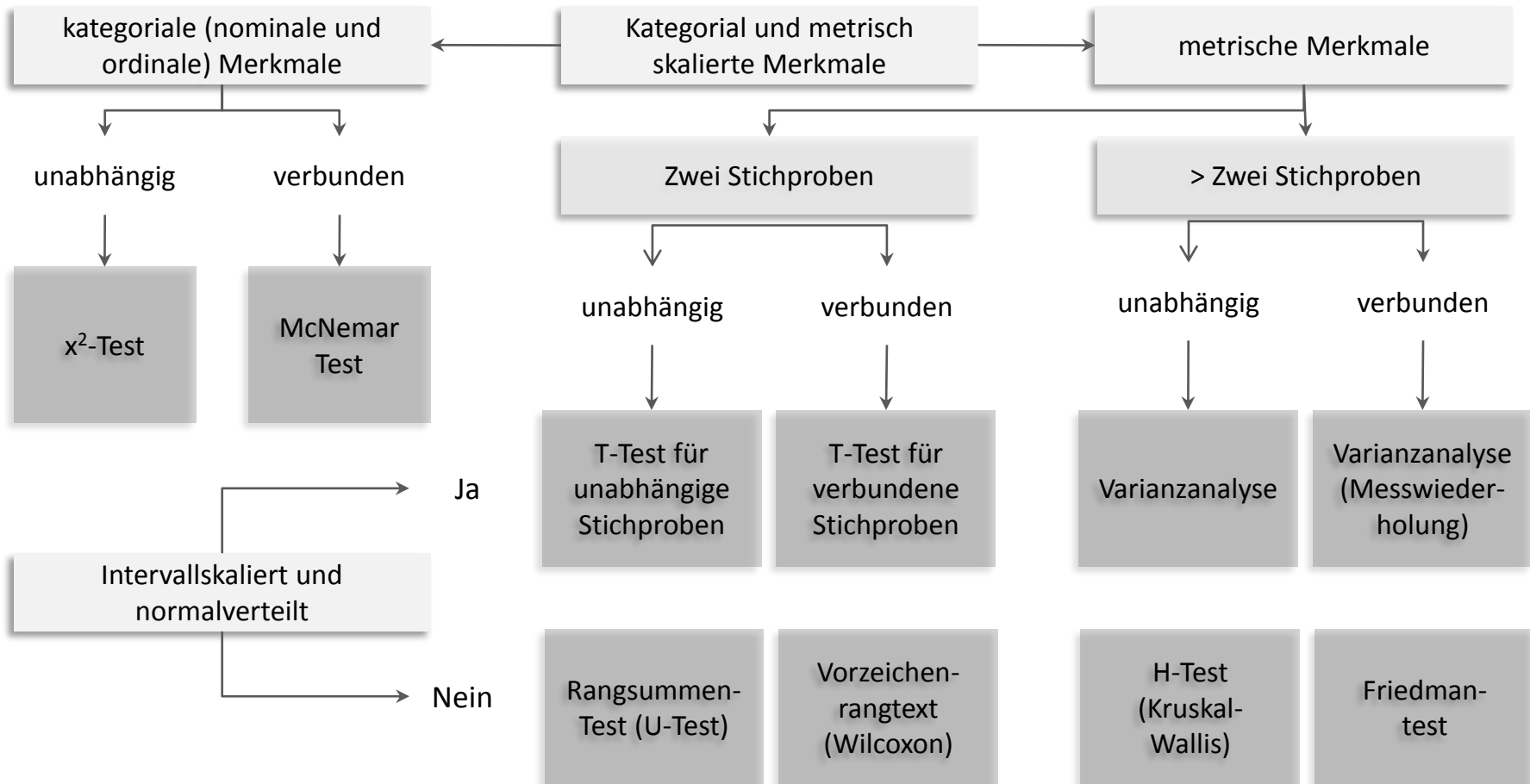
	V_1_1	V_1_2	V_1_3	V_2_1	var	var	var	var	var	var	var
1	Osnabrück	49076	22	Zufrieden							
2	Hameln	31789	24	Zufrieden							
3	Rinteln	31737	19	Sehr unzuf...							
4	Blomberg	32825	19	neutral							
5	Hannover	30453	27	Unzufrieden							
6	Berlin	12524	23	Sehr zufrie...							
7	Osnabrück	49086	22	Sehr zufrie...							
8	Osnabrück	49090	28	Zufrieden							
9	Osnabrück	49090	30	Zufrieden							
10	Minden	42423	25	Sehr zufrie...							

# Variablendefinition und Dateneingabe

## ➔ Übersicht gültiger Testverfahren



Hochschule Osnabrück  
University of Applied Sciences



# Aufgabe 1 - Erstellung einer Rohwerttabelle

Fall1	
Alter:	25
Geschlecht:	Weiblich
Stadt:	Osnabrück
Lieblingsfarbe:	Grün
Punktzahl:	60

Fall 2	
Alter:	18
Geschlecht:	Männlich
Stadt:	Bielefeld
Lieblingsfarbe:	Blau
Punktzahl:	84

Fall 3	
Alter:	18
Geschlecht:	Weiblich
Stadt:	Bielefeld
Lieblingsfarbe:	Blau
Punktzahl:	56

Fall 4	
Alter:	-
Geschlecht:	Weiblich
Stadt:	Melle
Lieblingsfarbe:	Rot
Punktzahl:	35

Fall 5	
Alter:	33
Geschlecht:	Männlich
Stadt:	Bielefeld
Lieblingsfarbe:	-
Punktzahl:	75

Fall 6	
Alter:	18
Geschlecht:	Weiblich
Stadt:	Bielefeld
Lieblingsfarbe:	Gelb
Punktzahl:	0

# Variablendefinition und Dateneingabe

## ➔ Mögliche Lösung für Aufgabe 1



Hochschule Osnabrück  
University of Applied Sciences

### Variablenansicht

Name	Typ	Spaltenf...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W...	Spalten	Ausrichtung	Messniveau	Rolle
V1_1_Alter	Numerisch	8	0	Alter der Teilnehmer	Keine	-999	8	Rechts	Skala	Eingabe
V1_2_Geschlecht	Numerisch	8	0	Geschlecht der Teilnehmer	{0, Weiblich...	-999	8	Rechts	Nominal	Eingabe
V1_3_Ort	String	20	0	Wohnort der Teilnehmer	Keine	Keine	8	Links	Nominal	Eingabe
V2_1_Lieblingsfarbe	Numerisch	8	0	Lieblingsfarbe der Teilnehmer	{1, Grün}...	-999	8	Rechts	Nominal	Eingabe
V2_2_Punktzahl	Numerisch	8	0	Punktzahl der Teilnehmer	Keine	-999	8	Rechts	Skala	Eingabe

### Datensicht

V1_1_Alter	V1_2_Geschlecht	V1_3_Ort	V2_1_Lieblingsfarbe	V2_2_Punktzahl
25	Weiblich	Osnabrück		-
18	Männlich	Bielefeld	Grün	-
18	Weiblich	Bielefeld	Blau	-
-999	Weiblich	Melle	Rot	-
33	Männlich	Bielefeld	Gelb	-
18	Weiblich	Bielefeld		-
28	Weiblich	Melle		-
-999	Männlich	Münster		-
54	Männlich	Melle		-
-	-	-	-	-

# Variablendefinition und Dateneingabe



Hochschule Osnabrück  
University of Applied Sciences

## ➔ Aufgabe 2 – Marktanalyse im Kleidungshandel

Marketing\_Kleidung.sav [DataSet3] - IBM SPSS Statistics Dateneditor

Datei Bearbeiten Ansicht Daten Transformieren Analysieren Direktmarketing Grafik Extras Fenster Hilfe

	Name	Typ	Spalten...	Dezimal...	Beschriftung	Werte	Fehlend	Spalten	Ausrichtung	Maß	Rolle
1					Identifikationsnummer						
2					Telefonnummer						
3					Geschlecht (0 - männlich, 1 - weiblich)						
4					Körpergröße in cm						
5					Kleidergröße (S, M oder L)						
6					Umsatz pro Monat in Euro						
7					Körpergewicht in kg						
8					Kundenzufriedenheit mit dem Service von 1 bis 10						
9					Alter in Jahren						
10											
11											
12											
13											
14											

Datenansicht Variablenansicht

IBM SPSS Statistics -Prozessor ist bereit Unicode:OFF

➔ Bitte ergänzen Sie die Variablenbeschreibung!

