

Kapitel 6

Häufigkeitsauszählungen

Ein erster Schritt bei der Analyse von Daten ist in der Regel eine Häufigkeitsauszählung. Wir wollen im folgenden eine Häufigkeitsauszählung am Beispiel der Datei *studium.sav* durchführen. Die Datei *studium.sav* befindet sich auf Ihrer Übungsdiskette bzw. in Ihrem Übungsverzeichnis *\SPSSBUCH*. Die Datei enthält die Ergebnisse einer Befragung zur psycho-sozialen Lage von Studierenden. StudentInnen wurden danach befragt, wie Sie ihre soziale Lage, ihre psychische Situation sowie Ihr Studium beurteilen. Erhoben wurden ferner die biografischen Daten Studienfach, Geschlecht, Alter und Staatsangehörigkeit.

6.1 Häufigkeitstabellen

- Laden Sie zunächst die Datei *studium.sav*; wählen Sie hierfür aus den Menüs

Datei
Öffnen...

Es erscheint die Dialogbox *Datei öffnen*.

- Wählen Sie die genannte Beispieldatei *studium.sav* aus, und bestätigen Sie Ihre Auswahl mit *Öffnen*. Die Datei erscheint im Daten-Editor.
- Wählen Sie aus den Menüs

Analysieren
Deskriptive Statistiken
Häufigkeiten...

Es erscheint die Dialogbox *Häufigkeiten* (siehe Bild 6.1).

- Übertragen Sie die Variable *psyche* mit Hilfe der Transport-Schaltfläche in die Zielvariablenliste, und bestätigen Sie mit *OK*.

Die Ergebnisse erscheinen im Viewer. Vor Ausgabe der eigentlichen Häufigkeitstabelle erscheint eine Übersicht über die gültigen und fehlenden Fälle; diese wird hier nicht abgebildet.

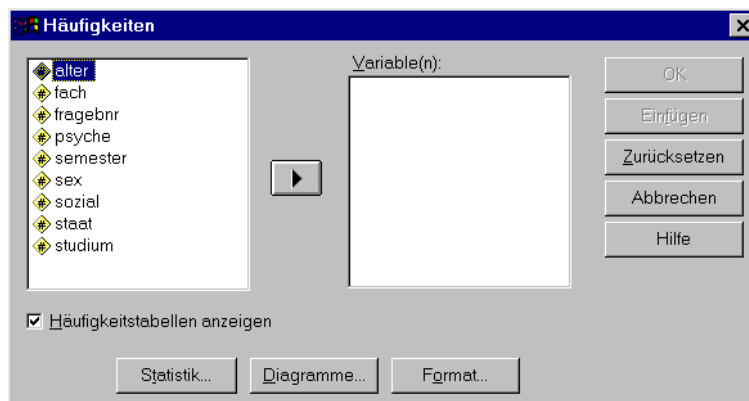


Bild 6.1: Dialogbox Häufigkeiten

Psychische Lage					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Äußerst labil	20	18,5	18,7	18,7
	Labil	40	37,0	37,4	56,1
	Stabil	41	38,0	38,3	94,4
	Sehr stabil	6	5,6	5,6	100,0
	Gesamt	107	99,1	100,0	
Fehlend	keine Angabe	1	,9		
Gesamt		108	100,0		

Jede Zeile der Häufigkeitstabelle beschreibt eine Merkmalsausprägung. Die Zeile mit dem Label »keine Angabe« repräsentiert diejenigen Fälle, bei denen keine Antwort vorlag. Insgesamt liegen 107 gültige Antworten vor, sowie ein Fall, bei dem die psychische Lage nicht bekannt ist (fehlend). Die erste Spalte enthält die Labels der einzelnen Merkmalsausprägungen (»äußerst labil«, »labil«, »stabil«, ...). In der zweiten Spalte erscheint unter »Häufigkeit« die Anzahl der jeweiligen Nennungen. So haben z.B. 20 Personen auf die Frage nach ihrer psychischen Befindlichkeit mit »äußerst labil« geantwortet, 40 Personen mit »labil«. In der dritten Spalte befindet sich die prozentuale Häufigkeit (Prozent) jeder Merkmalsausprägung. Die prozentuale Häufigkeit bezieht sich auf alle abgegebenen Antworten, d.h. inklusive der fehlenden Werte. In der vierten Spalte erscheint der gültige Prozentsatz (Gültige Prozente). Der gültige Prozentsatz klammert bei der prozentualen Berechnung die fehlenden Werte aus. Die letzte Spalte enthält den kumulativen Prozentsatz (Kumulierte Prozente). Der kumulative Prozentsatz summiert zeilenweise die prozentuale Häufigkeit der gültigen Antworten auf. So beträgt z.B. der Prozentsatz derjenigen, die »äußerst labil« oder »labil« geantwortet haben, 56,1%. Die Zahl ergibt sich aus der Berechnung: $18,7\% + 37,4\% = 56,1\%$. In der letzten Zeile steht die Gesamtsumme der einzelnen Spalten (Gesamt).

