### SPSS – Kurs online Lösungen

3) Erzeugen Sie eine Variable, die die Differenz zwischen fef50 und fef75 angibt, Name: diff\_fef, Berechnung: fef50 – fef75

Transformieren → Variable berechnen



5) Variable geb\_jz (Jahreszeit Geburt) ausgehend von der Variable gebmo erzeugen:

Transformieren  $\rightarrow$  Umcodieren in andere Variablen

| <b>t</b>         | Umcodieren in andere Variablen                                                                                                                                                                                      | × |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|
| Bildungsstand El | Numerische Var> Ausgabevar.:   gebmo> geb_jz   geb_jz   Beschriftung:   Geburt Jahreszeit     Ändern     Alte und neue Werte     Ealls (optionale Fallauswahlbedingung)     Einfügen     Zurücksetzen     Abbrechen |   |

#### 5ff) Schaltfläche "Alte und neue Werte"

| Umcodieren in an                                                                                                                                                                                                                                                                       | dere Variablen: Alte und neue Werte                                                                                                                                                                                                                                                                                                                                                                    | × |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|
| Alter Wert            Wert:             Systemdefiniert fehlend             System- oder benutzerdefiniert fehlende Werte             Bereich:             bis             Bereich, KLEINSTER bis Wert:             Bereich, Wert bis <u>G</u> RÖSSTER:             Alle anderen Werte | Neuer Wert            • Wert:             • Systemdefiniert fehlend             • Alte Werte kogieren             Alt> Neu:             Alt> Neu:             Alt> Neu:             3 thru 5> 'Frühjahr'         6 thru 8> 'Sommer'             9 thru 11> 'Herbst'             ELSE> 'Winter'             Entfernen             Mum. Zeichenfolgen in Zahlen umwandeln ('5'->5)             Abbrechen |   |

<u>Beachten</u>: Häkchen setzen bei "Ausgabe der Variablen als Zeichenfolgen" oder anstatt "Winter, etc." eine Codierung (1,2,…) nehmen.

6) Filter setzen: männliche Kinder mit mindestens 2 Vorerkrankungen (Anz\_Vorerkr1) auswählen

Daten → Fälle auswählen, Schaltfläche "Falls Bedingung zutrifft",

Formel wie gezeigt eintragen.

| <b>t</b> a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Fälle auswählen: Falls      | X         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-----------|
| <ul> <li>Image: Arrow of the second second</li></ul> | Anz_Vorerkr1 >= 2 & sex = 1 | z<br>Ien: |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Weiter Abbrechen Hilfe      |           |

 Unterschied zwischen alten und jungen Kindern hinsichtlich der Anzahl an Vorerkrankungen grafisch und per Kreuztabelle zeigen. Grafik → Diagrammerstellung



1ff) Kreuztabelle: Analysieren → Deskriptive Statistiken → Kreuztabellen, Schaltfläche "Statistiken" Häkchen bei Chi-Quadrat; Schaltfläche Zellen, Häkchen bei Häufigkeiten Beobachtet und Erwartet.

#### Anz\_Vorerkr1 \* Geburtsjahr in Klassen Kreuztabelle

|              |   |                  | Geburtsjahr |       |        |
|--------------|---|------------------|-------------|-------|--------|
|              |   |                  | alt         | jung  | Gesamt |
| Anz_Vorerkr1 | 0 | Anzahl           | 375         | 631   | 1006   |
|              |   | Erwartete Anzahl | 363,0       | 643,0 | 1006,0 |
|              | 1 | Anzahl           | 127         | 226   | 353    |
|              |   | Erwartete Anzahl | 127,4       | 225,6 | 353,0  |
|              | 2 | Anzahl           | 43          | 91    | 134    |
|              |   | Erwartete Anzahl | 48,4        | 85,6  | 134,0  |
|              | 3 | Anzahl           | 13          | 35    | 48     |
|              |   | Erwartete Anzahl | 17,3        | 30,7  | 48,0   |
|              | 4 | Anzahl           | 1           | 7     | 8      |
|              |   | Erwartete Anzahl | 2,9         | 5,1   | 8,0    |
| Gesamt       |   | Anzahl           | 559         | 990   | 1549   |
|              |   | Erwartete Anzahl | 559,0       | 990,0 | 1549,0 |

#### Chi-Quadrat-Tests

|                                    | Wert   | df | Asymptotisch<br>e Signifikanz<br>(zweiseitig) |
|------------------------------------|--------|----|-----------------------------------------------|
| Chi-Quadrat nach<br>Pearson        | 5,164ª | 4  | ,271                                          |
| Likelihood-Quotient                | 5,601  | 4  | ,231                                          |
| Zusammenhang linear-<br>mit-linear | 4,170  | 1  | ,041                                          |
| Anzahl der gültigen Fälle          | 1549   |    |                                               |

 a. 1 Zellen (10,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 2,89.

