

wurde erstmals auf ein ganzes Halbssemester erstreckt. Die Daten enthalten die Prüfungsergebnisse nach der neuen Kursform.

- Lassen Sie den Schluss zu, dass der Anteil der Studierenden, die die Prüfung bestehen, bei der neuen Kursform höher ist (bisher bestanden 40% der Studierenden die Prüfung nicht)?

4. *Lottozahlen in Österreich

Datenfile: lotto0708.sav

Im Datenfile sind folgende Variablen enthalten:

```
lottozahl  Lottozahl (1-45)
h2007      Ziehungshäufigkeit 2007
h2008      Ziehungshäufigkeit 2008
```

In Österreich ist Lotto sehr beliebt. Sechs Zahlen plus Zusatzzahl werden von einer Maschine aus der Menge der Zahlen von 1 bis 45 gezogen. Natürlich erwartet man, dass jede Zahl die gleiche Wahrscheinlichkeit hat, gezogen zu werden. Im Datenfile sind die Ziehungshäufigkeiten für alle 45 Lottozahlen des österreichischen Lottos für die Jahre 2007 und 2008 gegeben.

- Kann man aus den Daten schließen, dass der Mechanismus, mit dem die Lottozahlen ermittelt werden, fair ist, d.h., dass für alle Zahlen die Auswahlwahrscheinlichkeit gleich ist?

5. *Multiple-Choice-Prüfung

Datenfile: mchoice.sav

Im Datenfile ist folgende Variable enthalten:

```
richtig  Richtige Antwort auf MC-Frage
```

Georg steht vor einer Prüfung, die als Multiple-Choice-Prüfung durchgeführt wird. Wie üblich hat Georg keine Ahnung vom Prüfungsstoff und er hofft, die Prüfung durch reines Raten zu bestehen. Allerdings hat Georg eine frühere Prüfung desselben Prüfers mit markierten richtigen Antworten erhalten.

- Verteilt der Prüfer die richtigen Antworten zufällig über alle fünf Auswahlmöglichkeiten?

6. *Notenverteilung

Datenfile: noten.sav

Im Datenfile ist folgende Variable enthalten:

```
note  Note bei einer Prüfung
```

Die Noten, die von einem BW-Professor vergeben werden, folgten bisher einer symmetrischen Verteilung: 5% Sehr gut, 25% Gut, 40% Befriedigend, 25% Genügend und 5% Nicht genügend. Dieses Jahr wird eine Stichprobe von 150 Noten gezogen.

- Kann man daraus schließen (5% Signifikanzniveau), dass sich die Notenverteilung dieses Jahr von der früherer Jahre unterscheidet?

7. ***Autoklasse und Unfallhäufigkeit** Datenfile: unfaelle.sav

Im Datenfile ist folgende Variable enthalten:

auto Klasse des Autos

Zulassungsdaten aus einem Land zeigen, dass 15% der Autos Kleinwagen, 25% Kompaktmodelle, 40% Mittelklassemodelle und der Rest größere oder Sondermodelle sind. Eine Zufallsstichprobe von Verkehrsunfällen mit Autos wird gezogen.

- Kann man schließen, dass bestimmte Größenklassen von Autos häufiger in Verkehrsunfälle verwickelt sind, als es die Zulassungszahlen vermuten lassen?

8. ***Fehlerrate bei Lügendetektoren**

Nach wie vor gibt es in den USA Bemühungen, den perfekten Lügendetektor zu entwickeln. Neuere Ansätze stammen von Pavlidis et. al. (2002), die versuchten, mit einer hochauflösenden, temperatursensiblen Kamera aus Gesichtsaufnahmen Lügen zu entdecken. Rosenfeld (2002) verwendete sogenannte ERPs (ereigniskorrelierte Potentiale), bestimmte Gehirnaktivitätssignale, die mittels an der Kopfhaut angebrachten Elektroden gemessen werden. Er untersuchte Studierende, die unter anderem Sätze, wie *Verwenden Sie einen gefälschten Ausweis?* vorlesen mussten. Es wurde erwartet, dass bei Studierenden, die tatsächlich einen gefälschten Ausweis verwenden, ein entsprechendes Hirnsignal auftritt. Von insgesamt $N = 17$ „Schuldigen“ wurden 13 (77%) richtig erkannt. Bei einer vergleichbaren Studie des amerikanischen Verteidigungsministeriums wurden 75% der „Schuldigen“ mithilfe eines traditionellen klassischen Lügendetektors (Polygraphen) richtig erkannt.