6.2 Ausgabe statistischer Kennwerte

Um deskriptive Statistiken für numerische Variablen zu erhalten, können Sie in der Dialogbox *Häufigkeiten* auf den Schalter *Statistik...* klicken. Es öffnet sich die Dialogbox *Häufigkeiten: Statistik*.

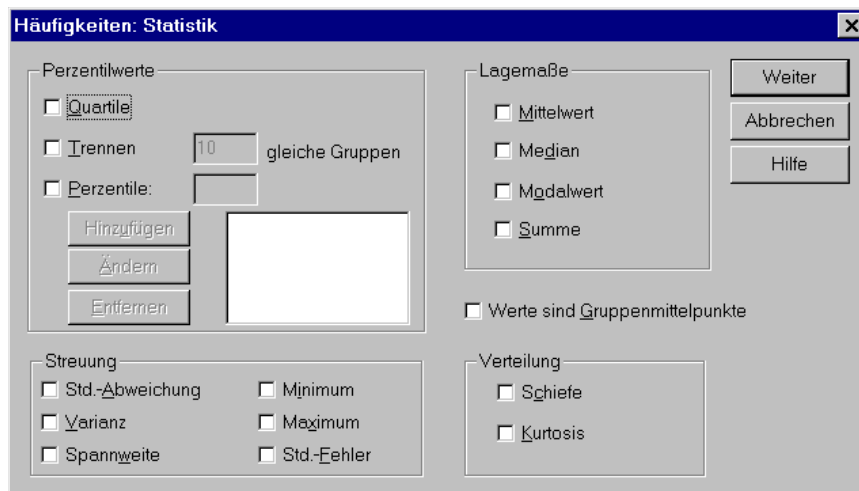


Bild 6.2: Dialogbox *Häufigkeiten: Statistik*

Im Auswahlkasten *Perzentilwerte* können Sie unter folgenden Optionen wählen:

- **Quartile:** Zeigt das 1., 2. und 3. Quartil an. Das 1. Quartil (Q_1) ist derjenige Punkt der Meßwertskala, unterhalb dessen 25% der Meßwerte liegen. Das 2. Quartil (Q_2) ist derjenige Punkt der Meßwertskala, unterhalb dessen 50% der Meßwerte liegen. Das 2. Quartil bezeichnet man auch als Median. Das 3. Quartil ist derjenige Punkt der Meßwertskala, unterhalb dessen 75% der Meßwerte liegen. Liegen die Daten nur in Form einer Rangordnung vor, wird der mittlere Quartilabstand als Streuungsmaß benutzt. Er ist definiert als

$$Q = \frac{Q_3 - Q_1}{2}$$

- **Trennen:** Erzeugt Perzentilwerte, welche die Stichprobe in Fallgruppen gleicher Klassenbreite unterteilen. Die voreingestellte Anzahl der Gruppen ist 10. Tragen Sie z.B. eine 4 ein, werden die Quartile, d.h. das 25., 50. und 75. Perzentil, angezeigt. Die Anzahl der angezeigten Perzentile ist somit um eins kleiner als die angegebene Zahl der Gruppen.

- *Perzentile*: Hierbei handelt es sich um benutzerdefinierte Perzentilwerte. Geben Sie einen Perzentilwert zwischen 0 und 100 ein, und klicken Sie anschließend auf *Hinzufügen*. Wiederholen Sie den Vorgang für jeden gewünschten Perzentilwert. Die Werte erscheinen in sortierter Folge in der Perzentilliste. Tragen Sie z.B. die Werte 25, 50 und 75 ein, so erhalten Sie Quartile. Sie können beliebige Perzentilwerte eingeben, z.B. 37 und 83. Im ersten Fall (37) wird der Wert der gewählten Variablen angezeigt, unterhalb dessen 37% der Werte liegen, im zweiten Fall (83) ist es der Wert, unterhalb dessen sich 83% der Werte befinden.