#### p-Wert >0.05, also keine sign. Unterschiede zwischen Altersgruppen

2) Häufigkeit allergische Atemwegserkrankungen, Kehlkopfentzündung und wie häufig beides zusammen

Analysieren  $\rightarrow$  Deskriptive Statistiken  $\rightarrow$  Häufigkeiten



Hiermit erhält man die einzelnen Häufigkeiten. Um zu erfahren, bei welchen Probanden beides zusammen auftritt: neue Variable erstellen, z.B. über **Transformieren →** 

Variable berechnen, Numerischer Ausdruck =aller+kehle, und sich dann auf dem gleichen Weg wie oben die Häufigkeit angeben lassen. Die "2er" haben beide Vorerkrankungen.

# Ubung 2

3) Kinder in drei Größengruppen nach Geschlecht einteilen Transformieren  $\rightarrow$  Umcodieren in andere Variablen Schaltfläche "Alte und neue Werte", zuerst Eingabe der Werte für die Mädchen

| maderiem                                               | Umcodieren in andere Variablen: Alte und neue Werte                                                                                                                                                                                       | ×                                                                       |
|--------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Hier<br>nacheinander<br>die jeweiligen<br>Bereiche der | Alter Wert   O Wert:   O Systemdefiniert fehlend   O System- oder benutzerdefiniert fehlende Werte   O Bereich:   D Bereich, KLEINSTER bis Wert:   O Bereich, KLEINSTER bis Wert:   O Bereich, Wert bis GRÖSSTER:   O Alte order on Worte | Hier<br>nacheinander<br>die Werte 1, 2<br>und 3<br>eingeben,<br>jeweils |
| angeben                                                | Alle anderen werte     Ausgabe der Variablen als Zeichenfolgen Breite:     Num. Zeichenfolgen in Zahlen umwandeln (5'->5)     Weiter Abbrechen Hilfe                                                                                      | "Hinzufügen"<br>klicken                                                 |

Durch Klicken von "Weiter" kehrt man wieder zu dem vorherigen Fenster zurück.

3ff) Schaltfläche "Falls…" und Eintragen der Bedingung "sex=2", damit die vorher festgelegten Werte nur für die Mädchen

angewendet werden.

Für die Jungen wird der Vorgang wiederholt, die Grenzen der Körpergröße dementsprechend anpassen (vorherige Folie) und bei "Falls" als Bedingung sex=1 angeben. Der Variablenname bleibt gleich!!!



4) Bildungsstand der Eltern als Balkendiagramm Muster, Dicke der Balken, y-Achsenabschnitt verändern. Anmerkung einfügen; "rauva" als Unterteilungsvariable

Grafik → Diagrammerstellung, einfache Balken in das Vorschaufenster ziehen, Variable "sozio" auf x-Achse ziehen, OK klicken.

Dann Grafik doppelt klicken, Balken doppelt klicken, im Eigenschaften-Fenster auf *Füllung und Rahmen*, Muster verändern; dann zu *Optionen für Balken*, mit Schieber Dicke verändern; Eigenschaften Fenster schließen.

Dann y-Achse doppelt klicken, im Eigenschaften-Fenster auf *Skala*, dort die Werte verändern. Fenster schließen.

#### 4ff) Dann die Anmerkung einfügen:



Für die Unterteilung nach "rauva" zurück in Grafik →

Diagrammerstellung, die Gruppierten Balken oben ins Vorschaufenster ziehen, "rauva" als Clustervariable angeben.