# Variablendefinition und Dateneingabe

## ➔ Mögliche Lösung für Aufgabe 2



Hochschule Osnabrück  
University of Applied Sciences

Marketing\_Kleidung.sav [DataSet3] - IBM SPSS Statistics Dateneditor

	Name	Typ	Spaltenf...	Dezimal...	Beschreibung	Werte	Fehlend	Spalten	Ausrichtung	Maß	Rolle
1	Kunden_ID	Numerisch	8	0	Identifikationsnummer	Keine	-999	8	Rechts	Nominal	Eingabe
2	Telefonnummer	Zeichenfolge	16	0	Telefonnummer	Keine	-999	8	Links	Nominal	Eingabe
3	Geschlecht	Numerisch	8	0	Geschlecht (0 - männlich, 1 - weiblich)	{0, männlich}...	-999	8	Rechts	Nominal	Eingabe
4	Körpergröße	Numerisch	8	0	Körpergröße in cm	Keine	-999	8	Rechts	Skala	Eingabe
5	Kleidergröße	Numerisch	8	0	Kleidergröße (S, M oder L)	{1, S}...	-999	8	Rechts	Ordinal	Eingabe
6	Umsatz	Numerisch	8	2	Umsatz pro Monat in Euro	Keine	-999,00	8	Rechts	Skala	Eingabe
7	Gewicht	Numerisch	8	2	Körpergewicht in kg	Keine	-999,00	8	Rechts	Skala	Eingabe
8	Zufriedenheit	Numerisch	8	0	Kundenzufriedenheit mit dem Service von 1 bis 10	Keine	-999	8	Rechts	Ordinal	Eingabe
9	Alter	Numerisch	8	0	Alter in Jahren	Keine	-999	8	Rechts	Skala	Eingabe
10											
11											
12											
13											
14											

Datenansicht Variablenansicht

IBM SPSS Statistics -Prozessor ist bereit Unicode:OFF

Wertbeschriftungen

Wert:

Beschreibung:

Rechtschreibung...

Hinzufügen Ändern Entfernen

0 = "männlich"  
1 = "weiblich"

OK Abbrechen Hilfe

Wertbeschriftungen

Wert:

Beschreibung:

Rechtschreibung...

Hinzufügen Ändern Entfernen

1 = "S"  
2 = "M"  
3 = "L"

OK Abbrechen Hilfe

Variablenansicht:  
Beschreibt die  
Struktur der  
Daten

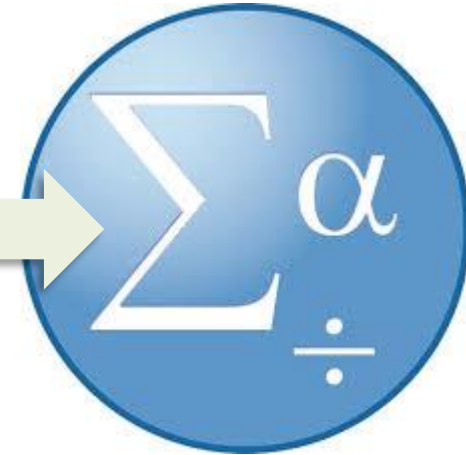
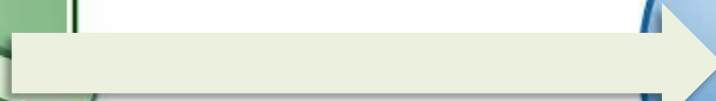
# Datenaufbereitung

➔ Dateneingabe

**[tinyurl.com/spss-bsp](http://tinyurl.com/spss-bsp)**



MS Excel



IBM SPSS



# Deskriptive Statistik

## ➔ Maße der zentralen Tendenz - Lageparameter



Hochschule Osnabrück  
University of Applied Sciences

- Geben über bestimmte Eigenschaften eines Datenkollektivs oder einer Verteilung **zusammenfassende** Auskunft
- Aus einer Vielzahl von Einzelwerten werden einige wenige Werte gebildet
- **Maße der zentralen Tendenz**
  - **Modalwert:** am häufigsten vorkommende Wert einer Verteilung. Wahrscheinlichster Wert bei zufälliger Ziehung. Grafische Darstellung über Histogramm. Lediglich Nominalskalenniveau erforderlich.
  - **Medianwert:** Wert, von dem die übrigen Werte im Durchschnitt am wenigsten abweichen. Teilt eine Verteilung in zwei Hälften. → Rangfolge bilden: bei ungerader Anzahl von Fällen der mittlere Wert. Ansonsten numerisches Mittel der beiden mittleren Zahlen. Ordinalskalenniveau erforderlich.
  - **Arithmetisches Mittel (Mittelwert):** Gebräuchlichstes Maß der zentralen Tendenz. Summe aller Werte dividiert durch die Anzahl der Fälle ( $n$ ). Summe der quadrierten Abweichungen zum Mittelwert ist minimal. Metrisches Skalenniveau erforderlich.

# Deskriptive Statistik

## ➔ Dispersionsmaße - Streuungsmaße



Hochschule Osnabrück  
University of Applied Sciences

## Dispersionsmaße

- **Variationsbreite:** Gibt Größe des Bereichs an, in dem die Messwerte liegen.

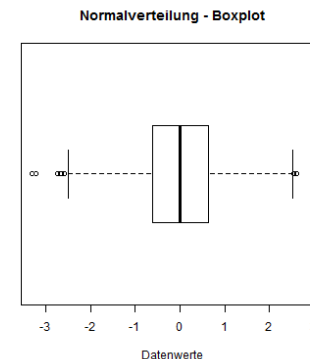
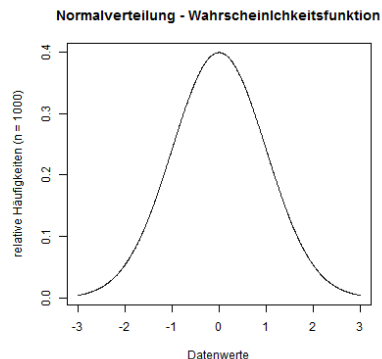
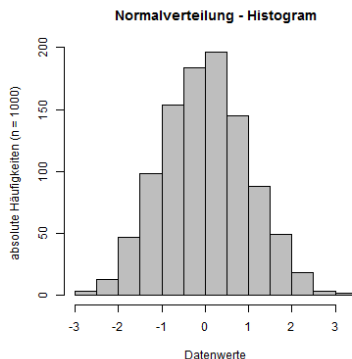
➔ Differenz zwischen Größtem und kleinsten Wert

- **Varianz:** Wichtigstes Dispersionsmaß.

➔ Summe der quadrierten Abweichungen aller Messwerte vom Arithmetischen Mittel dividiert durch die Anzahl aller Messwerte minus Eins. Metrisches Skalenniveau notwendig.

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Varianz

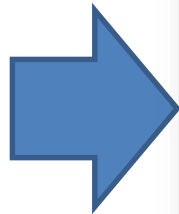


# Deskriptive Statistik

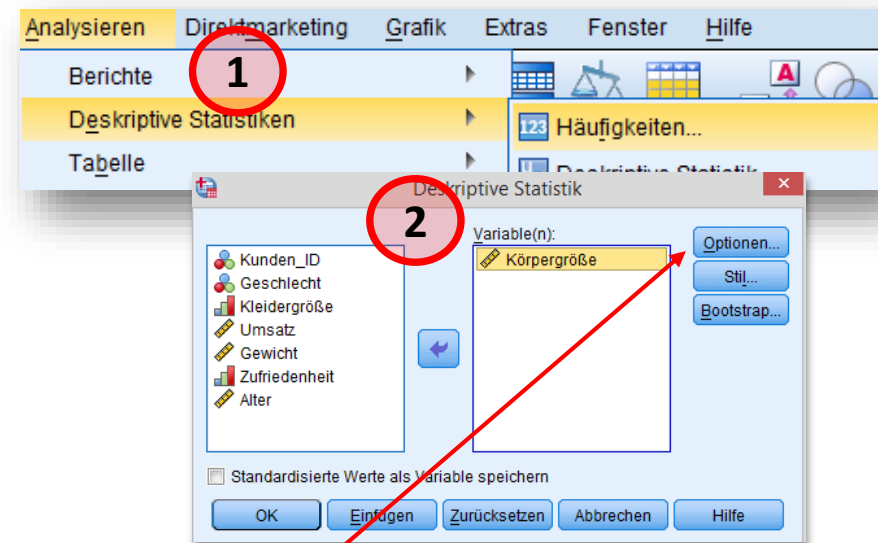
➔ Beschreibende Statistik

- Deskriptive Statistik ist ein Verfahren zur **Beschreibung von Zahlen** durch andere Zahlen
- Viele Daten werden durch wenige **Kennwerte** beschrieben

Körpergröße
184
181
187
176
171
194
180
181
188
187
188
⋮
183
185
176
187



Statistiken		
Körpergröße in cm		
N	Gültig	148
	Fehlend	2
Mittelwert		177,96
Median		178,50
Standardabweichung		8,561
Bereich		41
Minimum		157
Maximum		198