Im Auswahlkasten *Streuung* können Sie optional zwischen folgenden Streuungsmaßen wählen:

- *Standardabweichung*: Die Standardabweichung ist ein Maß für die Streuung der Meßwerte; sie ist die Quadratwurzel aus der Varianz. Trägt man die Standardabweichung zu beiden Seiten des Mittelwertes auf, so liegen bei normalverteilten Werten ca. 67% der Werte in diesem Intervall.
- *Varianz*: Die Varianz ist das Quadrat der Standardabweichung und somit ebenfalls ein Maß für die Streuung der Meßwerte. Sie wird berechnet aus der Summe der Abweichungsquadrate aller Meßwerte von ihrem arithmetischen Mittel, dividiert durch die um 1 verminderte Anzahl der Werte.
- *Spannweite*: Die Spannweite ist die Differenz zwischen dem größten Wert (Maximum) und dem kleinsten Wert (Minimum).
- *Minimum*: Der kleinste Wert.
- *Maximum*: Der größte Wert.
- *Standardfehler*: Es handelt sich hierbei um den Standardfehler des Mittelwertes. Trägt man den Standardfehler zu beiden Seiten des Mittelwertes auf, liegt mit etwa 67%-iger Wahrscheinlichkeit der Mittelwert der Grundgesamtheit in diesem Intervall. Der Standardfehler errechnet sich aus der Standardabweichung dividiert durch die Quadratwurzel des Stichprobenumfangs.

Die üblichen Streuungsmaße für intervallskalierte und normalverteilte Variablen sind die Standardabweichung und der Standardfehler. Die Standardabweichung dient, wie schon erwähnt, dazu, einen Streubereich für die einzelnen Werte anzugeben. Nach einer Faustregel liegen im einfachen Bereich der Standardabweichung (jeweils einmal links und rechts vom Mittelwert die Standardabweichung abgetragen) etwa 67 % der Werte, im doppelten Bereich der Standardabweichung etwa 95 % und im dreifachen Bereich etwa 99 % der Werte.

Der Standardfehler hingegen dient dazu, ein Konfidenzintervall für den Mittelwert anzugeben. Tragen Sie den doppelten Standardfehler links und rechts vom Mittelwert ab, so liegt mit etwa 95%-iger Wahrscheinlichkeit der Mittelwert der Grundgesamt in diesem Bereich. Mit etwa 99%-iger Wahrscheinlichkeit liegt er im dreifachen Bereich des Standardfehlers. Meist gibt man nur eines der beiden genannten Streuungsmaße an, wobei man sich gerne für den Standardfehler entscheidet, weil dieser kleinere Werte liefert. Auf alle Fälle sollten Sie deutlich machen, für welches der beiden Streuungsmaße Sie sich entschieden haben.

Im Auswahlkasten *Lagemaße* können Sie unter folgenden Begriffen wählen:

- **Mittelwert:** Der Mittelwert ist das arithmetische Mittel der Meßwerte und berechnet sich daher aus der Summe der Meßwerte geteilt durch ihre Anzahl. Liegen z.B. zwölf Meßwerte vor und beträgt die Summe der Meßwerte 600, so ist der Mittelwert $x = 600 : 12 = 50$.
- **Median:** Der Median ist derjenige Punkt der Meßwertskala unterhalb und oberhalb dessen jeweils die Hälfte der Meßwerte liegen. Liegen die Meßwerte einzeln vor, z.B.

3 7 8 5 4 6 3 9 2 8 4,

so schreibt man diese zunächst der Größe nach sortiert auf:

2 3 3 4 4 5 6 7 8 8 9.

In diesem Fall ist der Meßwert 5 der Median. Es liegen insgesamt $n = 11$ Meßwerte vor, so daß der 6. Meßwert der Median ist. Es liegen nämlich dann 5 Meßwerte unterhalb und 5 Meßwerte oberhalb dieses Wertes. Bei ungeradem n ist der Median also ein tatsächlich auftretender Meßwert. Bei geradem n ist der Median das arithmetische Mittel zweier benachbarter Meßwerte. Liegen z.B. die Meßwerte

3 4 4 5 6 7 8 8 9 9

vor, so ist der Median in diesem Falle: $(6 + 7) : 2 = 6,5$.

- **Modalwert:** Der Modalwert ist der am häufigsten auftretende Wert in einer Stichprobe. Weisen mehrere Werte dieselbe maximale Häufigkeit auf, so wird nur der kleinste Wert angezeigt.
- **Summe:** Die Summe aller Werte.