# 1) Im Menü **Datei** → Öffnen → Daten, es öffnet sich folgendes Fenster:

| 👍 Daten öffne                                                          | n                                                                |         |                 | Damit die Excel-                                                                  |
|------------------------------------------------------------------------|------------------------------------------------------------------|---------|-----------------|-----------------------------------------------------------------------------------|
| Suchen in:                                                             | SPSSKurs                                                         | 🔽 🙆 🔯 🔚 |                 |                                                                                   |
| i vaso-Date<br>atemweg<br>atenwege<br>DatenFra<br>Fragebog<br>VarFrage | eien<br>e20.sav<br>e19.sav<br>gebogen.sav<br>en.sav<br>bogen.sav |         |                 | Dateien angezeigt<br>werden, unten die<br>Auswahl des Dateityps<br>setzen (gelb). |
| Dateiname:                                                             |                                                                  |         | Ö <u>f</u> fnen | Abfraça ab arata Zaila                                                            |
| Dateityp:                                                              | SPSS Statistics (*.sav)                                          | •       | Einfügen        | Abfrage, ob erste Zelle                                                           |
| String-La                                                              | SPSS Statistics (*.sav)<br>SPSS/PC+ (*.sys)                      |         | Abbrechen       | die Variablennamen                                                                |
|                                                                        | Systat (*.syd, *.sys)<br>Portable (*.por)                        |         |                 | sind: ja!                                                                         |
|                                                                        | Excel (*.xls, *.xlsx, *.xlsm)                                    |         |                 |                                                                                   |
|                                                                        | Lotus (*.w*)                                                     |         |                 |                                                                                   |
|                                                                        | dBase (*.dbf)                                                    | *       |                 | _                                                                                 |

#### 2) Überprüfen und vervollständigen Sie die Variablenübersicht Eintragen der Variablenlabels und bei "vaso" der Wertelabels gemäß Infoblatt der Daten

| 🔚 *Unb        | a *Unbenannt2 [DatenSet1] - IBM SPSS Statistics Daten-Editor |                 |                                |             |            |                            |                      |                              |                    |             |            |           |  |
|---------------|--------------------------------------------------------------|-----------------|--------------------------------|-------------|------------|----------------------------|----------------------|------------------------------|--------------------|-------------|------------|-----------|--|
| <u>D</u> atei | <u>B</u> earbeiten                                           | A <u>n</u> sich | t Da <u>t</u> en T <u>r</u> ai | nsformieren | Analysiere | n Direkt <u>m</u> arketing | ) Dia <u>g</u> ramme | E <u>x</u> tras <u>F</u> ens | ster <u>H</u> ilfe |             |            |           |  |
|               | 😑 H 🖨 🛄 🗠 🛥 📓 🏪 🗐 🕅 📓 🖉 🚍 🖧 🗒 📲 🎝 🍋 🤲                        |                 |                                |             |            |                            |                      |                              |                    |             |            |           |  |
|               |                                                              | Name            | Тур                            | Spaltenf    | Dezimal    | Variablenlabel             | Wertelabels          | Fehlende W                   | Spalten            | Ausrichtung | Messniveau | Rolle     |  |
| 1             | vol                                                          |                 | Numerisch                      | 12          | 6          | Eingeatmete Lu             | Keine                | Keine                        | 12                 | Rechts      | 🛷 Skala    | 🖒 Eingabe |  |
| 2             | rate                                                         |                 | Numerisch                      | 12          | 6          | Geschwndigkei              | Keine                | Keine                        | 12                 | Rechts      | 🛷 Skala    | 🖒 Eingabe |  |
| 3             | vaso                                                         |                 | Numerisch                      | 12          | 0          | Auftreten von V            | {1, keine Va         | Keine                        | 12                 | Rechts      | 💑 Nominal  | 🖒 Eingabe |  |
| A             |                                                              |                 |                                |             |            |                            |                      |                              |                    |             |            |           |  |

#### 3) Menüpunkt Analysieren → Deskriptive Statistiken → Häufigkeiten

#### Statistiken

Auftreten von Vasokonstriktion

| Ν | Gültig  | 39 |
|---|---------|----|
|   | Fehlend | 0  |

|   | Auftreten von Vasokonstriktion |                                |            |         |                     |                        |            |  |  |  |
|---|--------------------------------|--------------------------------|------------|---------|---------------------|------------------------|------------|--|--|--|
|   |                                |                                | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |            |  |  |  |
| ۲ | Gültig                         | keine Vasokonstriktion         | 20         | 51,3    | 51,3                | 51 Zun                 | n Aktivier |  |  |  |
|   |                                | Vasokonstriktion<br>beobachtet | 19         | 48,7    | 48,7                | 100, <sup>dor</sup>    | opelklicke |  |  |  |
|   |                                | Gesamt                         | 39         | 100,0   | 100,0               |                        |            |  |  |  |

#### Häufigkeiten

# Ubung 3

3) Führen Sie deskriptive Analyseverfahren durch: wie häufig treten bei der Variable "vaso" 1 und 2 auf. In welchem Bereich liegen die Werte von "vol" und "rate" (Spannweite)? Sind die Werte plausibel?

Menüpunkt Analysieren → Deskriptive Statistiken → Deskriptive Statistik, Variablenauswahl vol und rate.