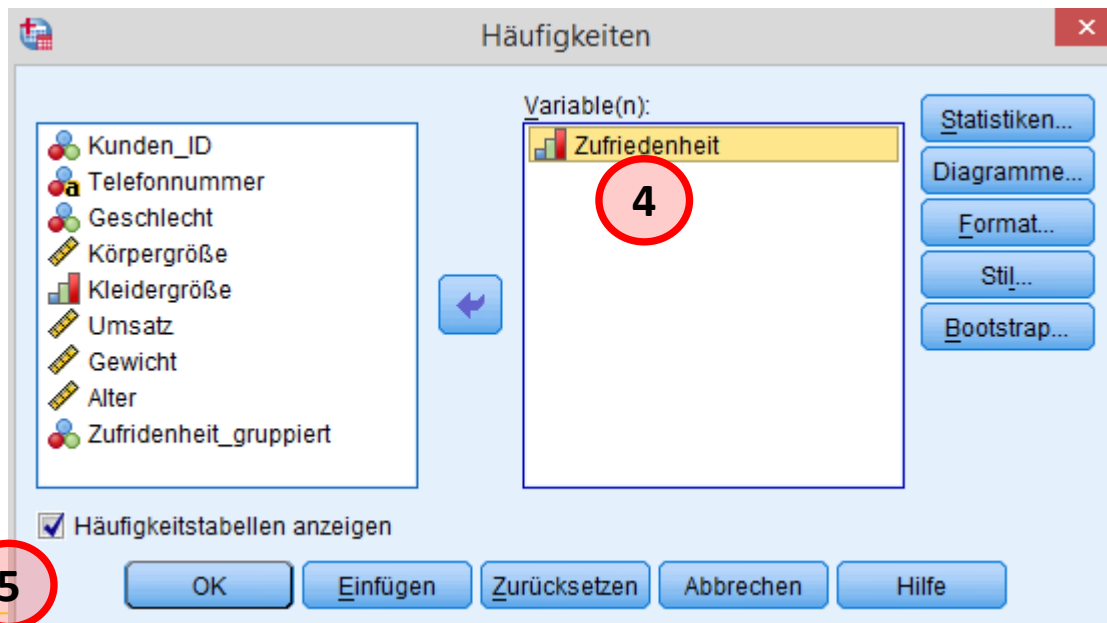
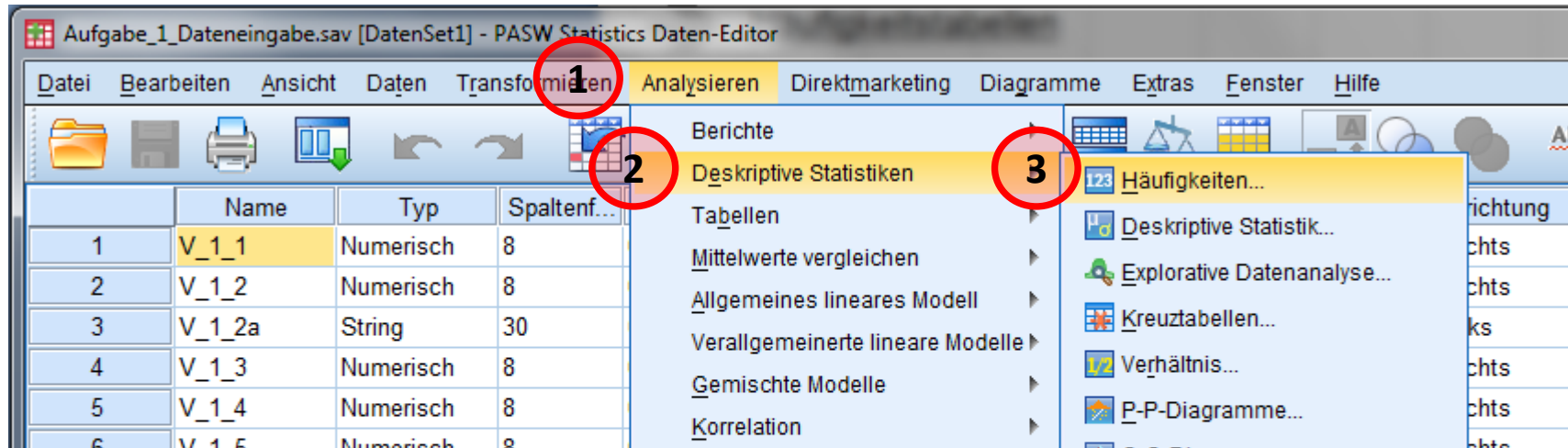


Unter Optionen können Sie verschiedene Kennwerte auswählen

**!** Tipp: Rechtsklick und die Auswahl von „deskriptiver Statistik“ gibt einen schnellen Überblick

# Deskriptive Statistik

➔ Berechnung der Häufigkeiten



# Deskriptive Statistik

➔ Berechnung von Häufigkeiten

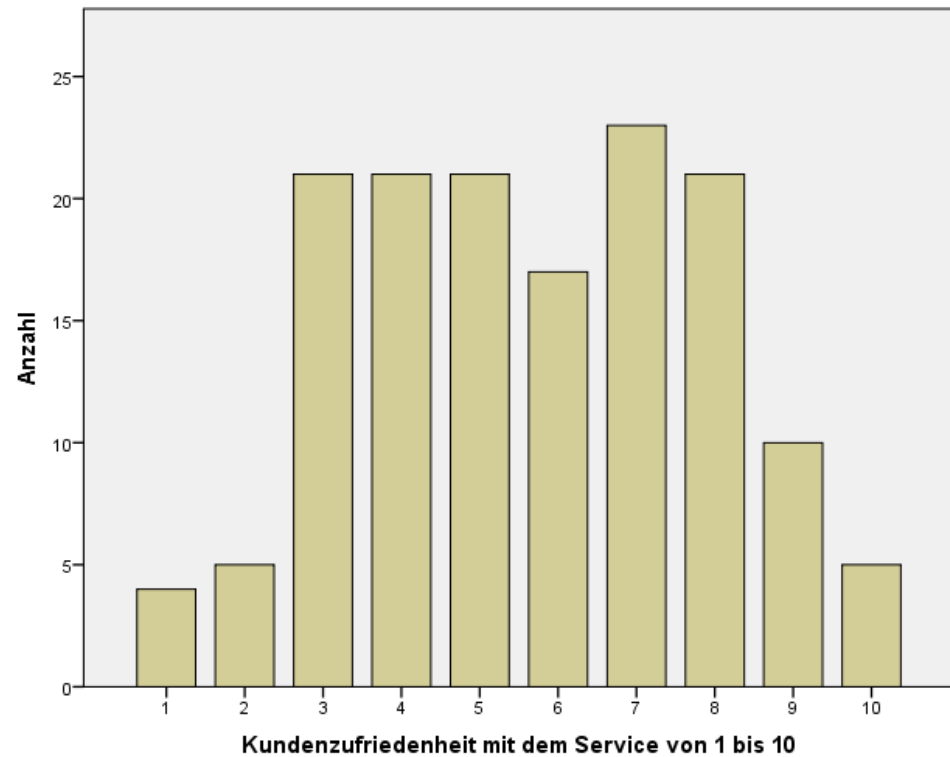


Hochschule Osnabrück  
University of Applied Sciences

## Statistiken

Kundenzufriedenheit mit dem Service

N	Gültig	148
	Fehlend	2
Mittelwert		5,66
Median		6,00
Modalwert		7
Bereich		9
Minimum		1
Maximum		10



# Aufgabe 3 - Häufigkeitstabellen



Hochschule Osnabrück  
University of Applied Sciences



Geben Sie an, wie die Ausprägungen in der Variable „Geschlecht“ und „Kleidungsgröße“ verteilt sind

# Lösung Aufgabe 3 - Häufigkeitstabellen



Hochschule Osnabrück  
University of Applied Sciences

Häufigkeitstabellen:

**Geschlecht (0 - männlich, 1 - weiblich)**

	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig männlich	82	54,7	54,7	54,7
weiblich	68	45,3	45,3	100,0
Gesamtsumme	150	100,0	100,0	

**Kundenzufriedenheit mit dem Service von 1 bis 10**

	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig 1	4	2,7	2,7	2,7
2	5	3,3	3,4	6,1
3	21	14,0	14,2	20,3
4	21	14,0	14,2	34,5
5	21	14,0	14,2	48,6
6	17	11,3	11,5	60,1
7	23	15,3	15,5	75,7
8	21	14,0	14,2	89,9
9	10	6,7	6,8	96,6
10	5	3,3	3,4	100,0
Gesamtsumme	148	98,7	100,0	
Fehlend -999	2	1,3		
Gesamtsumme	150	100,0		



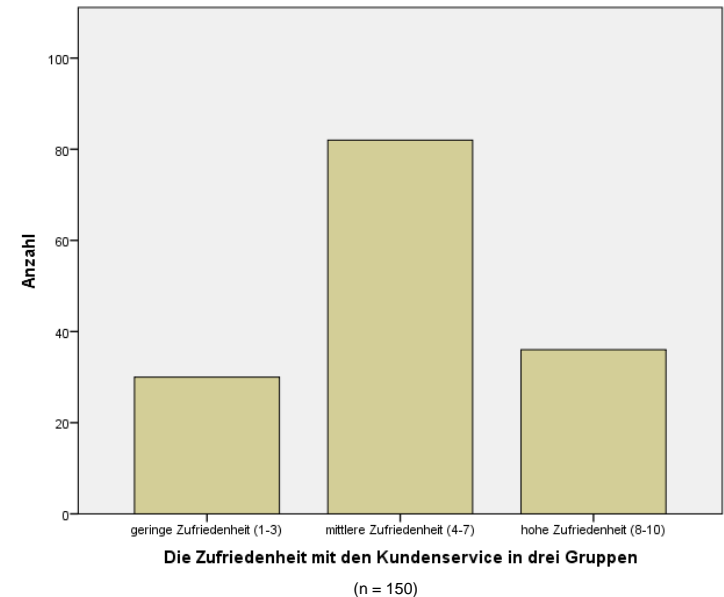
# Datenaufbereitung

## ➔ Klassenbildung

Wenn Sie ein aussagekräftiges Diagramm für die Verteilung eines Merkmals erstellen wollen, dieses aber viele (ungleiche) Werte ausweist, eignet sich die Klassierung zur Datenaggregation.



