

## Statistik

Der Weg zur Datenanalyse

Bearbeitet von

Ludwig Fahrmeir, Christian Heumann, Rita Künstler, Iris Pigeot, Gerhard Tutz

8. Auflage 2016. Taschenbuch. XVI, 581 S. Softcover

ISBN 978 3 662 50371 3

Format (B x L): 19,3 x 26 cm

[Weitere Fachgebiete > Mathematik > Stochastik > Mathematische Statistik](#)

Zu [Inhaltsverzeichnis](#)

schnell und portofrei erhältlich bei

**beck-shop.de**  
DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.

# Kapitel 1

## Einführung

### 1.1 Wo braucht man Statistik?

Zunächst sollen einige Beispiele, die später ausführlicher behandelt werden, typische Fragestellungen und Anwendungssituationen veranschaulichen, bei denen statistische Methoden eingesetzt werden. Die statistischen Begriffe, die in diesem Zusammenhang fallen, werden in den entsprechenden Kapiteln eingeführt.

#### Münchener Absolventenstudie 2011

#### Beispiel 1.1

Im Laufe der letzten Jahre wurden zur Beurteilung der Berufsaussichten von Studienabgängerinnen und -abgängern mit insbesondere sozialwissenschaftlicher Ausrichtung an einigen Universitäten Deutschlands Befragungen der Absolventen durchgeführt. Um die Berufsaussichten einzuschätzen zu können, ist eine Fülle an Informationen erforderlich. Am Institut für Soziologie der Ludwig-Maximilians-Universität München wurde daher für die Absolventenstudie 2011 ein spezieller Fragebogen konzipiert, der insgesamt 75 Fragen umfasst. Die Fragen betrafen folgende Sachverhalte: Wie kann das Studienverhalten der Soziologie-Absolventen der LMU München eingeschätzt werden? Welchen Tätigkeiten gehen die Absolventen nach ihrem Studium nach? Wie ist der Übergang in die Arbeitswelt erfolgt? Der Fragebogen deckte zahlreiche konkrete Aspekte ab wie etwa den Studienverlauf (z.B. Anzahl absolviertes Semester, Gesamtnote, Wechsel des Studienorts, Praktika), mögliche Zusatzqualifikation, aber auch Aspekte zur Person wie z.B. Geschlecht, Alter, Familienstand und berufliche Stellung der Eltern.

Der Fragebogen wurde im Jahr 2011 an 716 Personen verschickt und zugestellt. Insgesamt wurden Daten von 441 Fragebögen vom Institut für Soziologie zur Verfügung gestellt. Einen Auszug von 36 zufällig ausgewählten Personen und fünf Merkmalen, die im Folgenden erläutert werden, findet man in Tabelle 1.1.

Wir verwenden diesen Teildatensatz zur Veranschaulichung einiger elementarer Methoden der Statistik. Die fünf Variablen sind: ( $G$  mit 1 = männlich, 2 = weiblich), Studiendauer in Semestern ( $S$ ), Frage: Wie häufig benötigten bzw. benötigen Sie folgende Studieninhalte bei Ihrer jetzigen Stelle (nur Fach Statistik)? ( $Stat$  mit vier Kategorien: 1 = ständig, 2 = häufig, 3 = gelegentlich, 4 = nie), hauptsächliche Finanzierungsquelle während des Studiums ( $F$  mit vier möglichen Ausprägungen: 1 = Zuwendung von Eltern, Verwandten, 3 = BAföG, 4 = Eigenes Einkommen, 5 = Sonstige) sowie Gesamtnote ( $N$ ).