Eventuell hier, spätestens aber beim Punkt 4 (Grafikerstellung) müsste auffallen, dass 0,604316 ändern, dann

so aus:

|                                                      | Ν  | Minimum   | Maximum     | Mittelwert  | Standardabw<br>eichung |
|------------------------------------------------------|----|-----------|-------------|-------------|------------------------|
| Eingeatmete Luftmenge                                | 39 | -,916291  | 1,308333    | ,15959000   | ,537651922             |
| Geschwndigkeit, mit der<br>die Luft eingeatmet wurde | 39 | -3,506558 | 604316,0000 | 15495,54000 | 96767,96301            |
| Gültige Werte<br>(Listenweise)                       | 39 |           |             |             |                        |

39

Deskriptive Statistik

der Maximalwert von "rate" extrem hoch ist. Bei Betrachtung der Werte wird klar, dass das Komma vergessen wurde. Also: Wert 33 in Deskriptive Statistik Standardabw Ν Minimum Mittelwert eichung sieht die gleiche Abfrage Maximum Eingeatmete Luftmenge 39 -,916291 1,308333 .15959000 ,537651922 Geschwndiakeit, mit der 39 -3,506558 1,321756 ,27344592 ,899923407 die Luft eingeatmet wurde

Gültige Werte

(Listenweise)

4) Erzeugen Sie Grafiken zu "vol" und "rate", um die Unterschiede in den Werten bei den beiden "vaso"-Gruppen zu zeigen.

Statistisch aussagekräftig wären hier z.B. Boxplots. Menüpunkt **Diagramme → Diagrammerstellung**, "Boxplots" auswählen, Bild "einfacher Boxplot" nach oben ziehen, x-Achse: "vaso", y-Achse: nacheinander "vol" und "rate".



Was sagen die Boxplots aus? Die Werte bei "vol" (Eingeatmete Luftmenge) scheinen bei beiden Ausprägungen ("vaso" 1 und 2) relativ normalverteilt zu sein (schwarzer Balken (Median) in der Mitte der "Box"), allerdings ist der eine "whisker" bei "vaso" 1 viel kürzer als der andere, es besteht hier also eine leichte Schiefe in der Verteilung. Die Varianzen scheinen relativ homogen zu sein (d.h. "Größe der Boxen beider Gruppen ungefähr gleich"). Bei der "Geschwindigkeit mit der die Luft eingeatmet wurde" kann auch eine leichte Schiefe in der Verteilung beobachtet werden (schwarzer Balken nicht mittig), insgesamt dürfte dies aber noch als normalverteilt gelten. Varianzhomogenität würde man hier nicht

mehr annehmen, da die Boxen doch recht unterschiedlich groß

sind.

5) Kann statistisch ein Unterschied zwischen den Werten der Variablen "vol" und "rate" abgesichert werden hinsichtlich des Auftretens von Vasokonstriktionen?

Zuerst muss man einen geeigneten Test auswählen. Die bei Punkt 4 gemachten Aussagen zu Verteilungsannahme und Varianzhomogenität sind hierzu entscheidend. Zusätzlich gibt es Tests, die diese objektiven Aussagen absichern können.

Der Shapiro-Wilk und der KS-Test testet die Normalverteilungsannahme: **Analysieren → Deskriptive Statistik → Explorative Datenanalyse** unter Punkt "Diagramme" Häkchen bei "Normalverteilungsdiagramm mit

Tests" setzen.

| Auttreten von<br>Vasokonstriktion |                                |           | -  |             | -         |    |             |  |
|-----------------------------------|--------------------------------|-----------|----|-------------|-----------|----|-------------|--|
|                                   |                                | Statistik | df | Signifikanz | Statistik | df | Signifikanz |  |
| Eingeatmete Luftmenge             | keine Vasokonstriktion         | ,137      | 20 | ,200        | ,932      | 20 | ,171        |  |
|                                   | Vasokonstriktion<br>beobachtet | ,122      | 19 | ,200        | ,980      | 19 | ,945        |  |
| Geschwndigkeit, mit der           | keine Vasokonstriktion         | ,112      | 20 | ,200        | ,961      | 20 | ,573        |  |
| die Luft eingeatmet wurde         | Vasokonstriktion<br>beobachtet | ,191      | 19 | ,068        | ,934      | 19 | ,204        |  |

Tests auf Normalverteilung

Kolmogorov-Smirnov<sup>a</sup>

Shapiro-Wilk

\*. Dies ist eine untere Grenze der echten Signifikanz.