Im Auswahlkasten *Verteilung* können Sie zwischen den beiden folgenden Schiefeitsmaßen wählen:

- **Schiefte:** Die Schiefe ist eine Bezeichnung für die Abweichung einer Häufigkeitsverteilung von einer symmetrischen Verteilung, also einer Verteilung, bei der innerhalb gleicher Abstände vom Mittelwert auf beiden Seiten jeweils gleich viele Werte liegen. Die Schiefe ist Null, wenn die beobachtete Verteilung eine Normalverteilung ist. Diese Aussage kann zum Test auf Normalverteilung benutzt werden: Ist die Schiefe signifikant von Null verschieden, ist die Hypothese, daß die Daten aus einer normalverteil-

ten Grundgesamtheit stammen, zu verwerfen. Liegt die Spitze einer schiefen Verteilung bei den kleineren Meßwerten, so spricht man von positiver, im anderen Fall von negativer Schiefe.

- *Kurtosis (Exzeß)*: Der Exzeß gibt an, ob eine Verteilung breitgipflig (hoher Wert) oder schmalgipflig ist. Der Exzeß ist Null, wenn die beobachtete Verteilung eine Normalverteilung ist. Diese Aussage kann ebenfalls zum Testen auf Normalverteilung benutzt werden: Ist die Kurtosis signifikant von Null verschieden, ist die Hypothese, daß die Daten aus einer normalverteilten Grundgesamtheit stammen, zu verwerfen.

Im allgemeinen benutzt man bei intervallskalierten und normalverteilten Variablen als Lagemaß den Mittelwert und als Streuungsmaß die Standardabweichung oder den Standardfehler und bei ordinalskalierten oder nicht normalverteilten Variablen als Lagemaß den Median und als Streuungsmaß das 1. und 3. Quartil. Bei nominalskalierten Variablen ist außer dem Modalwert kein sinnvoller Kennwert anzugeben.

Ferner gibt es noch folgende Optionen:

- *Werte sind Gruppenmittelpunkte*: Kreuzen Sie diese Option an, werden bei der Berechnung des Medians und der übrigen Perzentilwerte die Abschätzungen dieser Kennwerte für gehäufte Daten berechnet. Dieser Problematik ist ein eigener Abschnitt gewidmet.

Wir wollen für die Variable *alter* folgende Kennwerte ausgeben: Mittelwert, Median, Modalwert, Quartile, Standardabweichung, Varianz, Spannweite, Minimum, Maximum, Standardfehler, Schiefe und Kurtosis. Gehen Sie wie folgt vor:

- Wählen Sie aus den Menüs

Analysieren
Deskriptive Statistiken
Häufigkeiten...

- Klicken Sie in der Dialogbox *Häufigkeiten* zunächst auf *Zurücksetzen*, um ältere Einstellungen zu löschen.
- Übertragen Sie die Variable *alter* in die Zielvariablenliste.
- Klicken Sie auf die Schaltfläche *Statistik...*
- Aktivieren Sie in der Dialogbox *Häufigkeiten: Statistik* die gewünschten Kennwerte. Klicken Sie anschließend auf *Weiter*. Sie gelangen zurück zur Dialogbox *Häufigkeiten*.
- Deaktivieren Sie in der Dialogbox *Häufigkeiten* die Option *Häufigkeitstabellen anzeigen*. Bestätigen Sie mit *OK*.

Sie sehen folgende Ergebnisse im Viewer:

Statistiken		
Alter		
N	Gültig	106
	Fehlend	2
Mittelwert		22,24
Standardfehler des Mittelwertes		,21
Median		22,00
Modus		21
Standardabweichung		2,19
Varianz		4,79
Schiefe		,859
Standardfehler der Schiefe		,235
Kurtosis		1,042
Standardfehler der Kurtosis		,465
Spannweite		11
Minimum		18
Maximum		29
Perzentile	25	21,00
	50	22,00
	75	23,00

Die Probanden der Studie zur psycho-sozialen Lage von Studierenden sind im Mittel 22,24 Jahre alt. Der Median beträgt 22. Die meisten der Befragten sind 21 Jahre alt (Modus). Der jüngste Befragte ist 18 Jahre alt (Minimum), der älteste 29 Jahre (Maximum). Der älteste Befragte ist 11 Jahre älter als der jüngste Befragte (Spannweite). Die Standardabweichung beträgt 2,19. Die Varianz als Quadrat der Standardabweichung ist demzufolge $(2,19)^2 = 4,79$. Schiefe und Kurtosis werden jeweils zusammen mit ihrem Standardfehler ausgegeben.

6.3 Median bei gehäuften Daten

Insbesondere bei Daten, die in Form einer Häufigkeitstabelle vorliegen, ist die Bestimmung des Medians und beliebiger anderer Perzentile allein nach der herkömmlichen Abzählmethode zu ungenau. In solchen Fällen besteht die Möglichkeit, den Median und andere beliebige Perzentile nach einer genaueren Methode zu berechnen. Dies soll anhand eines Beispiels aus der Zahnmedizin erläutert werden.