➔ Gruppieren



# Deskriptive Statistik

➔ Transformieren: Umkodieren von Variablen (bspw. um zu aggregieren)

The image shows the SPSS 'Transformieren' (Transform) menu and the 'Umkodieren in andere Variablen' (Recode into Different Variables) dialog box. The 'Umkodieren in andere Variablen...' option is highlighted in the menu, marked with a red circle and the number 1. The dialog box is open, showing the 'Umkodieren in andere Variablen' window. The 'Numerische Var. -> Ausgabevar.' (Numeric Variable -> Output Variable) section is active, with 'Körpergröße --> ?' selected. The 'Ausgabedatenvariable' (Output Variable) section shows 'Name: Körpergröße\_gruppiert' and 'Beschriftung: Körpergröße in zwei Gruppen z'. The 'Umkodieren in andere Variablen: Alte und neue Werte' (Recode into Different Variables: Old and New Values) sub-dialog box is also open, showing the 'Alter Wert' (Old Value) and 'Neuer Wert' (New Value) sections. The 'Bereich' (Range) option is selected under 'Alter Wert', with '0 thru 182' and '182 thru 190' as examples. The 'Neuer Wert' section shows 'Wert' (Value) as the selected option. The 'Alt -> Neu:' (Old -> New:) list shows the mapping: '0 thru 182 --> 1', '182 thru 190 --> 2', and '190 thru 230 --> 3'. The 'Ausgabe der Variablen als Zeichenfolgen' (Output Variable as Strings) checkbox is checked, and the 'Num. Zeichenfolgen in Zahlen umwandeln' (Convert numeric strings to numbers) checkbox is unchecked. The 'Breite' (Width) is set to 8. The 'Weiter' (Next) button is highlighted with a red circle and the number 3.

Transformieren   Analysieren   Direktmarketing   Diagramme

Variable berechnen...  
Werte in Fällen zählen...  
Werte verschieben...  
Umkodieren in dieselben Variablen...  
**Umkodieren in andere Variablen...**  
Automatisch umkodieren...  
Visuelles Klassieren...

Umkodieren in andere Variablen

Numerische Var. -> Ausgabevar.:  
Körpergröße --> ?

Ausgabedatenvariable  
Name: Körpergröße\_gruppiert  
Beschriftung: Körpergröße in zwei Gruppen z  
Ändern

Umkodieren in andere Variablen: Alte und neue Werte

Alter Wert  
☐ Wert:  
☐ Systemdefiniert fehlend  
☐ System- oder benutzerdefiniert fehlende Werte  
☒ Bereich:  
bis  
☐ Bereich, KLEINSTER bis Wert:  
☐ Bereich, Wert bis GRÖSSTER:  
☐ Alle anderen Werte

Neuer Wert  
☒ Wert:  
☐ Systemdefiniert fehlend  
☐ Alte Werte kopieren

Alt -> Neu:  
0 thru 182 --> 1  
182 thru 190 --> 2  
190 thru 230 --> 3

Hinzufügen  
Ändern  
Entfernen

☒ Ausgabe der Variablen als Zeichenfolgen   Breite: 8  
☐ Num. Zeichenfolgen in Zahlen umwandeln ('5' -> 5)

Weiter   Abbrechen   Hilfe

# Aufgabe 4



**Transformieren Sie die Variable „Zufriedenheit“  
in eine neue gruppierte Variable  
z.B. „Zufriedenheit\_gruppiert“**



Zufriedenheit	Gruppiert
1 – 3	Geringe Zufriedenheit
4 – 7	Mittlere Zufriedenheit
8 – 10	Hohe Zufriedenheit

# Lösung zur Aufgabe 4



Hochschule Osnabrück  
University of Applied Sciences

## 1. Rekodierung

Umcodieren in andere Variablen: Alte und neue Werte

**Alter Wert**

☐ Wert:

☐ Systemdefiniert fehlend

☐ System- oder benutzerdefiniert fehlende Werte

☒ Bereich:

bis

☐ Bereich, KLEINSTER bis Wert:

☐ Bereich, Wert bis GRÖSSTER:

☐ Alle anderen Werte

**Neuer Wert**

☒ Wert:

☐ Systemdefiniert fehlend

☐ Alte Werte kopieren

Alt --> Neu:

1 thru 3 --> 1  
4 thru 6 --> 2  
7 thru 10 --> 3

Hinzufügen  
Ändern  
Entfernen

**! Achtung bei kontinuierlichen Variablen**  
(siehe Folie "Transformieren: Umkodieren von Variablen")

☐ Ausgabe der Variablen als Zeichenfolgen Breite: 8

☐ Num. Zeichenfolgen in Zahlen umwandeln ('5' -> 5)

Weiter Abbrechen Hilfe

# Lösung zur Aufgabe 4



Hochschule Osnabrück  
University of Applied Sciences

## 2. Neue Variable benennen

The screenshot shows the 'Umkodieren in andere Variablen' (Recode into Different Variables) dialog box in SPSS. On the left, a list of variables is shown, including 'Zufriedenheit'. In the center, 'Numerische Var. -> Ausgabevar.: Zufriedenheit -> Zufriedenheit\_Gruppiert' is selected. On the right, the 'Ausgabevariable' (Output Variable) section is circled in red, showing 'Name: Zufriedenheit\_Gruppiert' and 'Beschriftung: Zufriedenheit in drei Gruppen'. Below this, there are buttons for 'Ändern' (Change), 'Alte und neue Werte...' (Old and New Values...), and 'Falls...' (Optional). At the bottom are buttons for 'OK', 'Einfügen' (Paste), 'Zurücksetzen' (Reset), 'Abbrechen' (Cancel), and 'Hilfe' (Help).

Erforderlich um die  
Umkodierung  
auszuführen

## 3. Wertelabels benennen

The screenshot shows the 'Wertbeschriftungen' (Value Labels) dialog box in SPSS. It has a 'Wert:' (Value) field and a 'Beschriftung:' (Label) field. A 'Rechtschreibung...' (Spelling...) button is located to the right of the 'Wert:' field. Below these fields, there are buttons for 'Hinzufügen' (Add), 'Ändern' (Change), and 'Entfernen' (Remove). A list box on the right contains the following entries: '1 = "gering"', '2 = "mittel"', and '3 = "hoch"'. At the bottom are buttons for 'OK', 'Abbrechen' (Cancel), and 'Hilfe' (Help).

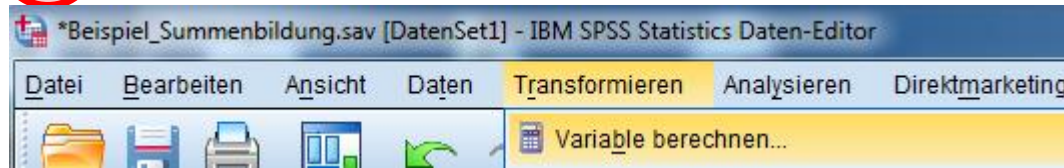
# Weitere Transformationen - Summenbildung



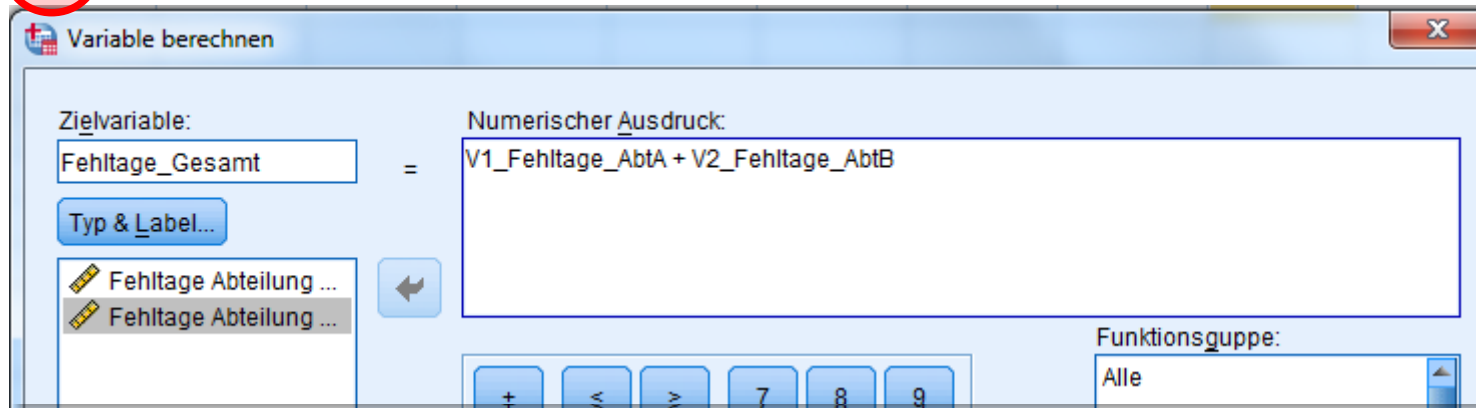
V1_Fehltage_AbtA	V2_Fehltage_AbtB
33	23
15	2
68	34
1	5
19	145
34	34
24	15
44	26
31	44
3	23
2	15
19	26
23	35
12	37
23	3

Die Verrechnung von Variablen funktioniert über „Variable berechnen“

1



2



## Aufgabe 4 - Weitere Transformationen



Hochschule Osnabrück  
University of Applied Sciences

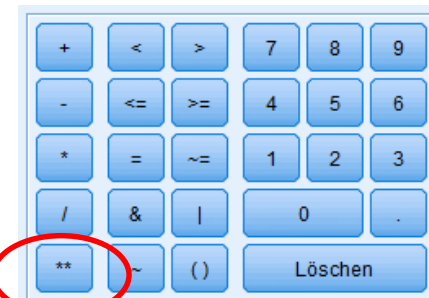
➔ Rechnen Sie die Körpergröße (momentan in cm) in Meter um und passen Sie die Variablenbeschreibung in der Variablenansicht an



$$\text{Körperlänge in m} = \frac{\text{Körperlänge in cm}}{100}$$

➔ Berechnen Sie den Body Maß Index für die Kunden

$$BMI = \frac{\text{Körpergewicht in kg}}{\text{Körperlänge in m}^2}$$





# Lösung Aufgabe 4 - Weitere Transformationen



Hochschule Osnabrück  
University of Applied Sciences

Umrechnen der Körpergröße:

Variable berechnen

Zielvariable: Körpergröße\_m = Numerischer Ausdruck: Körpergröße / 100

Typ & Beschriftung...