a. Signifikanzkorrektur nach Lilliefors

In unserem Fall haben wir eine Fallzahl <50, würden also eher auf den Shapiro-Wilk Test zurück greifen. Der p-Wert (Werte, die unter "Signifikanz" stehen) ist in allen Fällen größer als 0,05, also kann die Nullhypothese "die Daten folgen der Normalverteilung" nicht verworfen werden. Wir dürfen also parametrische Testverfahren verwenden (Normalverteilung wird vorausgesetzt). Der Test auf Varianzhomogenität (Levene Test) wird über die Auswahl von "nicht transformiert" bei "Streubreite vs. Mittleres Niveau mit Levene Test" in der Explorativen Datenanalyse durch geführt (aber auch automatisch beim Durchführen eines T-Tests). Analysieren  $\rightarrow$  Mittelwerte vergleichen  $\rightarrow$  T-Test bei unabhängigen Stichproben. Vaso ist die Gruppenvariable, vol und rate die Testvariablen.

#### Man erhält folgende Ausgabe:

|                                      | Test bei unabhängigen Stichproben |       |                                     |       |        |                 |            |                 |                       |                         |
|--------------------------------------|-----------------------------------|-------|-------------------------------------|-------|--------|-----------------|------------|-----------------|-----------------------|-------------------------|
| Levene-Test der<br>Varianzgleichheit |                                   |       | T-Test für die Mittelwertgleichheit |       |        |                 |            |                 |                       |                         |
|                                      |                                   |       |                                     |       |        |                 | Mittlere   | Standardfehle   | 95% Konfider<br>Diffe | nzintervall der<br>renz |
|                                      |                                   | F     | Signifikanz                         | Т     | df     | Sig. (2-seitig) | Differenz  | r der Differenz | Untere                | Obere                   |
| Eingeatmete Luftmenge                | Varianzen sind gleich             | 1,565 | ,219                                | 2,704 | 37     | ,010            | ,431361142 | ,159500823      | ,108181777            | ,754540507              |
|                                      | Varianzen sind nicht<br>gleich    |       |                                     | 2,718 | 36,262 | ,010            | ,431361142 | ,158686252      | ,109611155            | ,753111129              |
| Geschwndigkeit, mit der              | Varianzen sind gleich             | 1,818 | ,186                                | 2,140 | 37     | ,039            | ,362977463 | ,169593864      | ,019347655            | ,706607271              |
| die Luft eingeatmet wurde            | Varianzen sind nicht<br>gleich    |       |                                     | 2,127 | 34,221 | ,041            | ,362977463 | ,170641216      | ,016275211            | ,709679715              |

Zuerst besagt der Levene-Test der Varianzgleichheit, dass bei "Eingeatmete Luftmenge" Varianzgleichheit angenommen werden kann (p-Wert > 0,2), wir nehmen also die Student T-Test Statistik (Wert obere Zeile), also 0,01, das heißt, es bestehen signifikante Unterschiede zwischen den Vasokonstriktions-Gruppen. Bei der "Geschwindigkeit,…" ist der p-Wert des Levene's Test <0,2, wir gehen also nicht von homogenen Varianzen aus, nehmen daher die Welch T-Test Statistik (untere Zeile). Hier ist der p-Wert 0,041, also auch hier bestehen signifikante Unterschiede zwischen den beiden Gruppen auf dem 5% Niveau.

2) Erstellen Sie eine neue Variable "Ausland", diese soll die Summe der im Ausland bereisten Städte (Paris, Madrid) wider spiegeln.

z.B. über Transformieren -> Werte in Fällen zählen



dann noch bei der Schaltfläche "Werte definieren" die "1" angeben.

Alternativ kann man auch → Variable berechnen wählen und dann die Summe von Q2\_6 und Q2\_7 wählen.

3) Wie häufig wurden die einzelnen Städte bereist?
 Unelegante Lösung: Analysieren → Deskriptive Statistiken →
 Häufigkeiten, Auswahl der Variablen Q2\_1 bis Q2\_7
 Häufigkeitstabelle

|        | Kassel |            |         |                     |                        |  |  |  |
|--------|--------|------------|---------|---------------------|------------------------|--|--|--|
|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |  |  |  |
| Gültig | nein   | 12         | 60,0    | 60,0                | 60,0                   |  |  |  |
|        | ja     | 8          | 40,0    | 40,0                | 100,0                  |  |  |  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |  |  |  |

| На | nno | ver |
|----|-----|-----|
|    |     |     |

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 5          | 25,0    | 25,0                | 25,0                   |
|        | ja     | 15         | 75,0    | 75,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 1          | 5,0     | 5,0                 | 5,0                    |
|        | ja     | 19         | 95,0    | 95,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