- Liefert der neue Lügendetektor bessere Ergebnisse als der traditionelle?

9. ***Konzentrationsleistung von Studierenden** Datenfile: ktest.sav

Im Datenfile ist folgende Variable enthalten:

punkte Punktezahl bei einem Konzentrationstest

Bei einem Konzentrationstest kann man 0 bis 50 Punkte erzielen. Es ist bekannt, dass 15% der Personen mehr als 40 Punkte erzielen. Der Test wurde an 200 zufällig ausgewählten Studierenden durchgeführt.

- Kann man aus den Ergebnissen schließen, dass Studierende besser abschneiden als die Gesamtbevölkerung (1% Signifikanzniveau)?

Mehrere kategoriale Variablen

1. Arbeitszufriedenheit

Eine Arbeitnehmerorganisation erhebt in einer Befragung die Zufriedenheit mit den Arbeitsbedingungen. Die Angaben für drei Arbeitnehmerklassen ist folgender Tabelle zusammen gefasst.

	Zufriedenheit	
	zufrieden	nicht zufrieden
Management	14	3
Verwaltung	26	34
Produktion	47	28

- Erstellen Sie aus den Daten der Tabelle ein Datenfile.
- Gibt es Unterschiede zwischen den drei Klassen?

2. Zufriedenheit mit öffentlichen Verkehrsmitteln

Im Einzugsgebiet einer Großstadt wurde eine Untersuchung über die Zufriedenheit mit dem Angebot an öffentlichen Verkehrsmitteln durchgeführt. Ein Teilaspekt befasste sich mit der Frage, ob Bewohner des Stadtkerns zufriedener mit dem Angebot öffentlicher Verkehrsdienstleister (kurz: Öffis) sind als Bewohner des Stadtrandes bzw. der Umlandgemeinden.

Angebot	Wohnort		
	Stadtkern	Stadtrand	Umland
sehr gut	44	32	19
eher gut	79	57	52
eher schlecht	52	63	127
sehr schlecht	45	83	206

- Erstellen Sie aus den Daten der Tabelle ein Datenfile.
- Unterscheidet sich die Zufriedenheit je nach Wohnort?

3. Titanic

Datenfile: `titanic.sav`

Im Datenfile sind folgende Variablen enthalten:

Klasse	Klasse, die ein Passagier bewohnt hat
Schicksal	Schicksal des Passagiers 1 = Gerettet, 2 = Gestorben
Anzahl	Gewichtungsvariable

Der Datensatz beschreibt, wieviel Passagiere des gesunkenen Kreuzfahrtschiffes Titanic in den jeweiligen Klassen gerettet wurden bzw. gestorben sind.

- Ist die Überlebenschance in der 1.Klasse signifikant höher als in der 2.Klasse? (Wir nehmen die Titanic-Daten als Stichprobe für alle Kreuzfahrtschiffe)

4. Prüfung

Datenfile: pruefung.sav

Im Datenfile sind folgende Variablen enthalten:

Lernpensum	Lernaufwand in Stunden
Vorlesungsbesuch	regelmäßiger Besuch der Vorlesung 0 = nein, 1 = ja
Punkte	Punktezahl beim Test
Pruefung	Bestehen der Prüfung 0 = nein, 1 = ja

- Ist das Bestehen der Prüfung vom Lernpensum und vom Vorlesungsbesuch abhängig?