Dabei wurden diese Variablen nicht speziell in Hinblick auf die eigentlich interessierende Fragestellung ausgewählt, sondern aus eher illustrativen Gründen. □

Person $i$	$G$	$S$	Stat	$F$	$N$	Person $i$	$G$	$S$	Stat	$F$	$N$
1	2	11	4	4	2.7	2	2	12	2	4	2.3
3	2	11	1	1	2.4	4	1	12	3	4	2.74
5	2	10	2	4	1.51	6	2	12	4	1	2.15
7	1	9	2	4	2.03	8	1	16	3	1	2
9	2	11	3	1	1.35	10	2	11	1	4	3.18
11	2	14	2	1	2.12	12	1	11	4	1	2.01
13	1	12	4	4	1.82	14	2	12	4	3	1.52
15	2	11	1	1	1.91	16	2	11	3	1	1.7
17	2	11	4	1	2.05	18	2	14	2	5	1.89
19	2	12	3	1	2	20	2	11	2	1	2.29
21	1	12	3	1	1.8	22	2	13	2	3	1.75
23	1	15	1	1	1.9	24	2	11	3	3	1.24
25	1	12	3	4	2.02	26	2	14	3	1	1.3
27	1	19	3	4	2.5	28	2	9	3	1	2.7
29	2	12	4	1	2.3	30	1	14	3	4	1.8
31	2	13	3	4	1.56	32	1	14	1	4	3.27
33	2	11	4	1	2.25	34	2	10	2	1	1.43
35	1	14	2	1	2	36	2	11	4	1	2.5

Tabelle 1.1: Daten für 36 Absolventen der Münchener Absolventenstudie 2011 des Instituts für Soziologie der LMU München

## Beispiel 1.2

### Mietspiegel 2015

In vielen Städten und Gemeinden der Bundesrepublik werden sogenannte Mietspiegel erstellt. Sie bieten Mietern und Vermietern eine Marktübersicht zu Miethöhen, helfen in Mietberatungsstellen und werden, neben Sachverständigen, auch zur Entscheidung in Mietstreitprozessen herangezogen. Das frühere Miethöhengesetz wurde in das Bürgerliche Gesetzbuch (BGB) integriert. In §558, Absatz (2) ist die ortsübliche Vergleichsmiete definiert: „Die ortsübliche Vergleichsmiete wird gebildet aus den üblichen Entgelten, die in der Gemeinde oder einer vergleichbaren Gemeinde für Wohnraum vergleichbarer Art, Größe, Ausstattung, Beschaffenheit und Lage einschließlich der energetischen Ausstattung und Beschaffenheit in den letzten vier Jahren vereinbart oder, von Erhöhungen nach §560 abgesehen, geändert worden sind. Ausgenommen ist Wohnraum, bei dem die Miethöhe durch Gesetz oder im Zusammenhang mit einer Förderzusage festgelegt worden ist.“ Damit werden erstens die Grundgesamtheiten festgelegt, aus denen die Stichproben für die Erstellung von Mietspiegeln zu ziehen sind. Zweitens wird zugleich ein Hinweis auf die statistische Analysemethode gegeben: Sinngemäß bedeutet dies für die Nettomiete, dass ihr Durchschnittswert in Abhängigkeit von Merkmalen wie Art, Größe, Ausstattung, Beschaffenheit und Lage der Wohnung zu bestimmen bzw. zu schätzen ist. Wir beschränken uns hier auf die sogenannte Nettomiete, d.h. den monatlichen Mietpreis, der nach Abzug aller Nebenkosten übrigbleibt. Die wichtigsten Faktoren, mit denen ein Teil der erheblichen Streuung der Nettomieten um einen Durchschnittswert erklärt werden kann, sind die Wohnfläche und das Baualter. Diese beiden Merkmale finden in allen herkömmlichen Tabellenmietspiegeln Berücksichtigung. Darüber hinaus werden aber auch weitere Merkmale zur Lage und Ausstattung der Wohnung einbezogen, die zu begründeten Zu- oder Abschlägen führen.

Zur Erstellung eines Mietspiegels wird aus der Gesamtheit aller nach dem Gesetz relevanten Wohnungen der Stadt eine repräsentative Stichprobe gezogen und die interessierenden Daten werden von Interviewern in Fragebögen eingetragen. Das mit der Datenerhebung beauftragte Institut, in München Infratest, erstellt daraus eine Datei, die der anschließenden statistischen Beschreibung,

Auswertung und Analyse zugrunde liegt. Die Präsentation der Ergebnisse erfolgt schließlich in einer Mietspiegelbroschüre bzw. im Internet.