#### München

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 9          | 45,0    | 45,0                | 45,0                   |
|        | ja     | 11         | 55,0    | 55,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

Berlin

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 2          | 10,0    | 10,0                | 10,0                   |
|        | ja     | 18         | 90,0    | 90,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

Paris

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 8          | 40,0    | 40,0                | 40,0                   |
|        | ja     | 12         | 60,0    | 60,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

Madrid

|        |        | Häufigkeit | Prozent | Gültige<br>Prozente | Kumulierte<br>Prozente |
|--------|--------|------------|---------|---------------------|------------------------|
| Gültig | nein   | 12         | 60,0    | 60,0                | 60,0                   |
|        | ja     | 8          | 40,0    | 40,0                | 100,0                  |
|        | Gesamt | 20         | 100,0   | 100,0               |                        |

#### 3ff) elegante Lösung: Analysieren → Mehrfachantworten → Variablensets definieren

| <b>t</b> a                                                                                                                                                                                                                                                                                                             | Mehrfachantwortsets                                                                                                                                                                                                                                                                                                              |                                        | ×                                    |  |  |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|--------------------------------------|--|--|
| Setdefinition –<br>monatlich<br>lch reise<br>lch mag S<br>Alleine für<br>Ich mag f<br>lch trinke<br>lch lese g<br>Größe in<br>Größe in<br>Variablen coo<br>Dichotomi<br>Kategorien                                                                                                                                     | Variablen im Set:         gerne [         Sport [Q         hle ich         neisses         regelm         ierne [Q $minchen [Q2_2]$ $minchen [Q2_3]$ $minchen [Q2_4]$ $minchen [Q2_5]$ $minchen [Q2_6]$ $minchen [Q2_6]$ $minchen [Q2_7]$ diert als         en Gezählter Wert: $minchen [Dis [Dis [Dis [Dis [Dis [Dis [Dis [Dis$ | M<br>Hinzufügen<br>Ändern<br>Entfernen | ehrfachantwort <u>s</u> ets:<br>\$Q2 |  |  |
|                                                                                                                                                                                                                                                                                                                        | haraista Städta                                                                                                                                                                                                                                                                                                                  |                                        |                                      |  |  |
| Anmerkung: Die hier definierten Sets sind nur in den Prozeduren "Mehrfachantworten: Häufigkeiten"<br>und "Mehrfachantworten: Kreuztabellen" verfügbar. Verwenden Sie "Mehrfachantwortsets definieren" im<br>Menü "Daten", um Sets zu definieren, die in anderen Prozeduren verwendet werden können.<br>Schließen Hilfe |                                                                                                                                                                                                                                                                                                                                  |                                        |                                      |  |  |

#### 3ff) Dann Analysieren → Mehrfachantworten → Häufigkeiten

Das Set Q2, auswählen, also in den rechten Kasten schieben. Ergebnis:

|                              |          | Antworten |         | Prozent der |
|------------------------------|----------|-----------|---------|-------------|
|                              |          | Ν         | Prozent | Fälle       |
| bereiste Städte <sup>a</sup> | Kassel   | 19        | 10,6%   | 47,5%       |
|                              | Hannover | 30        | 16,7%   | 75,0%       |
|                              | Hamburg  | 33        | 18,3%   | 82,5%       |
|                              | München  | 25        | 13,9%   | 62,5%       |
|                              | Berlin   | 30        | 16,7%   | 75,0%       |
|                              | Paris    | 24        | 13,3%   | 60,0%       |
|                              | Madrid   | 19        | 10,6%   | 47,5%       |
| Gesamt                       |          | 180       | 100,0%  | 450,0%      |

#### Häufigkeiten von \$Q2

a. Dichotomie-Gruppe tabellarisch dargestellt bei Wert 1.