5. *Reisebegleitung im Haupturlaub

In einer Stichprobe ergab die Aufteilung in Männer und Frauen, die in Urlaub fahren, folgende Tabelle:

Reisebegleitung	Frau	Mann
PartnerIn	2273	2418
Familienurlaub	1212	1023
Gruppenurlaub	960	744
Allein	454	325
Anderes	151	93

- Erstellen Sie ein Datenfile mit den Daten obiger Tabelle!
- Ist die Reisebegleitung bei Frauen und Männern unterschiedlich?

6. *Spaß am Sex bei Ehepartnern

In einer amerikanischen Untersuchung, 1987 (Quelle: Agresti, 1990, adaptiert), wurde beiden Ehepartnern unter anderem die folgende Frage gestellt: „Sex macht mir und meinem Partner Spaß (1) nie oder selten (2) manchmal (3) sehr oft oder immer“. Die folgenden Daten beschreiben die Häufigkeiten der Antworten, kreuzklassifiziert nach den Antworten der Ehefrauen und Ehemänner.

Ehemann	Ehefrau			gesamt
	selten	manchmal	oft	
selten	7	7	5	19
manchmal	2	8	10	20
oft	3	13	36	52
gesamt	12	28	51	91

- Gibt es einen Zusammenhang zwischen den Antworten der Ehepartner?

7. *Waffenregistrierung und Einstellung zu Todesstrafe

In den USA wurden im Rahmen des 1982 General Social Survey Einstellungen zu Waffenregistrierung und Todesstrafe erhoben (aus Agresti 1990, p.29).

Waffenregistrierung	Todesstrafe	
	dafür	dagegen
dafür	784	236
dagegen	311	66

- Gibt es einen Zusammenhang zwischen den Einstellungen zu diesen Themen?
- Wenn ja, wie ist der Zusammenhang?

8. *Lehrveranstaltungsbesuch

Der Besuch von Lehrveranstaltungen kostet Zeit und wird von Studierenden gern auf das Notwendigste beschränkt. Allerdings wird durch aktive geistige Präsenz im Hörsaal ein Grundstein zur Erfassung und Verständnis der Lehrinhalte gelegt. Dieses Verständnis ist in gewissen Fächern durch Selbststudium allein nur schwer zu erlangen.

In einem Kurs mit 170 Teilnehmern soll untersucht werden, ob Unterschiede im Prüfungsergebnis zwischen jenen Studierenden, die regelmäßig die Kurse besucht haben, und jenen, die nur im Selbststudium gearbeitet haben, existieren. Ein Ergebnis, bei dem nur zwischen Bestehen und Nichtbestehen der Prüfung unterschieden wird, ist in folgender Tabelle zusammen gefasst:

	Kurs + Vorbereitung	nur Selbststudium
bestanden	79	55
nicht bestanden	12	24

- Man bestimme das Odds-Ratio für das Bestehen der Prüfung bei Kursbesuch im Vergleich zu Selbststudium.
- Man bestimme ein Konfidenzintervall für dieses Odds-Ratio! Kann daraus geschlossen werden, dass die Chancen nicht gleich sind?

Eine und mehrere metrische Variablen

1. ***Alter bei Amtsantritt** Datenfile: us-president.sav

Im Datenfile ist das Alter der US-Präsidenten bei Amtsantritt angegeben.

- Beschreiben Sie den Datensatz mit Histogramm, Boxplot und Maßzahlen!

2. ***VwGH-Verfahrensdauer 1. Instanz** Datenfile: vwgh.sav

Im Datenfile ist u.a. die Länge des Verfahrens in der ersten Berufungsinstanz angegeben (`dauer2`).

- Dauern die Verfahren in der 1. Instanz länger als vor 20 Jahren, als sie im Schnitt 2 Jahre und 7 Monate dauerten?
- Ist das Ergebnis nur deshalb signifikant, weil nicht wenige Ausreißerwerte vorliegen?