Kunden\_ID  
Telefonnummer  
Geschlecht  
Körpergröße  
Kleidergröße  
Umsatz  
Gewicht  
Zufriedenheit  
Alter  
Körpergröße\_cm

+ < > 7 8 9  
- <= >= 4 5 6  
\* = ~= 1 2 3  
/ & | 0 .  
\*\* ~ ( ) Löschen

Funktionsgruppe:  
Alle  
Arithmetisch  
Verteilungsfunktionen  
Umwandlung  
Aktuelles Datum/aktuelle Uhrzeit  
Datumsarithmetik  
Datumserstellung

Funktionen und Sondervariablen:

Falls... (optionale Fallauswahlbedingung)

OK Einfügen Zurücksetzen Abbrechen Hilfe

# Lösung Aufgabe 4 - Weitere Transformationen

Berechnen des BMI:



Hochschule Osnabrück  
University of Applied Sciences

Variable berechnen

Zielvariable: BMI

Numerischer Ausdruck: Gewicht/Körpergröße\_m \*\* 2

Typ & Beschriftung...

Kunden\_ID  
Telefonnummer  
Geschlecht  
Körpergröße  
Kleidergröße  
Umsatz  
Gewicht  
Zufriedenheit  
Alter  
Körpergröße\_cm  
Körpergröße\_m

+ < > 7 8 9  
- <= >= 4 5 6  
\* = ~= 1 2 3  
/ & | 0 .  
\*\* ~ ( ) Löschen

Funktionsgruppe:  
Alle  
Arithmetisch  
Verteilungsfunktionen  
Umwandlung  
Aktuelles Datum/aktuelle Uhrzeit  
Datumsarithmetik  
Datumserstellung

Funktionen und Sondervariablen:

Falls... (optionale Fallauswahlbedingung)

OK Einfügen Zurücksetzen Abbrechen Hilfe

# Fälle auswählen

Sollen Variablen nach den Ausprägungen einer anderen Variable gefiltert werden, bspw. nur männliche Teilnehmer, so funktioniert das über „Fälle auswählen“.



1

Daten Transformieren

- Variableneigenschaften
- Messniveau für unbekannt
- Dateneigenschaften kopieren
- Neues benutzerdefiniertes Format definieren...
- Mehrfachantworten-Sets
- Validierung
- Doppelte Fälle ermitteln...
- Ungewöhnliche Fälle identifizieren...
- Daten-Sets vergleichen...
- Fälle sortieren...
- Variablen sortieren...
- Transponieren...
- Dateien zusammenfügen
- Umstrukturieren...
- Aggregieren...
- Orthogonales Design
- Datenblatt kopieren
- Datei aufteilen...
- Fälle auswählen...
- Fälle gewichten...

2

Fälle auswählen

Auswählen

- ☐ Alle Fälle
- ☒ Falls Bedingung zutrifft
- ☐ Zufallsstichprobe
- ☐ Nach Zeit- oder Fallbereich

Falls...

Stichprobe...

Kunden\_ID

Geschlecht

Körpergröße

Kleidergröße

Umsatz

Gewicht

Zufriedenheit

Alter

3

Fälle auswählen: Falls

Geschlecht = 0

+ < > 7 8 9

	Kunden_ID	Geschlecht
43	43	0
44	44	0
45	45	0
46	46	0
47	47	0
48	48	0
49	49	0
50	50	1
51	51	0
52	52	0
53	53	0
54	54	1
55	55	0
56	56	0
57	57	0
58	58	0
59	59	1
60	60	1
61	61	0
62	62	0
63	63	0

# Aufgabe 5 - Fälle auswählen



Hochschule Osnabrück  
University of Applied Sciences



Wählen Sie für aus dem aktuellen Datensatz nur die Kunden mit der Größe **S** aus!

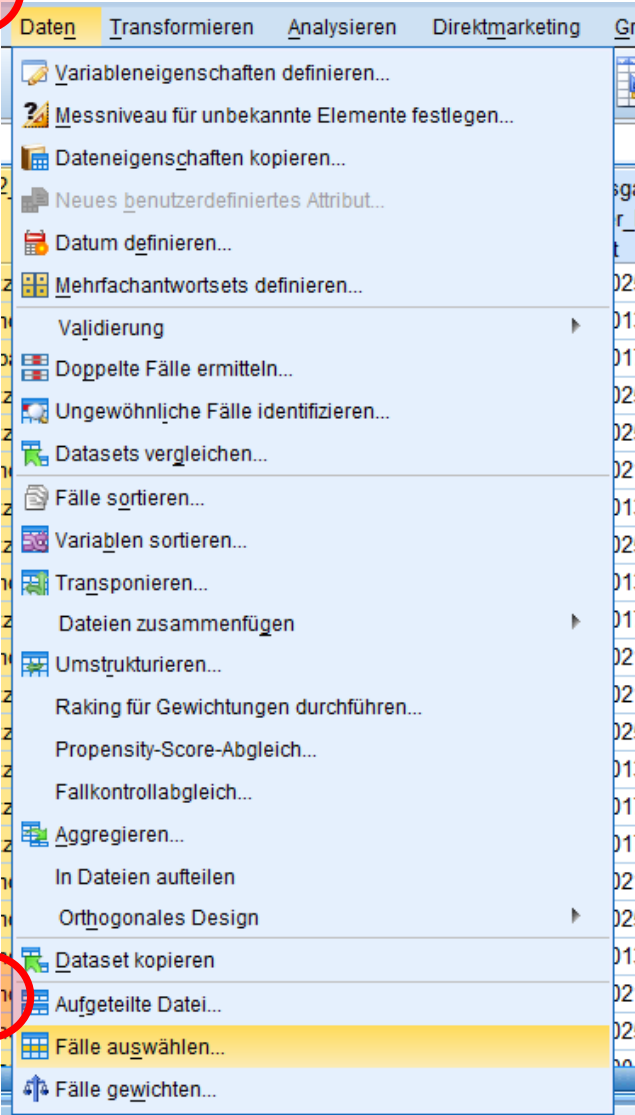
# Lösung Aufgabe 5 - Fälle auswählen



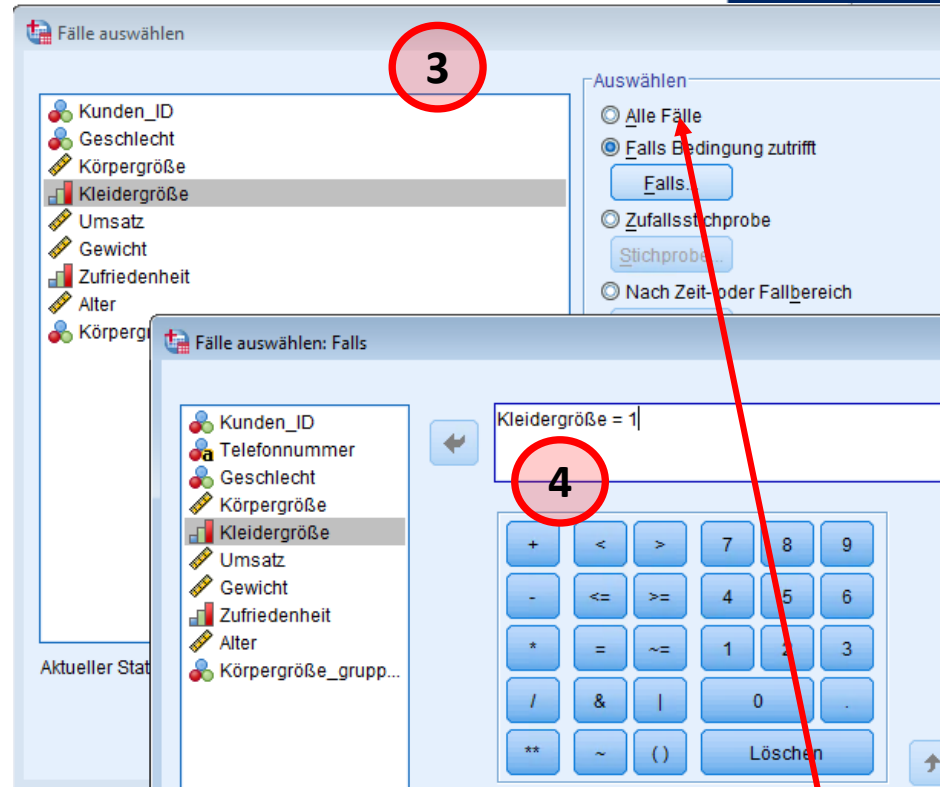
Hochschule Osnabrück  
University of Applied Sciences