4) Erstellen Sie eine neue Variable "fitness". Diese soll bei allen vorab "2" betragen. Wenn Q3\_2 mit 1 oder 2 beantwortet wurde erfolgt eine Aufwertung auf "3", wenn Q3\_5 mit 1 oder 2 beantwortet wurde erfolgt eine Abwertung um 1 des vorherigen Wertes

Dies geht über **Umkodieren in andere Variablen**. Ausgangsvariable ist Q3\_2, die neue Variable soll "fitness" heißen. Die alten und neuen Werte definiert man folgendermaßen (man setzt hier quasi in einem

| 🔚 Umcodieren in andere Variablen: Alte und neue Werte                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                       |  |  |  |  |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|--|--|--|
| Alter Wert          O Wert:         O Systemdefiniert fehlend         O System- oder benutzerdefiniert fehlende Werte         Image: Bereich:         Image: Bereich: | Neuer Wert<br>Wert:<br>Systemdefiniert fehlend<br>Alte Werte kogieren<br>Alt> Neu:<br>1 thru 2> 3<br>3 thru 5> 2<br>Hinzufügen<br>Ändern<br>Entfernen |  |  |  |  |  |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Ausgabe der Variablen als Zeichenfolgen Breite: 8                                                                                                     |  |  |  |  |  |
| Weiter Abbrechen Hilfe                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                       |  |  |  |  |  |

Schritt die Werte auf 2 und führt die Aufwertung auf 3 durch, wenn Q3\_2 mit oder 2 beantwortet wurde).

4ff) Nach Erzeugung der Variable fitness führt man die Abwertung durch: **Umkodieren in dieselben Variablen**, Ausgangsvariable ist fitness. Unter "Alte und neue Werte" gibt man an,

dass aus einer 2 eine 1 und aus einer 3 eine 2 werden soll, "Falls…" Q3\_5 mit 1 oder 2 beantwortet wurde.

Man könnte dies auch über die Syntax lösen, s. nächste Seite



4ff) \*Erzeugen der Variable fitness, alle Werte erhalten eine 2. COMPUTE fitness=2.

\*Aufwertung durch Frage Q3\_2.

IF (Q3\_2 = 1 | Q3\_2=2) fitness= fitness+1.

\*Abwertung durch Frage Q3\_5.

```
IF (Q3_5=1|Q3_5=2) fitness=fitness-1.
EXECUTE.
```

5) *Einfügen der Antwortmöglichkeit "weiß nicht" (6)*: In der Variablenansicht auf die Spalte *"*Werte" auf die Zeile *"*Q3\_1" klicken, hier den entsprechenden Wert und die Beschriftung einfügen.

Errechnen des Durchschnittswertes: Den Durchschnittswert ("Mittelwert") errechnet man z.B. über Analysieren → Deskriptive Statistiken → Deskriptive Statistik

Zu beachten ist allerdings, dass die Antworten mit einer "6" (weiß nicht) in diese Berechnung mit eingehen, was nicht sinnvoll ist, da ja nur die Skala 1-5 eine Bewertung an sich ist. Um das zu vermeiden, muss man in der Variablenansicht unter "Fehlend" die "6" als fehlenden Wert definieren! Möchte man im folgenden *Balkendiagramm* für Q3\_1 den Balken für die "6" mit angezeigt bekommen, muss man dieses wieder rückgängig machen!

#### 6) Aufsummieren der nicht beantworteten Fragen: Transformieren → Werte in Fällen zählen



Dann über die Schaltfläche "Werte definieren" als zu zählende Werte die "6" als einzelnen fehlenden Wert und "systemdefiniert fehlend" (SYSMIS) angeben.

7) Kreuztabelle zu der Einkommensgruppe und der Anzahl der im Ausland besuchten Städte; Trend/Zusammenhang?
Analysieren → Deskriptive Statistiken → Kreuztabellen
Zeile "Q1", Spalte "Ausland", Häkchen setzen bei "Statistiken" → ChiQuadrat und bei "Zellen" → Erwartet.

|                               |                  |                  | im Ausland bereiste Städte |      |      |        |
|-------------------------------|------------------|------------------|----------------------------|------|------|--------|
|                               |                  |                  | 0                          | 1    | 2    | Gesamt |
| monatliches<br>Nettoeinkommen | weniger als 800€ | Anzahl           | 2                          | 2    | 0    | 4      |
|                               |                  | Erwartete Anzahl | 8,                         | 1,9  | 1,2  | 4,0    |
|                               | 800 bis 1200 €   | Anzahl           | 5                          | 2    | 0    | 7      |
|                               |                  | Erwartete Anzahl | 1,4                        | 3,4  | 2,2  | 7,0    |
|                               | 1200 bis 1500    | Anzahl           | 1                          | 5    | 5    | 11     |
|                               |                  | Erwartete Anzahl | 2,3                        | 5,4  | 3,4  | 11,0   |
|                               | 1500 bis 2000    | Anzahl           | 0                          | 5    | 2    | 7      |
|                               |                  | Erwartete Anzahl | 1,4                        | 3,4  | 2,2  | 7,0    |
|                               | mehr als 2000    | Anzahl           | 0                          | 5    | 5    | 10     |
|                               |                  | Erwartete Anzahl | 2,1                        | 4,9  | 3,1  | 10,0   |
| Gesamt                        |                  | Anzahl           | 8                          | 19   | 12   | 39     |
|                               |                  | Erwartete Anzahl | 8,0                        | 19,0 | 12,0 | 39,0   |