Im Folgenden betrachten wir einen Ausschnitt aus dem Mietspiegel München 2015. Eigentliches Zielmerkmal ist die monatliche Nettomiete (NM), auf die gesamte Wohnfläche oder pro Quadratmeter bezogen. Weitere Merkmale sind unter anderem: Wohnfläche (in qm), Baualter (Baujahr oder Baualterskategorie), Zentralheizung (ja/nein), Warmwasserversorgung (ja/nein), gehobene Ausstattung des Bades bzw. der Küche. Tabelle 1.2 enthält durchschnittliche Nettomieten/qm, gegliedert nach Baualters- und Wohnflächenkategorien (klein, mittel, groß). Sie gibt somit einen ersten Einblick in die Datenlage, wie sie sich in einer Teilstichprobe von 3065 Wohnungen in München darstellt. □

Nettomiete/qm				
Wohnfläche				
Baualter	bis 38 qm	39 bis 80 qm	81 bis 120 qm	121 qm und mehr
bis 1918	11.92 (13)	10.92 (159)	10.56 (124)	9.46 (51)
1919 bis 48	12.60 (12)	9.51 (220)	9.99 (96)	11.07 (9)
1949 bis 65	12.59 (86)	10.16 (607)	9.86 (186)	9.33 (10)
1966 bis 77	12.89 (79)	10.30 (365)	9.73 (149)	10.21 (8)
1978 bis 89	13.71 (19)	10.96 (161)	10.22 (72)	10.64 (8)
ab 1990	13.90 (17)	11.86 (381)	12.01 (208)	12.98 (25)

Tabelle 1.2: Einfacher Tabellen-Mietspiegel, in Klammern Anzahl der einbezogenen Wohnungen

## Luftschadstoffe

Die Überwachung, Kontrolle und Reduzierung von Luftschadstoffen ist ein wichtiges umwelt- und gesundheitspolitisches Ziel. In vielen Städten und Gemeinden werden daher an ausgewählten Messstellen entsprechende Messwerte aufgezeichnet und ausgewertet. Abbildung 1.1 zeigt die mittleren Monatswerte der Konzentration von Stickoxiden (Stickstoffdioxid und -monoxid) und Kohlenmonoxid für den Zeitraum 1990 bis 2015. Die Lücken in den Messwerten in den ersten Jahren bedeuten fehlende Werte. Die Einheiten sind ppb (Parts per Billion) bei Stickoxiden und Mikrogramm/Kubikmeter bei Kohlenmonoxid. Die Messstation befindet sich in der zentral gelegenen und mäßig verkehrsbelasteten Stampfenbachstraße in Zürich. Sie ist Teil eines größeren Netzwerks in der Ostschweiz und Liechtenstein (Quelle: OSTLUFT.ch). Typische Fragestellungen sind: Lässt sich ein längerfristiger Trend feststellen? Wie groß sind saisonale Einflüsse? Lassen sich Erfolge umweltpolitischer Maßnahmen erkennen? Wir werden solchen Fragen in Kapitel 14 nachgehen. □

## Beispiel 1.3

## Politische Umfragen

Befragungen zur Beliebtheit von Politikern, zur Beurteilung der wirtschaftlichen Lage oder darüber, welche Partei man wählen würde, werden regelmäßig von bekannten Instituten durchgeführt. Abbildung 1.2 vergleicht die Beliebtheit von Merkel und Steinbrück für die Monate des Jahres 2013. Die Grafik vermittelt einen sehr eindeutigen Eindruck von sich abzeichnenden Tendenzen. Da die Daten aus Stichproben gewonnen werden, ist jedoch Vorsicht bei der Interpretation geboten: Wie groß ist der Anteil der Stichprobenfehler, also zufälliger Veränderungen, gegenüber substanziellem Veränderungen?