3. **US-Masters 2009: Runden 1 und 2** Datenfile: augusta2009.sav

- Beschreiben Sie die Verteilung für die in den ersten zwei Runden im US-Masters in Augusta 2009 benötigten Schläge (Variablen `R1` und `R2` enthalten die zwei ersten Rundenergebnisse).
- In welchem Intervall würde man aufgrund dieser Stichprobe die notwendigen Schläge für die zwei ersten Runden annehmen?
- Ist die Anzahl Schläge für die ersten zwei Runden normalverteilt? Welches Ergebnis zeigt der Kolmogorov-Smirnov-Test? Weist der Q-Q-Plot Auffälligkeiten auf?

4. **Büromieten** Datenfile: miete.sav

Im Datenfile sind folgende Variablen enthalten:

`Leer` Anteil leer stehender Büros
`Miete` Büromiete in Euro pro Quadratmeter

Ein Ökonom untersucht die Beziehung zwischen Büromieten und Anteil leer stehender Büroflächen. Dazu nimmt er die Büromieten von zufällig ausgewählten Büros (in 30 verschiedenen Städten) und die Anteile leer stehender Büros in diesen Städten.

- Gibt es einen Zusammenhang zwischen dem Anteil nicht vermieteter Büroflächen und den Mieten?

- Wie hoch wäre eine durchschnittliche Büromiete in einer Stadt, in der 10% der Büros leer stehen?

5. Filmproduktion

Datenfile: film.sav

Im Datenfile sind folgende Variablen enthalten:

EINNAHM	Einnahmen aus einem Film (in Mio.)
PRODCOST	Produktionskosten (in Mio.)
PROMOT	Werbeausgaben für den Film (in Mio.)
BUCH	Dummyvariable für Verfilmung eines Bestsellers 0 = nein, 1 = ja

Der Erfolg eines Filmes wird oft rein ökonomisch gemessen, in Hitlisten der erfolgreichsten Filme geht es meist nur um die Einspielergebnisse. Können diese aus anderen Daten über den Film vorhergesagt werden?

Im Datenfile sind die Einspielergebnisse, die Produktionskosten und die Werbeausgaben für 25 Filme aufgelistet, zusätzlich gibt eine Variable an, ob der Stoff des Films schon als Buch bekannt war.

- Man finde eine passende Gleichung, aus der die Einspielergebnisse eines Films prognostiziert werden können.
- Sind diese signifikant höher, wenn es sich um eine Verfilmung eines Buches handelt und wenn ja, um wie viel?

6. Prüfung

Datenfile: pruefung.sav

Im Datenfile sind folgende Variablen enthalten:

Lernpensum	Lernaufwand in Stunden
Vorlesungsbesuch	regelmäßiger Besuch der Vorlesung 0 = nein, 1 = ja
Punkte	Punktezahl beim Test
Pruefung	Bestehen der Prüfung 0 = nein, 1 = ja

- Kann die Punktezahl durch Lernpensum und Vorlesungsbesuch erklärt werden?
- Hilft regelmäßiger Besuch der Vorlesung dabei, viel Punkte bei der Prüfung zu erreichen?

7. Mobilität und Scheidung

Datenfile: mobilitaet.sav

Im Datenfile sind folgende Variablen enthalten:

region Region der USA

mobility Mobilitätsrate

divorce Scheidungsrate

Soziologen untersuchten den Zusammenhang zwischen Mobilitätsrate (Prozentsatz an Personen, die innerhalb der letzten 5 Jahre die Wohnung gewechselt haben) und der Scheidungsrate (Anzahl Scheidungen eines Jahres pro 1000 Einwohner) in mehreren Regionen der USA.

- Beschreiben Sie den Zusammenhang zwischen den zwei Raten?
- Gibt es einen signifikanten Zusammenhang zwischen den beiden Raten?