#### monatliches Nettoeinkommen \* im Ausland bereiste Städte Kreuztabelle

7ff) Der Chi-Quadrat-Test zeigt einen signifikanten Zusammenhang zwischen der Einkommensgruppe und der im Ausland besuchten Städte an (p<0,05).

|                                    | Wert                | df | Asymptotisch<br>e Signifikanz<br>(zweiseitig) |
|------------------------------------|---------------------|----|-----------------------------------------------|
| Chi-Quadrat nach<br>Pearson        | 21,450 <sup>a</sup> | 8  | ,006                                          |
| Likelihood-Quotient                | 24,236              | 8  | ,002                                          |
| Zusammenhang linear-<br>mit-linear | 11,433              | 1  | ,001                                          |
| Anzahl der gültigen Fälle          | 39                  |    |                                               |

a. 14 Zellen (93,3%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist ,82.

Allerdings haben quasi alle Zellen eine erwartete Häufigkeit <5 und man sollte dieses Ergebnis nicht verwenden. Die Zusammenfassung von Einkommensgruppen könnte hier helfen.

1) Die Vermutung liegt nahe, dass die Körpergröße (gross) eng mit der Vitalkapazität (fvc) zusammenhängt. Zeigen Sie dies durch eine Korrelationsanalyse.

Bei vorherigen Analysen hatten wir festgestellt, dass die Werte der Vitalkapazität nicht normalverteilt sind. Man würde daher eher auf einen nicht-parametrischen Korrelationskoeffizienten

zurück greifen, z.B. Spearman. Analysieren → Korrelation → bivariat

Dies ergibt einen Korrelationskoeffizienten von 0,941, der auf dem 0,1%-Niveau signifikant ist. Es liegt also eine sehr starke, positive Korrelation vor.



#### Lineare Regression von gross und fvc Analysieren $\rightarrow$ Regression $\rightarrow$ Linear

| t.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Lineare Regression                                                                                                                                                                                                                                                                           | ×                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| <ul> <li>✓ nr</li> <li>✓ Umweltbelastun</li> <li>✓ aller</li> <li>✓ kehle</li> <li>✓ raumu</li> <li>✓ rauva</li> <li>✓ Bildungsstand El</li> <li>✓ Schnu</li> <li>✓ huste</li> <li>✓ gebtg</li> <li>✓ gebta</li> <li>✓ Gebta</li></ul> | Abhängige Variable:<br>Abhängige Variable:<br>Vorherige<br>Unabhängige Variable(n):<br>Vorherige<br>Unabhängige Variable(n):<br>Methode:<br>Einschluss<br>Auswahlvariable:<br>Bedingung<br>Fallbeschriftungen:<br>WLS-Gewichtung:<br>WLS-Gewichtung:<br>Einfügen Zurücksetzen Abbrechen Hill | Statistiken<br>Diagramme<br>Sp <u>e</u> ichern<br>Optionen<br>Sti <u>I</u><br>Boot <u>s</u> trap |

Unter "Diagramme" noch das Residuendiagramm (ZPRED gegen ZRESID) sowie ein Normalverteilungsdiagramm der standardisierten Residuen anfertigen lassen und unter "Statistiken" den Durbin-Watson Test auswählen.

Man erhält ein R<sup>2</sup> von 0,862, der Wert der Durbin-Watson Statistik von 1,853 weist nicht auf eine Autokorrelation der Daten hin.

Das Modell lautet fvc=-0,51\*gross - 4,675

Das P-P-Diagramm zeigt, dass die Residuen relativ normal verteilt sind.

Das Streudiagramm der standardisierten Residuen gegen die standardisierten geschätzten Werte zeigt allerdings wieder eine gekrümmte Punktwolke.

Genau genommen empfiehlt es sich daher, es mit einer Linearisierung der Variable fvc zu versuchen.

Transformieren → Variable berechnen, Verwendung von z.B. lg10 als Funktion. Die Verwendung der Variable log\_fvc in der Regression führt zu einem R<sup>2</sup> von 0,885 unter Erfüllung der Voraussetzung an die Residuen.