## Beispiel 1.4

Tabelle 1.3 zeigt für den Befragungszeitraum (10.12.2013- 12.12.2013) Ergebnisse zur sogenannten Sonntagsfrage „Welche Partei würden Sie wählen, wenn am nächsten Sonntag Bundestagswahlen wären?“. Dazu sind zu einer Stichprobe von 1349 Personen entsprechende Prozentzahlen

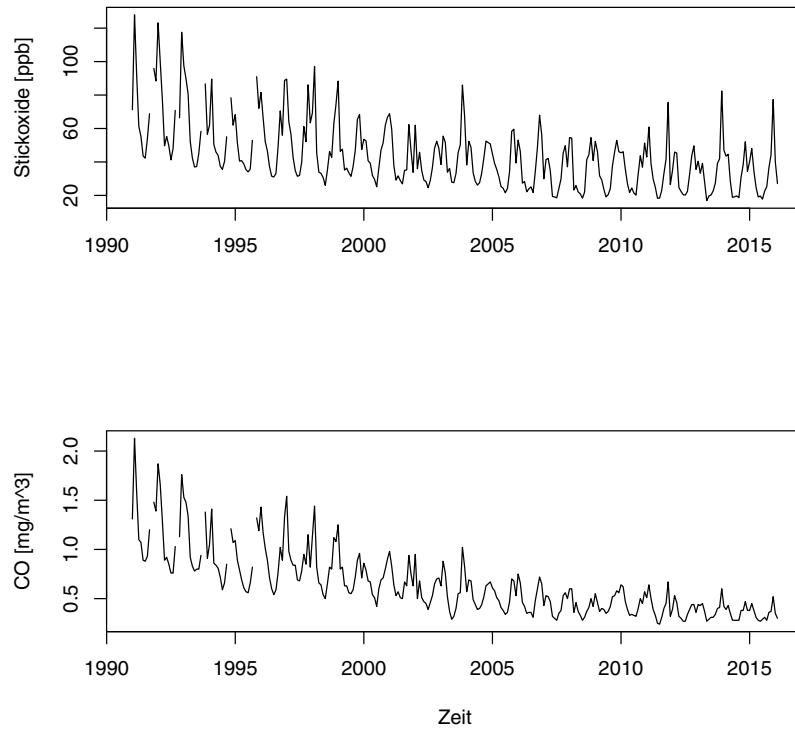


Abbildung 1.1: Mittlere Monatswerte der Konzentration von Stickoxiden (Stickstoffdioxid und -monoxid) und Kohlenmonoxid

getrennt nach Geschlecht bzw. insgesamt angegeben. Es entsteht der Eindruck, dass die Parteipräferenz vom Geschlecht abhängt. Da es sich um eine Stichprobe handelt, entsteht wieder die Frage, ob dies Zufall oder „statistisch signifikant“ ist.  $\square$

	CDU/CSU	SPD	FDP	Linke	Grüne	Rest	
Männer	38.8	24.3	5.4	14.8	7.9	8.8	100
Frauen	44.5	25.0	1.6	11.4	13.0	4.5	100
	41.4	24.6	3.6	13.2	10.3	6.9	

Tabelle 1.3: Prozentzahlen der Parteipräferenz bei der Sonntagsfrage. Quelle: Jung, Matthias; Schroth, Yvonne; Wolf, Andrea (2015): Politbarometer 2013 (Cumulated Data Set, incl. Flash). GESIS Data Archive, Cologne. ZA5677 Data file Version 1.1.0, doi:10.4232/1.12171

### Beispiel 1.5 Kreditwürdigkeitsprüfung und Insolvenzprognose

Bei der Kreditvergabe ist es für Banken offensichtlich wichtig, dass die Rückzahlung ordnungsgemäß abgewickelt wird. Um zu vermeiden, dass es zu Verzögerungen der Ratenzahlungen oder gar zum Kreditausfall kommt, ist es daher nötig, die zukünftige Bonität eines potenziellen Kreditnehmers abzuschätzen und die Vergabe eines Kredits davon abhängig zu machen. Die Bank steht vor