➔ Wählen Sie die Kunden mit der Kleidergröße **S** aus

1



2



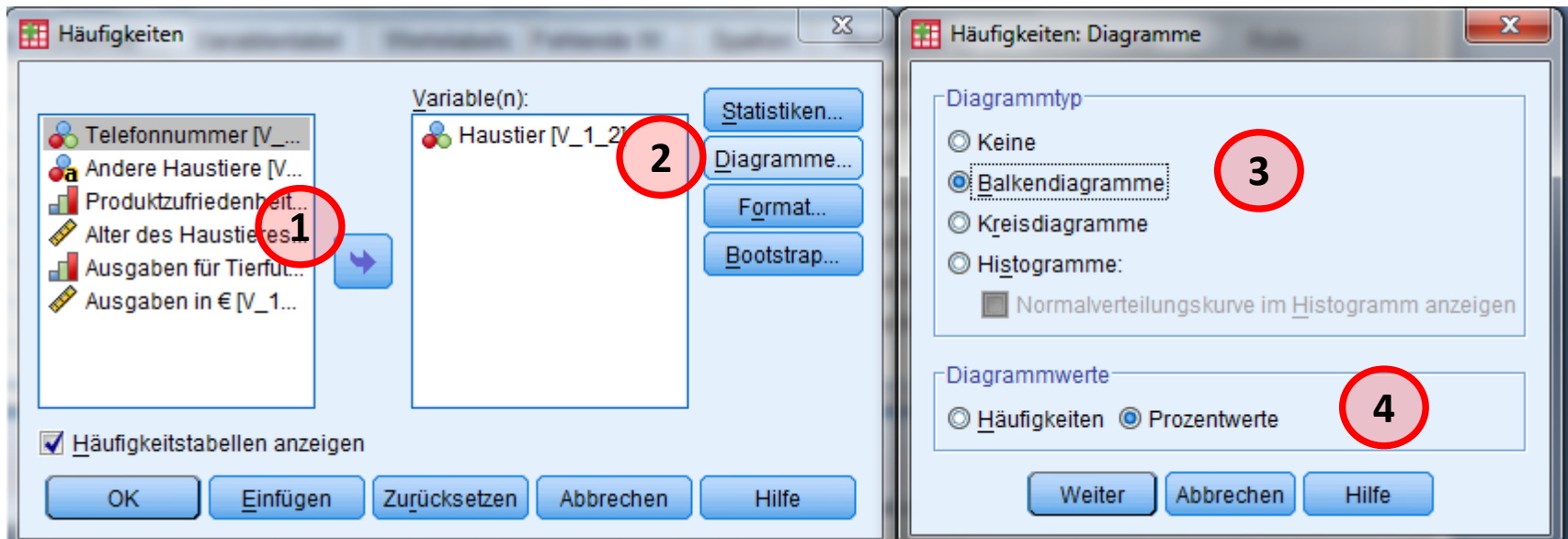
Die Fallauswahl kann im Menüfenster „Fälle auswählen“ mit „Alle Fälle“ rückgängig gemacht werden

# Deskriptive Statistik

## ➔ Diagramme



Hochschule Osnabrück  
University of Applied Sciences



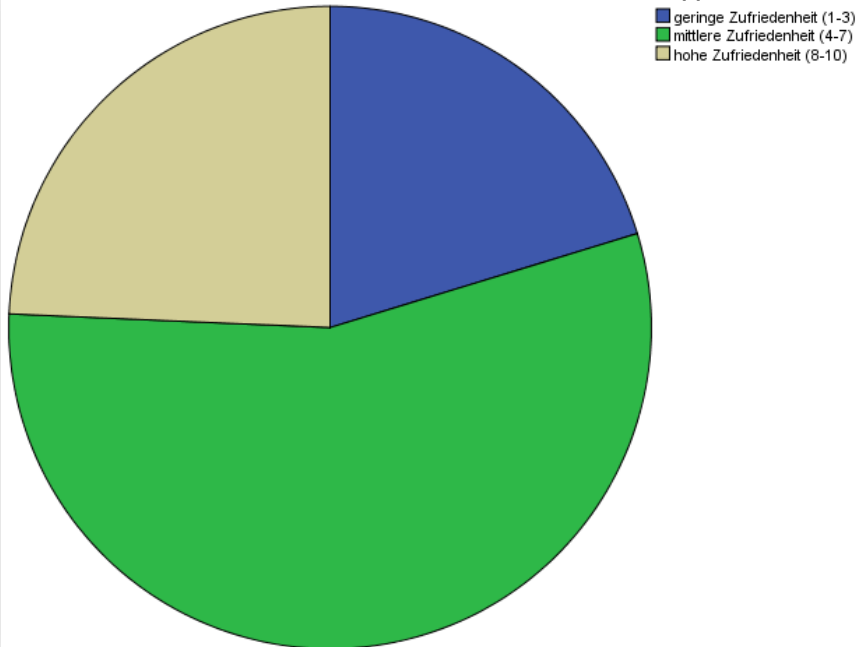
➔ Kreisdiagramm zumeist bei Variablen mit nur zwei Ausprägungen

# Aufgabe 6 - Diagramme

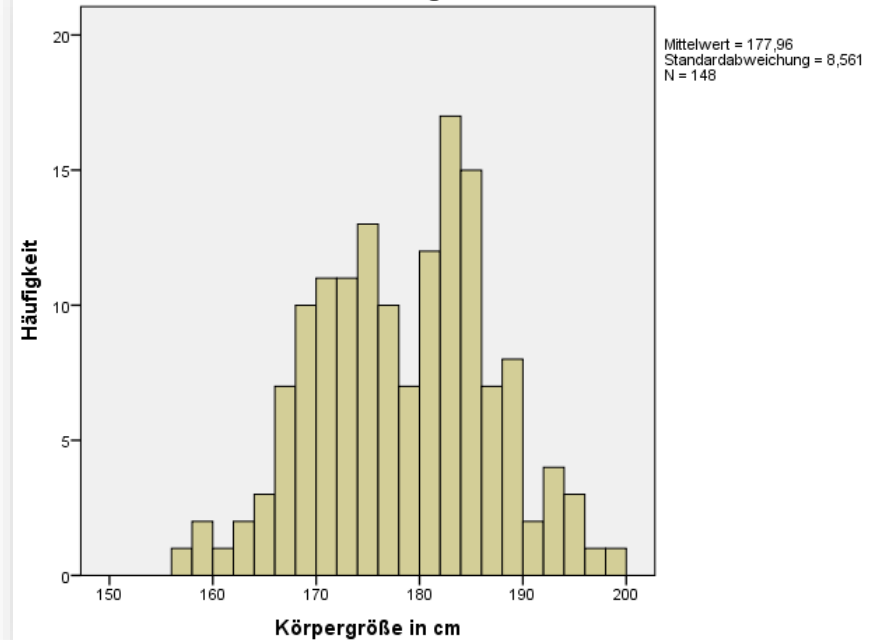
➔ Bilden Sie passende Diagramme für die Variablen „Körpergröße“ und “Zufriedenheit\_gruppiert“ um die Häufigkeiten zu visualisieren.

# Lösung Aufgabe 6 - Diagramme

Die Zufriedenheit mit den Kundenservice in drei Gruppen



Histogramm





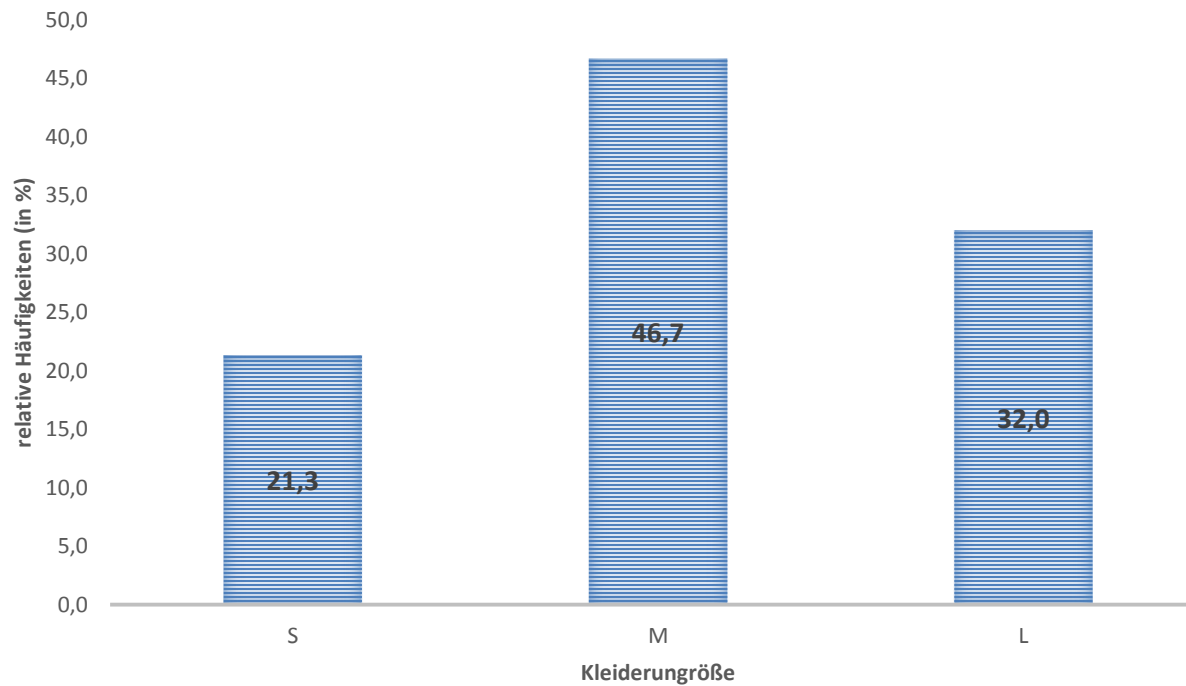
# Excel Grafiken

➔ Zu Präsentationszwecken Übertragung in Excel empfehlenswert...