Metrische und kategoriale Variablen

1. Alkohol und Gehalt

Datenfile: `gehaltalc.sav`

Im Datenfile sind folgende Variablen enthalten:

GEHALT Jahresgehalt (in 1000 Euro)
 ALKOHOL Alkoholkonsum, codiert mit
 0 = nie, 1 = gelegentlich, 2 = täglich
 RAUCH Zigaretten pro Tag

Die Einstellung zu Alkohol ist in Österreich sehr zwiespältig. Auf der einen Seite sind die Risiken des Alkoholkonsums nicht nur für die Gesundheit sondern auch für den sozialen Status wohl bekannt. Andererseits ist Alkohol ein fast allseits akzeptiertes Mittel, um soziale Kontakte aufzunehmen und zu pflegen. 70 Personen - nur Angestellte und Beamte - wurden befragt, wie viel Alkohol sie konsumieren (zusätzlich wurde auch der Zigarettenkonsum erhoben) und wie hoch ihr Gehalt ist.

- Kann man schließen, dass Alkoholkonsum und Gehalt in Beziehung stehen?

2. Personalfuktation

Datenfile: `personal.sav`

Im Datenfile sind folgende Variablen enthalten:

AUFNAHME Anzahl neu eingestellter Mitarbeiter
 ABGANG Anzahl von Mitarbeitern, die den Betrieb verlassen haben
 HOTELTYP Ausrichtung des Hotels primär auf
 1 = Seminargäste, 2 = Geschäftsreisende, 3 = Urlaubsgäste
 ORT Lage des Hotels
 1 = Stadt (mind. 50000 Einwohner), 2 = Land

Mitarbeiterzu- und -abgänge bieten Chancen und Risiken für Betriebe. In manchen Branchen sind sie häufiger als in anderen, in gewissen Branchen (etwa Saisonbetrieben im Tourismus) geradezu typisch, in anderen Branchen eher nur ein Anzeichen guter oder schlechter Konjunktur. Im vorliegenden Beispiel sind die Daten über Zugänge und Abgänge im Personal von 35 größeren Beherbergungsbetrieben (keine reinen Saisonbetriebe) gegeben.

- Wie hängt die Veränderung im Personalstand (Differenz zwischen Zu- und Abgängen) mit Lage und Typ des Hotels zusammen?

3. Bewertung von LVs

Datenfile: `lvbeurteilung.sav`

Im Datenfile sind folgende Variablen enthalten:

`punkte` Bewertung der LV mit Punkten (max. 10)
`prof` Geschlecht der(s) Lehrenden
`stud` Geschlecht der(s) Studierenden

An einer Uni können die Studierenden Lehrveranstaltungen und die Lehrenden mit Punkten bewerten.

- Zu welchem Schluss kommt man, wenn man nur jeweils den Einfluss eines Faktors (`prof` und `stud`) auf die Bewertung untersucht?
- Zu welchem Schluss kommt man bei Untersuchung des gemeinsamen Einflusses?

4. Filmbewertung

Datenfile: `filmbewertung.sav`

Im Datenfile sind folgende Variablen enthalten:

`PUNKTE` Punktescore von 0 (= sehr schlecht) bis 20 (= ausgezeichnet)
`SEX` Geschlecht 1 = weiblich, 2 = männlich

Filmbeurteilungen sind oft sehr subjektiv. Es kann vorkommen, dass das Urteil einer Gruppe sehr positiv, das einer anderen Gruppe stark negativ ausfällt. Besonders polarisierend sind Filme mit drastischer Darstellung von Gewalt und Sex.

Die Daten enthalten die Beurteilungen, gemessen auf einer Skala von 0 bis 20, eines Films mit mehreren brutalen Szenen.

- Existiert ein Unterschied in der Beurteilung zwischen Frauen und Männern?

5. Mitarbeit in Lehrveranstaltungen

Datenfile: `mitarbeit.sav`

Im Datenfile sind folgende Variablen enthalten:

`FRAGEN` Punktescore von 0 (= sehr schlecht) bis 20 (= ausgezeichnet)
`KURSZEIT` Beginnzeit des Kurses
1 = 09:00
2 = 13:00
3 = 16:00
`FORM` Sitzordnung
1 = in Reihen
2 = U-förmig

Die meisten LV-LeiterInnen befürworten aktive Mitarbeit von Studierenden. Im Idealfall stellen Studierende Fragen zum vorgetragenen Stoff und beantworten ihrerseits die Fragen, die an sie gerichtet werden. Der Leiter einer Statistik-LV weiß natürlich, dass viele Faktoren diese Mitarbeit beeinflussen,

zwei von ihnen möchte er untersuchen. Zum einen die Beginnzeit des jeweiligen Kurses, zum anderen die Anordnung der Sitze im Hörsaal.