- Laden Sie hierzu die Datei cpitn.sav, welche die Ergebnisse einer zahnmedizinischen Untersuchung enthält.

Neben den beiden Variablen schule und mhfreq, die den Schulabschluß bzw. die Häufigkeit des täglichen Zähneputzens wiedergeben, sind die sechs Variablen cpitn1 bis cpitn6 enthalten, die für das in Sextanten eingeteilte Gebiß den Grad der parodontalen Erkrankung, den sogenannten CPITN-Wert, angeben, und zwar nach folgendem Code:

0	gesundes Parodont
1	Blutung
2	Zahnstein
3	Taschentiefe von 3,5 bis 5,5 mm
4	Taschentiefe von 6 mm und mehr

■ Erstellen Sie mit Hilfe der Menüwahl

Analysieren

Deskriptive Statistiken

Häufigkeiten...

eine Häufigkeitstabelle etwa der Variablen cpitn1, und wünschen Sie sich die Berechnung von Mittelwert und Median, so erhalten Sie folgendes Ergebnis:

Statistiken

CPITN1

N	Gültig	2548
	Fehlend	0
Mittelwert		2,24
Median		2,00

CPITN1

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig gesund	109	4,3	4,3	4,3
Blutung	389	15,3	15,3	19,5
Zahnstein	921	36,1	36,1	55,7
Taschentiefe 3,5-5,5	1042	40,9	40,9	96,6
Taschentiefe >=6	87	3,4	3,4	100,0
Gesamt	2548	100,0	100,0	

Als Median wird der Wert 2 ausgewiesen, was nach dem herkömmlichen Abzählverfahren zwar formal in Ordnung ist, aber natürlich einen völlig unbefriedigenden, weil nicht genügend differenzierenden Wert ergibt. Für einen solchen Fall gehäufte Daten gibt es für den Median eine entsprechende Abschätzformel:

$$\text{Median} = u + \frac{b}{f_m} \cdot \left(\frac{n}{2} - F_{m-1} \right)$$

Dabei bedeuten:

n	Anzahl der Meßwerte
m	Klasse, in welcher der Median liegt
u	untere Grenze der Klasse m

f_m	absolute Häufigkeit in der Klasse m
F_{m-1}	kumulative Häufigkeit bis zur vorangehenden Klasse m-1
b	Klassenbreite

Von entscheidender Wichtigkeit ist demnach die korrekte Wahl der Klassengrenzen; diese sind so zu wählen, daß die auftretenden Codezahlen jeweils die Klassenmitte bilden. Als Klassengrenzen im gegebenen Beispiel sind daher der Reihe nach

-0,5 0,5 1,5 2,5 3,5 4,5

zu wählen. Die Klassenbreite ist 1.

Damit ergibt sich im vorliegenden Beispiel:

$$\begin{aligned} n &= 2548 \\ m &= 3 \text{ (da der Median in der 3. Klasse liegt)} \\ u &= 1,5 \\ f_m &= 921 \\ F_{m-1} &= 109 + 389 = 498 \\ b &= 1 \end{aligned}$$

$$\text{Median} = n + \frac{1}{2} \cdot \left(\frac{2548 - 498}{921} \right) = 498 + 2,32$$

Vergleicht man diesen Wert mit dem Mittelwert (2,24), so wird die Regel bestätigt, daß bei rechtsgipfligen Verteilungen (und um eine solche handelt es sich hier) der Median größer als der Mittelwert ist.

Diese genaue Berechnung des Medians erhalten Sie, wenn Sie in der Dialogbox *Häufigkeiten: Statistik* neben der Berechnung des Medians die Option *Werte sind Gruppenmittelpunkte* aktivieren.

Sie erhalten jetzt den korrekten Wert des Medians angezeigt (2,32).

Der Median ist definitionsgemäß derjenige Wert, unterhalb und oberhalb dessen jeweils 50 % der ihrer Größe nach geordneten Werte liegen. Eine Verallgemeinerung dieser Prozentzahl führt zu den sogenannten Perzentilen. So könnte man nach dem Wert fragen, unterhalb dessen 10 % der Werte liegen (oberhalb liegen dann 90 % der Werte). Gebräuchlich sind vor allem das 25 % – und 75 % – Perzentil, auch als 1. Quartil bzw. 3. Quartil bezeichnet.

Beliebige Perzentilwerte können in der Dialogbox *Häufigkeiten: Statistik* nacheinander eingegeben werden. Beim Vorliegen gehäufte Daten ist wieder die Option *Werte sind Gruppenmittelpunkte* zu aktivieren.