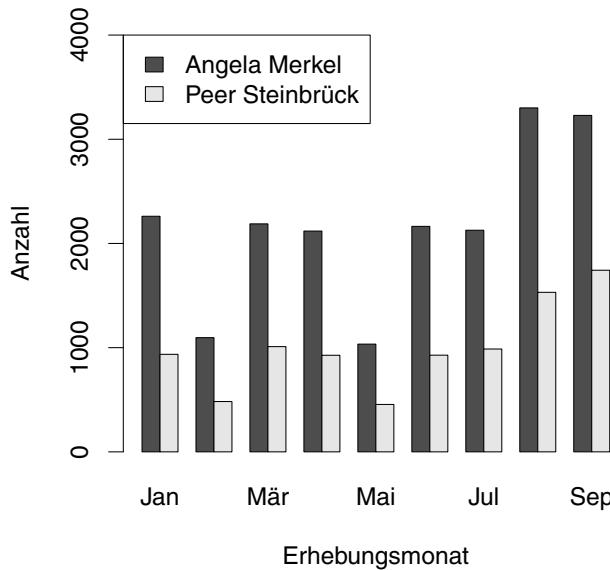


Abbildung 1.2: Kanzlerpräferenz

dem Problem, einen potenziellen Kreditnehmer entweder der Klasse der problemlosen Kreditnehmer zuzuordnen und den Kredit zu vergeben, oder ihn der Klasse der Problemfälle zuzuordnen und auf das Kreditgeschäft zu verzichten bzw. eine genauere Prüfung vorzunehmen.

Für eine datengestützte und automatische Entscheidungshilfe bei der Kreditvergabe werden bereits abgewickelte Kredite als sogenannte Lernstichproben herangezogen. Für jeden Kunden aus dieser Stichprobe ist seine „Kreditwürdigkeit“  $Y$  (ja = 0, nein = 1) bekannt. Zusätzlich werden bei der Kreditvergabe weitere Merkmale zum Kredit und zur Person des Kunden erhoben, die Hinweise darauf geben sollen, ob der Kunde als kreditwürdig einzustufen ist oder nicht. Wir betrachten folgende Merkmale aus einer Lernstichprobe von 1000 Konsumentenkrediten einer süddeutschen Großbank mit 300 schlechten und 700 guten Krediten:

- $X_1$  Laufendes Konto bei der Bank (nein = 1, mittel = 2, gut = 3)
- $X_2$  Laufzeit des Kredits in Monaten
- $X_3$  Kredithöhe in Euro
- $X_4$  Rückzahlung früherer Kredite (gut/schlecht)
- $X_5$  Verwendungszweck (privat/beruflich)
- $X_6$  Geschlecht (weiblich/männlich)

Die folgende Tabelle 1.4 gibt für die 300 schlechten ( $Y = 1$ ) und die 700 guten ( $Y = 0$ ) Kredite jeweils die Prozentzahlen für die Ausprägungen der ausgewählten Merkmale  $X_1$ ,  $X_3$ ,  $X_4$  und  $X_5$  an. Sieht man sich die Prozentzahlen für die Variable  $X_1$  in beiden „Schichten“  $Y = 1$  und  $Y = 0$  an, so erkennt man, dass sich diese Variable in den beiden Schichten deutlich unterschiedlich verhält. Dagegen ist zum Beispiel für die Variable  $X_3$  kein deutlicher Unterschied erkennbar.

Automatische Systeme zur Kreditwürdigkeitsprüfung nutzen Merkmale mit deutlichen Unterschieden in beiden Schichten so, dass ein neuer Kunde aufgrund seiner Ausprägungen in diesen

		$Y$	
$X_1$ : laufendes Konto		1	0
nein		45.0	19.9
gut		15.3	49.7
mittel		39.7	30.2
$X_3$ : Kredithöhe		$Y$	
in Euro		1	0
$< \dots \leq 500$		1.00	2.14
$500 < \dots \leq 1000$		11.33	9.14
$1000 < \dots \leq 1500$		17.00	19.86
$1500 < \dots \leq 2500$		19.67	24.57
$2500 < \dots \leq 5000$		25.00	28.57
$5000 < \dots \leq 7500$		11.33	9.71
$7500 < \dots \leq 10000$		6.67	3.71
$10000 < \dots \leq 15000$		7.00	2.00
$15000 < \dots \leq 20000$		1.00	.29
$X_4$ : Frühere Kredite		$Y$	
		1	0
gut		82.33	94.85
schlecht		17.66	5.15
$X_5$ : Verwendungszweck		$Y$	
		1	0
privat		57.53	69.29
beruflich		42.47	30.71

Tabelle 1.4: Prozentzahlen ausgewählter Merkmale zur Kreditwürdigkeit

Merkmale als kreditwürdig oder nicht kreditwürdig eingestuft wird. Dabei sollen Fehleinstufungen mit möglichst geringen Wahrscheinlichkeiten auftreten.  $\square$