Von 6 parallelen Lehrveranstaltungen eines Semesters mit jeweils ungefähr 40 Teilnehmern wurden 2 mit Beginnzeit 9 Uhr, zwei mit Beginnzeit 13 Uhr und zwei mit Beginnzeit 16 Uhr angesetzt. Von den jeweils zwei Kursen mit gleicher Beginnzeit fand einer in einem konventionellen Hörsaal mit einer Anordnung der Sitze in Reihen statt. Der jeweils andere Kurs fand in einem Seminarraum statt, in denen eine U-förmige Anordnung der Sitze vorgenommen wurde.

In jedem dieser Kurse wurde an 5 Tagen die Anzahl gestellter und beantworteter Fragen erhoben.

- Haben Beginnzeit und/oder Sitzanordnung relevanten Einfluss auf die Mitarbeit der Studierenden?

6. Investitionen

Datenfile: `investition.sav`

Im Datenfile sind folgende Variablen enthalten:

INVEST	Investitionen der letzten zwei Jahre (in 1000 Euro)
WELLNESS	Gab es Investitionen für Wellnessangebote 1 = Ja, 2 = Nein
REGION	Region des Betriebes 1 = Nord, 2 = Süd, 3 = West

Im Tourismus ist die Wellnessbewegung nicht unbeachtet geblieben. Viele Betriebe haben ihr Angebot an diesen neuen Markt mit seinen neuen Bedürfnissen angepasst. Hier liegen Investitionssummen vor, die Tourismusbetriebe innerhalb von zwei Jahren getätigt haben. Dabei wurde zum einen unterschieden, in welcher Region der Betrieb liegt, zum anderen, ob Investitionen auch in Wellnessangebote erfolgt sind.

- Gibt es Unterschiede in den Investitionssummen, die auf die Regionen zurückzuführen sind?
- Gibt es Unterschiede in den investierten Beträgen, die darauf zurückzuführen sind, dass Wellnessrichtungen erstellt wurden?
- Besteht eine Wechselwirkung?

7. *Haushaltsarbeit bei Teenagern

Datenfile: `teenagework.sav`

Im Datenfile sind folgende Variablen enthalten:

stunden	Haushaltsarbeit pro Woche (in Stunden)
mutter	Berufstätigkeit der Mutter (1=nein, 2=ja)
sex	Geschlecht der Teenager (1=weiblich, 2=männlich)

- Gibt es Unterschiede zwischen den Geschlechtern in der Mitarbeit im Haushalt?

8. ***Haushaltsarbeit bei Teenagern** Datenfile: `teenagework.sav`

Wir arbeiten mit dem Datenfile aus dem vorigen Beispiel.

- Ist der Unterschied in der Mitarbeit im Haushalt zwischen männlichen und weiblichen Teenagern anders, je nachdem ob die Mutter berufstätig ist oder nicht?

9. ***Verfahrensdauer am VwGH nach Senaten** Datenfile: `vwgh.sav`

Im Datenfile `vwgh.sav` über die Entscheidungen des Verwaltungsgerichtshofs sind auch die Verfahrensdauer vor dem VwGH (Variable `dauer3`) und der Senat, der über die Berufung beraten hat (Variable `senat`), angegeben.

- Unterscheidet sich die Dauer der Verfahren zwischen den verschiedenen Senaten?

10. ***Kurierdienste** Datenfile: `kurier.sav`

In den letzten Jahren sind viele Firmen dazu übergegangen, ihre Korrespondenz ganz oder teilweise von privaten Kurierdiensten befördern zu lassen. Ein großes Unternehmen plant diesen Schritt ebenfalls und möchte unter drei Kurierdiensten einen fest auswählen. Unter anderem ist auch die Zeit, in der Aufträge erledigt werden, ein sehr wichtiges Kriterium.