Die Abschätzungsformel für ein beliebiges Perzentil lautet:

$$\text{Perzentil} = u + \frac{b}{h_m} \cdot (P - H_{m-1})$$

Dabei bedeuten:

m	Klasse, in der das Perzentil liegt
u	untere Grenze der Klasse m
P	zum Perzentil gehörige Prozentzahl
h_m	prozentuale Häufigkeit in der Klasse m
H_{m-1}	prozentuale kumulative Häufigkeit in der Klasse m-1
b	Klassenbreite

Für das 50 % – Perzentil ($P = 50$) ergibt sich nach einigen Umformungen die für den Median aufgezeigte Formel.

In Balken-, Linien-, Flächen- und Kreisdiagrammen, in denen die Darstellung von Median und anderen Perzentilen vorgesehen ist, wird dieser modifizierten Berechnung im Falle des Vorliegens gehäufelter Daten Rechnung getragen (siehe Kap. 22.1.1).

6.4 Formate für Häufigkeitstabellen

- Laden Sie die Datei *studium.sav* (siehe Kap. 6.1).

Wir wollen eine Häufigkeitstabelle für die Variable *fach* sortiert nach fallender Ordnung der Häufigkeiten ausgeben. Gehen Sie wie folgt vor:

- Wählen Sie aus den Menüs

Analysieren
Deskriptive Statistiken
Häufigkeiten...

- Übertragen Sie die Variable *fach* in die Zielvariablenliste.
- Klicken Sie auf *Format...* Es öffnet sich die Dialogbox *Häufigkeiten: Format*.

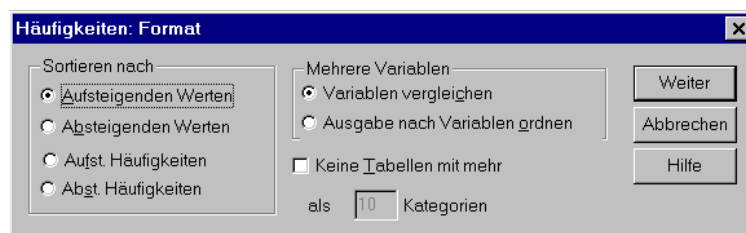


Bild 6.3: Dialogbox *Häufigkeiten: Format*

Im Auswahlkasten *Sortieren nach* können Sie die Reihenfolge, nach der die Datenwerte in einer Häufigkeitstabelle angezeigt werden sollen, bestimmen. Sie können eine der folgenden Optionen wählen:

- ▶ *Aufsteigende Werte*: Sortiert die Merkmalsausprägungen nach steigender Ordnung der Werte. Dies ist die Voreinstellung.
- ▶ *Absteigende Werte*: Sortiert die Merkmalsausprägungen nach fallender Ordnung der Werte.
- ▶ *Aufsteigende Häufigkeiten*: Sortiert die Merkmalsausprägungen nach steigender Ordnung der Häufigkeiten.
- ▶ *Absteigende Häufigkeiten*: Sortiert die Kategorien nach fallender Ordnung der Häufigkeiten.

Sie haben ferner die Gelegenheit, durch Aktivierung der Option *Keine Tabellen mit mehr als ... Kategorien* die Ausgabe langer Häufigkeitstabellen zu unterdrücken.

- Klicken Sie auf *Abst. Häufigkeiten*.
- Bestätigen Sie mit *Weiter*.
- Klicken Sie auf *OK*, um die Berechnungen durchzuführen.

Sie erhalten folgende Ausgabe:

Fachbereich		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Geistesw.	25	23,1	23,1	23,1
	Jura	22	20,4	20,4	43,5
	Wirtschaftsw.	19	17,6	17,6	61,1
	Psychologie	11	10,2	10,2	71,3
	Medizin	10	9,3	9,3	80,6
	Theologie	9	8,3	8,3	88,9
	Naturw.	9	8,3	8,3	97,2
	Ingenieurw.	2	1,9	1,9	99,1
	Sonstige	1	,9	,9	100,0
	Gesamt	108	100,0	100,0	

Die Hauptstudienfächer der Probanden werden nach fallender Ordnung der Häufigkeiten sortiert angezeigt.

6.5 Grafische Darstellung

Die Ergebnisse von Häufigkeitsverteilungen können grafisch aufbereitet werden. Als Beispiel wollen wir ein Balkendiagramm der Hauptstudienfächer erstellen. Gehen Sie wie folgt vor:

- Wählen Sie aus den Menüs

Analysieren
Deskriptive Statistiken
Häufigkeiten...