## Beispiel 1.6

### Aktienkurse

Ein wichtiges Anwendungsgebiet der Statistik ist die Analyse von Finanzmarktdaten, z.B. von Aktien-, Wechsel- und Zinskursen. Ziele sind etwa das Aufspüren von systematischen Strukturen, der Einfluss bestimmter politischer oder wirtschaftlicher Ereignisse, schnelles Erkennen von Veränderungen und Entscheidungsunterstützung für An- und Verkauf sowie für Portfoliogestaltung.

Neben Kursen einzelner Aktien bilden dazu auch die Werte eines Aktienindex eine wichtige Datengrundlage. Bekannte Aktienindizes sind etwa der Deutsche Aktienindex (DAX), der MDAX, in dem mittelgroße Unternehmen (deutsche Midcaps) vertreten sind, international der S&P Global 1200 (Standard & Poor's Global 1200), welcher 1200 der größten Aktiengesellschaften der Welt enthält, oder der MSCI World, der Aktien von Unternehmen aus 23 Ländern enthält.

Abbildung 1.3 zeigt die täglichen Kurse der BMW-Aktie von Januar 2000 bis Juni 2015.

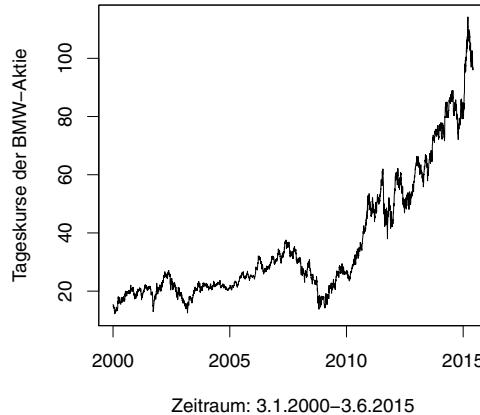


Abbildung 1.3: Tageskurse der BMW-Aktie

Durch Vergleich mit dem DAX lassen sich sogenannte *Beta-Faktoren* für die Aktien berechnen. Diese werden zur Steuerung von Aktienportfolios verwendet. Die Berechnung der Beta-Faktoren basiert auf dem „*Capital Asset Pricing Model*“ (CAPM) und wird in den Abschnitten 3.6 und 12.1 illustriert. Das CAPM spielt in der Kapitalmarkttheorie eine wichtige Rolle. Anhand dieses Modells wird versucht, Aussagen über die Rendite einer Aktie unter Berücksichtigung ihres Risikos zu treffen. □

*Beta-Faktoren*

*Capital Asset Pricing Model (CAPM)*

## Klinische und epidemiologische Studien

In medizinischen Studien ist oft die Beurteilung der Wirkung eines Medikaments bzw. einer Therapie oder das Erkennen bestimmter Risikofaktoren von primärem Interesse. In sogenannten randomisierten Studien werden Patienten z.B. zufällig entweder der Therapiegruppe oder einer Placebogruppe zugewiesen. Zusätzlich werden in der Regel noch personenspezifische Merkmale wie Alter, Geschlecht, allgemeiner Gesundheitszustand oder besondere Risikofaktoren mit erhoben. Ziel ist der statistische Nachweis der Wirksamkeit des Medikaments, der Therapie bzw. die Quantifizierung des Einflusses anderer Merkmale und Faktoren.

**Beispiel 1.7**

Als Beispiel betrachten wir eine am Klinikum der Technischen Universität München in den Jahren 1987 bis 1994 durchgeführte Studie zur Überlebenszeit nach einer Magenkrebsoperation, wobei der Tumor vollständig entfernt wurde. Von den 179 Patienten verstarben 82 an den Folgen der Tumorerkrankung, die restlichen Patienten überlebten bis zum Ende der Studie oder starben aufgrund anderer Ursachen. Neben dem Zeitpunkt der Operation, dem Alter und dem Geschlecht wurden von jedem Patienten prognostische Faktoren erhoben, die Art, Stärke, Lokalisation und Stadium der Tumorerkrankung betrafen. Von besonderem Interesse ist die Analyse des Einflusses, den solche Faktoren auf den weiteren Krankheitsverlauf und langfristige Überlebenschancen haben. □