Hochschule Osnabrück  
University of Applied Sciences

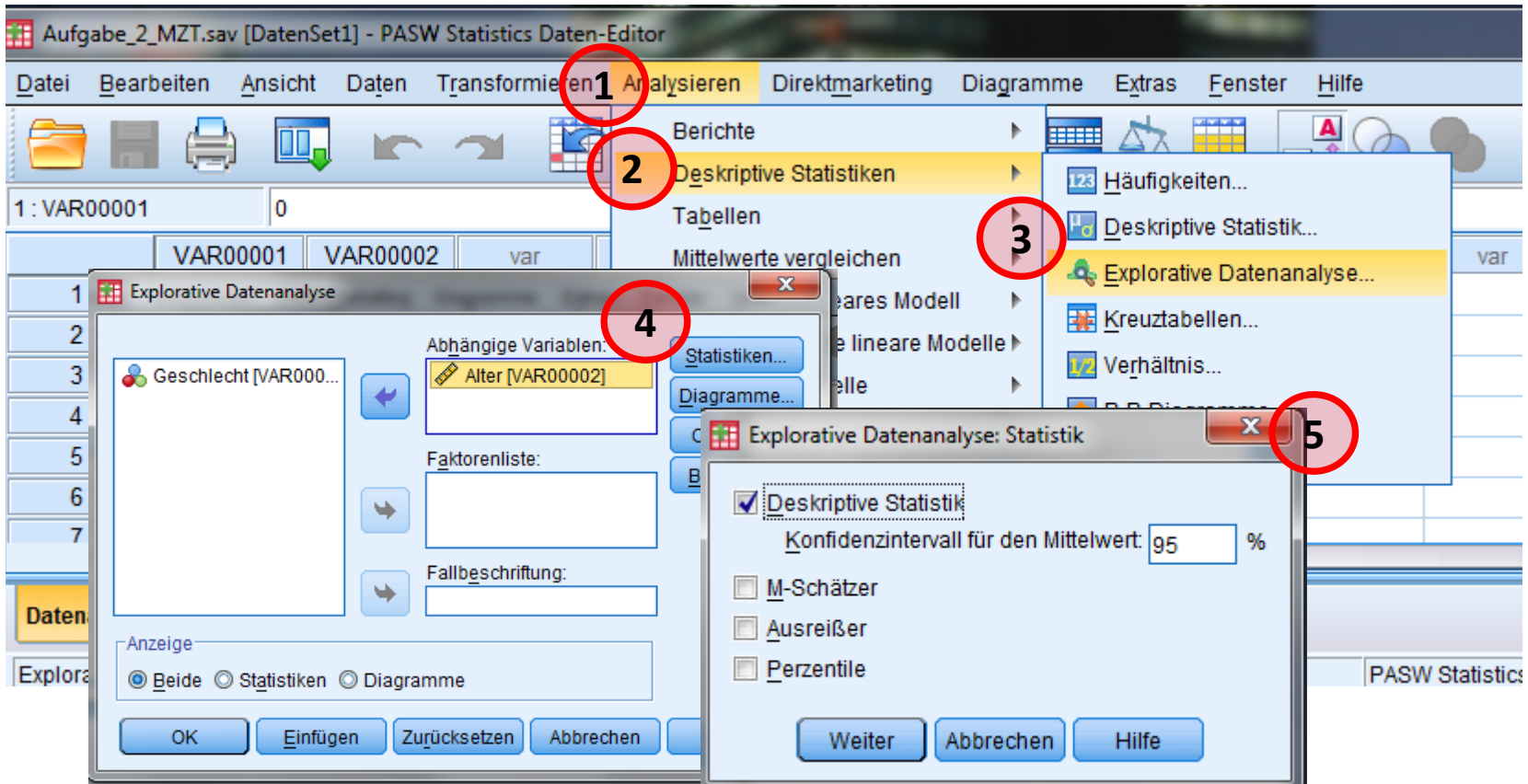
## VERTEILUNG DER KLEIDERGRÖÖE



# Deskriptive Statistik

➔ Berechnung von Mittelwerten und Konfidenz-Intervallen

„Analysieren“ → „Deskriptive Statistiken“ → „Explorative Datenanalyse“.  
Dort die interessierende Variable als abhängige Variable wählen und unter  
„Statistiken“ die Ausgabe eines 95%igen Konfidenzintervalls bestimmen



# Deskriptive Statistik



## Berechnung von Konfidenz-Intervallen

### Explorative Datenanalyse

[DatenSet1] C:\Users\Lab4Apps\Desktop\Lehre\WiConnect\SPSS-Seminar\_JDL\



Hochschule Osnabrück  
University of Applied Sciences

Verarbeitete Fälle

	Fälle					
	Gültig		Fehlend		Gesamt	
	N	Prozent	N	Prozent	N	Prozent
Alter	28	93,3%	2	6,7%	30	100,0%

Deskriptive Statistik

		Statistik	Standardfehler
Alter	Mittelwert	23,43	,906
	95% Konfidenzintervall des Mittelwerts		
	Untergrenze	21,57	
	Obergrenze	25,29	
	5% getrimmtes Mittel	23,21	
	Median	22,00	
	Varianz	22,995	
	Standardabweichung	4,795	
	Minimum	17	
	Maximum	34	
	Spannweite	17	
	Interquartilbereich	7	
	Schiefe	,872	,441
	Kurtosis	-,362	,858

# Deskriptive Statistik

➔ Berechnung von Konfidenz-Intervallen



Hochschule Osnabrück  
University of Applied Sciences

Sie können für eine numerische Variable einen Vertrauensbereich für den **Erwartungswert** berechnen. Der **Erwartungswert** ist eine (Ihnen unbekannte) Kenngröße der Grundgesamtheit, das **arithmetische Mittel** ist eine (Ihnen bekannt) Kenngröße der Stichprobe.

Der Vertrauensbereich enthält einen unbekannten Parameter  $\mu$ , hier den Erwartungswert, einer Verteilung mit einer Sicherheit von z.B. 95% und einer entsprechenden Irrtumswahrscheinlichkeit von  $\alpha$  5%.

**Das berechnete Konfidenz-Intervall enthält den wahren Schätzer mit großer Wahrscheinlichkeit...**



## Aufgabe 6 - Berechnung statistischer Kennwerte

- a) Lassen Sie sich jeweils eine Häufigkeitstabelle für die Variablen „Kleiderungsgröße“.
- b) Bestimmen Sie Mittelwert, sowie Variationsbreite (Spannweite) und die Standardabweichung des Umsatz in Euro.
- c) Welches Diagramm bietet sich zur Visualisierung an?
- d) Erzielen zufriedene Kunden einen höheren Umsatz? (Arbeiten Sie mit der Funktion „Analyse“ -> „explorative Datenanalyse“ und nutzen Sie die Faktorliste)

# Lösung Aufgabe 7 a + b + c



Hochschule Osnabrück  
University of Applied Sciences

**Kleidergröße (S, M oder L)**

a)

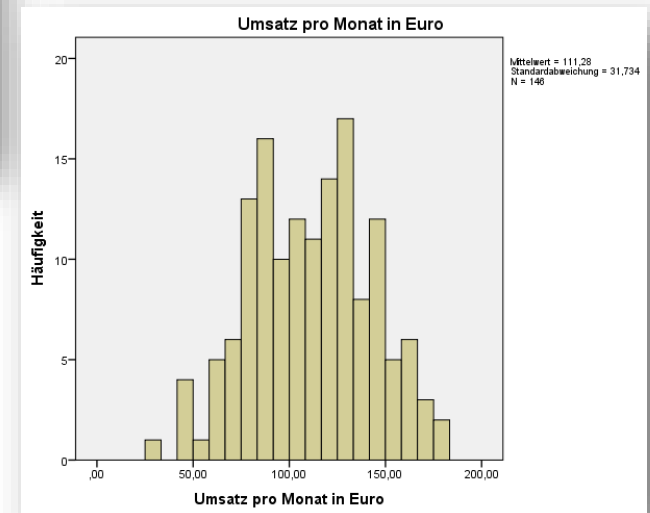
	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig S	32	21,3	21,3	21,3
M	70	46,7	46,7	68,0
L	48	32,0	32,0	100,0
Gesamtsumme	150	100,0	100,0	

**Deskriptive Statistiken**

b)

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Umsatz pro Monat in Euro	146	28,73	182,20	111,2769	31,73382
Gültige Anzahl (listenweise)	146				

c)