- Übertragen Sie die Variable fach in die Zielvariablenliste.
- Klicken Sie auf *Diagramme...* Es öffnet sich die Dialogbox *Häufigkeiten: Diagramme*.

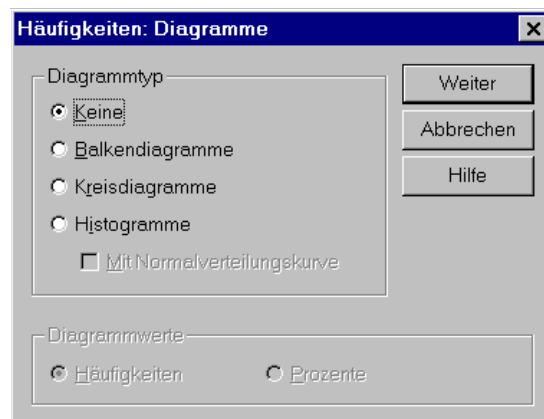


Bild 6.4: Dialogbox Häufigkeiten: Diagramme

- Aktivieren Sie *Balkendiagramme* und unter *Diagrammwerte* die Option *Prozente*. Bestätigen Sie mit *Weiter*. Sie gelangen zurück zur Dialogbox *Häufigkeiten*.
- Deaktivieren Sie in der Dialogbox *Häufigkeiten* die Option *Häufigkeitstabellen anzeigen*. Bestätigen Sie anschließend mit *OK*. Das Diagramm wird im Viewer angezeigt (siehe Bild 6.5).

Wir wollen dieses Balkendiagramm noch etwas verschönern.

- Klicken Sie doppelt auf den Bereich des Balkendiagramms, um die Grafik zu bearbeiten. Das Balkendiagramm wird im Diagramm-Editor angezeigt.
- Klicken Sie in der Formatleiste des Diagramm-Editors auf das Balken-Label-Symbol:



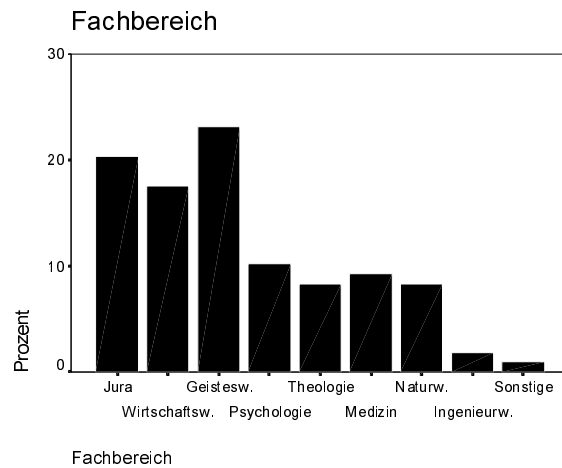


Bild 6.5: Balkendiagramm im Viewer.

Es öffnet sich die Dialogbox *Balkenbeschriftung*. Wählen Sie hier *Rahmen*, klicken Sie auf *Allen zuw.* und anschließend auf *Schließen*. Die einzelnen Balken erhalten Etiketten mit den jeweiligen Prozentwerten.

- Klicken Sie mit der Maus in die Mitte der Fläche eines beliebigen Balkens. Sie erkennen zwei kleine schwarze Vierecke an den oberen Kanten der Balken. Die Balkenflächen sind damit zum Editieren freigegeben.
- Klicken Sie mit der Maus auf das Füllmuster-Symbol:



Es öffnet sich die Dialogbox *Füllmuster*.



Bild 6.6: Dialogbox Füllmuster

- Wählen Sie hier ein beliebiges schraffiertes Füllmuster aus. Bestätigen Sie mit *Zuweisen*, und schließen Sie die Dialogbox.

Die Balken werden mit dem ausgewählten Füllmuster versehen.

- Klicken Sie mit der Maus in der Formatleiste auf das Symbol Balkenart:



- Wählen Sie hier *Schlagschatten* aus, klicken Sie auf *Allen zuw.*, und schließen Sie anschließend die Dialogbox.
- Klicken Sie doppelt auf die Überschrift »Fachbereich« des Balkendiagramms. Es öffnet sich die Dialogbox *Titel* (siehe Bild 6.7).