Um die Entscheidung zu erleichtern, werden jedem der drei Kurierdienste 12 zufällig ausgewählte Briefe (zufällige Aufgabezeit, zufälliger Bestimmungsort) zur Beförderung übergeben. Im Datenfile sind folgende Variablen enthalten:

`zeit` benötigte Zeit (in Minuten)
`kurier` Code zur Unterscheidung der Kurierdienste

- Gibt es signifikante Unterschiede in den Zustellzeiten zwischen den drei Kurierdiensten?

11. ***Carter – Reagan** Datenfile: `us-election80.sav`

Eine Teilmenge von Variablen des 1982 General Social Survey, betreffend die Präsidentschaftswahl 1980 sind:

VOTE	Wahlverhalten bei den Präsidentschaftswahlen 1980 1 = Reagan, 2 = Carter oder andere
RACE	Hautfarbe 1 = weiß, 2 = nicht weiß
POLVIEW	Likertskala für politische Einstellung 1 = extrem liberal .. bis 7 = extrem konservativ
WEIGHT	Anzahl von Beobachtungen mit entsprechender Variablenkombination

- Finden Sie ein Modell für das Wahlverhalten.

Multivariate Daten

1. Freizeitaktivitäten

Datenfile: `freizeit.sav`

Das Datenfile enthält mehrere Angaben zum Freizeitverhalten (`essentrinken` bis `kultur`).

- Lassen sich die vielen Aktivitäten zusammenfassen?
- Gibt es bei diesen Unterschiede zwischen Frauen und Männern?

2. *Kundenzufriedenheit

Datenfile: `konsumenten.sav`

Eine Untersuchung bei 253 Personen zur Kundenzufriedenheit mit einer Einzelhandelskette im Südosten der USA enthält Variablen mit sozialstatistischen Daten der befragten Person, verschiedene Fragen zur Kundenzufriedenheit und spezifische Fragen, wie die Kundenzufriedenheit verbessert werden könnte (Variablen `perf_1` bis `perf_20`).

- Um die Dimensionalität der Kundenzufriedenheit zu erforschen, soll eine Hauptkomponentenanalyse durchgeführt werden.
- Weiters sollen Forschungshypothesen formuliert und untersucht werden, die Unterschiede zwischen Konsumentengruppen (gebildet aus den sozialstatistischen Daten, wie z.B. Geschlecht, Alter, Einkommen) bezüglich der neuen Variablen (Komponenten) zum Gegenstand haben.

3. *Bewerbungen

Datenfile: `bewerbung.sav`

Die Daten aus [Kendall M. (1975). *Multivariate analysis*. Griffin, London] beziehen sich auf 48 Bewerbungen um eine Position in einem Unternehmen. Diese Bewerbungen wurden anhand von 15 Variablen bewertet.

- | | |
|---------------------------------|-------------------|
| ■ Form des Bewerbungsschreibens | ■ Erfahrung |
| ■ Erscheinung | ■ Schwung |
| ■ Akademische Fähigkeiten | ■ Ambition |
| ■ Sympathie | ■ Auffassungsgabe |
| ■ Selbstvertrauen | ■ Potenzial |
| ■ Klarheit | ■ Eifer |
| ■ Ehrlichkeit | ■ Eignung |
| ■ Geschäftstüchtigkeit | |

(Je höher der Wert, desto stärker ist die Eigenschaft ausgeprägt.)

- Gibt es einige zugrunde liegende Komponenten, die mit den Haupteigenschaften der Bewerber korrespondieren?

- Falls ja, können die Bewerber leichter verglichen werden. Basierend auf den Ergebnissen der Hauptkomponentenanalyse: Welche Kandidaten würde man auswählen, wenn die zu besetzende Position im Verkauf, im Marketing oder aber in der Abteilungsleitung angesiedelt ist?