# Lösung Aufgabe 7 d

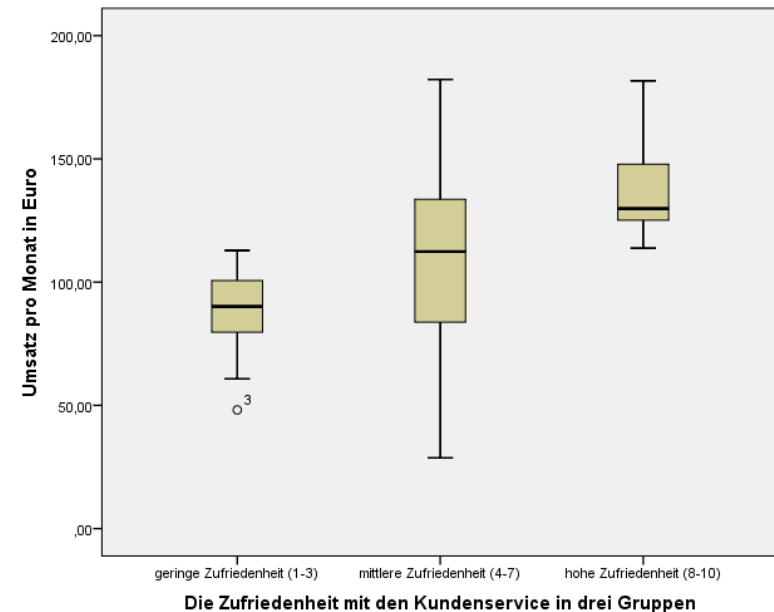


Hochschule Osnabrück  
University of Applied Sciences

## Deskriptive Statistik

Die Zufriedenheit mit den Kundenservice in drei Gruppen				Statistik	Standardfehler
Umsatz pro Monat in Euro	geringe Zufriedenheit (1-3)	Mittelwert		88,6823	2,85154
		95 % Konfidenzintervall für Mittelwert	Untergrenze	82,8503	
			Obergrenze	94,5144	
		5% getrimmter Mittelwert		89,4519	
		Median		90,1050	
		Varianz		243,938	
		Standardabweichung		15,61851	
		Minimum		48,13	
		Maximum		112,84	
		Bereich		64,71	
		Interquartilbereich		21,63	
		Schiefe		-,670	,427
		Kurtosis		,068	,833
	mittlere Zufriedenheit (4-7)	Mittelwert		108,5379	3,79422
		95 % Konfidenzintervall für Mittelwert	Untergrenze	100,9827	
			Obergrenze	116,0932	
		5% getrimmter Mittelwert		108,8406	
		Median		112,3800	
		Varianz		1122,897	
		Standardabweichung		33,50965	
		Minimum		28,73	
		Maximum		182,20	
		Bereich		153,47	
		Interquartilbereich		50,42	
		Schiefe		-,098	,272
		Kurtosis		-,598	,538
	hohe Zufriedenheit (8-10)	Mittelwert		137,4267	2,91675
		95 % Konfidenzintervall für Mittelwert	Untergrenze	131,5053	
			Obergrenze	143,3480	
		5% getrimmter Mittelwert		136,4360	
		Median		129,8050	
		Varianz		306,268	
		Standardabweichung		17,50050	
		Minimum		113,82	
		Maximum		181,67	
		Bereich		67,85	
		Interquartilbereich		22,97	
		Schiefe		,967	,393
		Kurtosis		,213	,768

Umsatz pro Monat in Euro



# SPSS Zugriff über den Virtual Desktop der HS Osnabrück



Hochschule Osnabrück  
University of Applied Sciences



Eine komplette Anleitung für den SPSS Zugriff ist [hier](#) verfügbar

Die HS besitzt eine Campuslizenz für SPSS

Mit SPSS kann auf allen Rechnern der HS gearbeitet werden (am Campus)

Von zu Hause kann über den [Virtual Desktop](#) mit der Benutzer Kennung zugegriffen werden

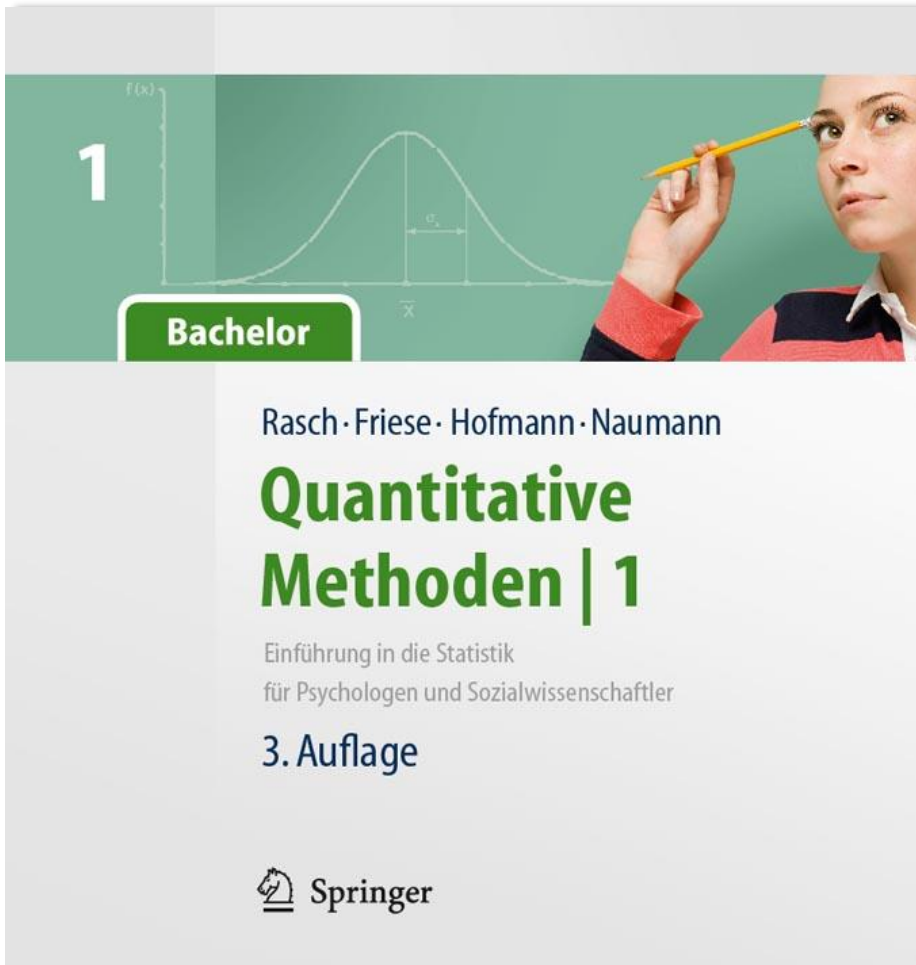
Bei Fragen zum Virtual Desktop bitte an wenden an folgende Email:  
[support@edvsz.hs-osnabrueck.de](mailto:support@edvsz.hs-osnabrueck.de)



# Empfehlungen Weiterführende Literatur



Hochschule Osnabrück  
University of Applied Sciences



In der Bibliothek an der Hochschule verfügbar

Im [Springerlink](#) der HS als PDF verfügbar

Internetauftritt mit Beispieldaten und Übungen

Weiterführender Teil ebenfalls im Springerlink verfügbar

# Methodenberatung



Hochschule Osnabrück  
University of Applied Sciences

**WiSo**



© MH/fotolia.com

FAKULTÄT WIRTSCHAFTS- UND  
SOZIALWISSENSCHAFTEN

**ZeMIT - Methodenberatung**

➔ FÜR STUDIERENDE, LEHRENDE UND ALLE MITARBEITER  
DER FAKULTÄT WISO

## ➔ KONTAKT

**Jan-David Liebe, M.A.**

Sedanstraße 1 (S1)

Tel.: 0541 / 969 - 3252

eMail: [j.liebe@hs-osnabrueck.de](mailto:j.liebe@hs-osnabrueck.de)



Homepage: <http://www.wiso.hs-osnabrueck.de/zemit.html>



# Weiterführende Literatur und Handbücher



Hochschule Osnabrück  
University of Applied Sciences

Bühl, A., SPSS 20, Einführung in die moderne Datenanalyse, 13. Auflage, München 2012.

Brosius, F., SPSS 20 für Dummies, Wiley-VCH Verlag, 2012.

Brosius, F., SPSS 19, Verlagsgruppe Hütig Jehle Rehm, Heidelberg 2011.

Akremit, L./Baur, N./Fromm, S., Datenanalyse mit SPSS für Fortgeschrittene, 3. Auflage, Wiesbaden 2011.

Fromm, S., Datenanalyse mit SPSS für Fortgeschrittene 2: Multivariate Verfahren für Querschnittsdaten, Wiesbaden 2010.

Backhaus, K./Erichson, B./Plinke, W./Weiber, R. (Hg.), Multivariate Analysemethoden. Eine anwendungsorientierte Einführung, 13. Auflage, Berlin 2011.

Janssen, J./Laatz, W., Statistische Datenanalyse mit SPSS für Windows, Berlin 2010.

Sarstedt, M./Schütz, T./Raithel, S., IBM SPSS Syntax - Eine anwendungsorientierte Einführung. 2. Auflage, Vahlen Verlag, München 2010.

Hatzinger, R./Nagel, H., PASW Statistics: Statistische Methoden und Fallbeispiele, München 2009.