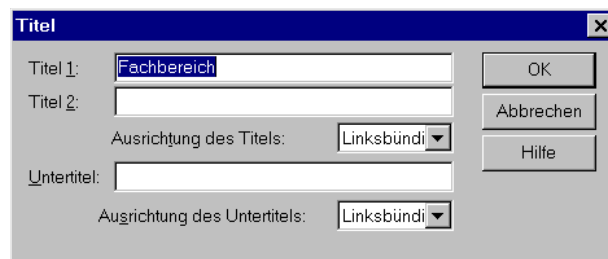


Bild 6.7: Dialogbox Titel

- Ändern Sie den Titel in »Hauptstudienfach« um, und bestätigen Sie mit *OK*.
- Aktivieren Sie abschließend im Menü *Diagramme* die Option *Äußerer Rahmen*. Schließen Sie den Diagramm-Editor; die entsprechend geänderte Grafik ist in Bild 6.8 dargestellt.

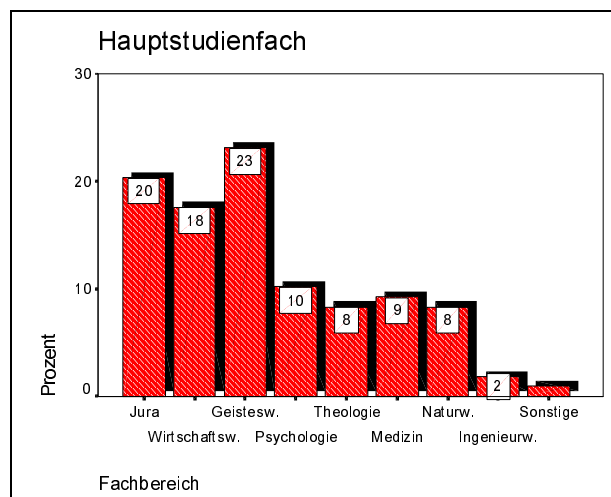


Bild 6.8: Bearbeitetes Balkendiagramm

Wir wollen ein weiteres Beispiel für die visuelle Präsentation einer Häufigkeitsauszählung betrachten.

- Wählen Sie aus den Menüs

Analysieren
Deskriptive Statistiken
Häufigkeiten...

- Klicken Sie auf *Zurücksetzen*, um die Standard-Einstellungen zu aktivieren.
- Übertragen Sie die Variable sozial in die Zielvariablenliste.
- Klicken Sie auf *Diagramme...* Wählen Sie *Histogramme* in der Dialogbox *Häufigkeiten: Diagramme*. Aktivieren Sie die Option *Mit Normalverteilungskurve*. Klicken Sie auf *Weiter*.
- Deaktivieren Sie in der Dialogbox *Häufigkeiten* die Option *Häufigkeitstabellen anzeigen*. Klicken Sie auf *OK*. Sie sehen das Histogramm im Viewer (siehe Bild 6.9).

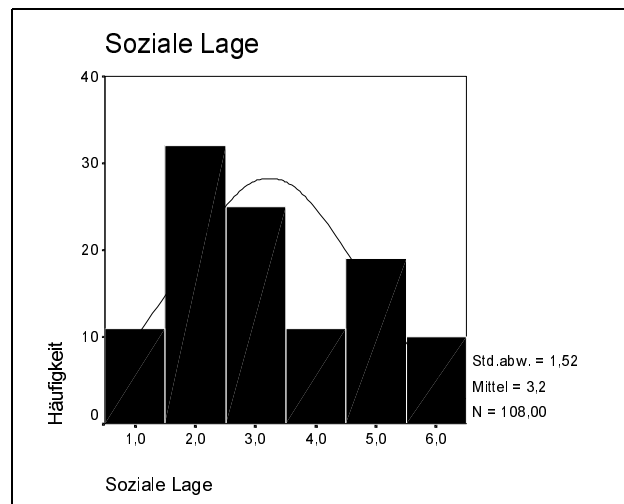


Bild 6.9: Histogramm

Die Häufigkeiten beim Histogramm werden durch Säulen gekennzeichnet, die im Unterschied zum Balkendiagramm nicht isoliert, sondern aneinandergereiht angezeigt werden. Ausgegeben werden die Standardabweichung, der Mittelwert und die zugrundeliegende Anzahl der Fälle (N). Ferner wird die Normalverteilungskurve ausgegeben.

- Klicken Sie doppelt auf den Bereich des Histogramms, um es im Diagramm-Editor Ihren Wünschen gemäß zu ändern. Die Grafik erscheint im Diagramm-Editor.
- Wählen Sie ein anderes Füllmuster aus, und versehen Sie die Balken mit Etiketten.
- Testen Sie weitere Optionen des Diagramm-Editors nach Belieben aus.

Wir beenden hiermit das Thema Häufigkeitsauszählungen. Führen Sie abschließend übungshalber eine Häufigkeitsverteilung der Variablen studium durch, und bereiten Sie die Ergebnisse der Häufigkeitsverteilung grafisch auf.