## Konjunkturtest des IFO-Instituts

Seit 1949 führt das IFO-Institut in München monatliche Unternehmensbefragungen durch, an denen sich zur Zeit rund 12000 Unternehmen des verarbeitenden und des Baugewerbes sowie des Groß-

**Beispiel 1.8**

Rücksendung des ifo Exemplars erbaten bis		Aktuelle Situation			Erwartungen für die nächsten 3 Monate			
		1)	2)	3)	7)	8)	9)	10)
1) Wir beurteilen unsere Geschäftslage für XY als	gut	<input type="checkbox"/>	befriedigend	<input type="checkbox"/>	schlecht	<input type="checkbox"/>	erhöht	<input type="checkbox"/>
2) Unseren Bestand an unverkauften Fertigwaren von XY empfinden wir als	zu klein	<input type="checkbox"/>	ausreichend (saisonüblich)	<input type="checkbox"/>	zu groß	<input type="checkbox"/>	nicht verändert	<input type="checkbox"/>
3) Unseren Auftragsbestand für XY empfinden wir als	verhältnismäßig groß	<input type="checkbox"/>	ausreichend (saisonüblich)	<input type="checkbox"/>	zu klein	<input type="checkbox"/>	gesenkt	<input type="checkbox"/>
	insgesamt	<input type="checkbox"/>	für den Export	<input type="checkbox"/>				
4) Tendenzen im vorangegangenen Monat	4) Die Nachfragesituation für XY hat sich	gebessert	<input type="checkbox"/>	nicht verändert	<input type="checkbox"/>	verschlechtert	<input type="checkbox"/>	
	5) Unser Auftragsbestand (In- u. Ausland, wertmäßig) für XY ist	gestiegen	<input type="checkbox"/>	etwa gleich geblieben (bzw. nicht üblich)	<input type="checkbox"/>	gesunken	<input type="checkbox"/>	
	6) Unsere inländische Produktionstätigkeit* bezüglich XY ist	gestiegen	<input type="checkbox"/>					
	7) Unsere Inlandsverkaufspreise (Netto) für XY wurden - unter Berücksichtigung von Konditionsveränderungen -	erhöht	<input type="checkbox"/>	nicht verändert	<input type="checkbox"/>	gesenkt	<input type="checkbox"/>	
	8) Unsere inländische Produktionstätigkeit* bezüglich XY wird voraussichtlich	steigen	<input type="checkbox"/>	etwa gleich bleiben	<input type="checkbox"/>	abnehmen	<input type="checkbox"/>	
	9) Unsere Inlandsverkaufspreise (Netto) für XY werden - unter Berücksichtigung von Konditionsveränderungen - voraussichtlich	steigen	<input type="checkbox"/>	etwa gleich bleiben	<input type="checkbox"/>	fallen	<input type="checkbox"/>	
	10) Der Umfang unseres Exportgeschäfts mit XY wird voraussichtlich - unter Berücksichtigung der bisherigen Exportabschlüsse und der laufenden Auftragsverhandlungen -	zunehmen	<input type="checkbox"/>	etwa gleich bleiben	<input type="checkbox"/>	abnehmen	<input type="checkbox"/>	
		wir exportieren XY nicht	<input type="checkbox"/>					
	11) Beschäftigte (nur inländische Betriebe) Die Zahl der mit der Herstellung von XY beschäftigten Arbeitnehmer wird	zunehmen	<input type="checkbox"/>	etwa gleich bleiben	<input type="checkbox"/>	abnehmen	<input type="checkbox"/>	

Abbildung 1.4: Fragebogen des IFO-Konjunkturtests für das verarbeitende Gewerbe

und Einzelhandels beteiligen. Ziel dieses sogenannten „Konjunkturtests“ ist es, direkt und schnell von den Unternehmen Auskunft z.B. über die Veränderungen von Urteilen, Plänen und Erwartungen zu bestimmten ökonomischen Größen zu bekommen, die sich als Indikatoren zur Konjunkturbeurteilung eignen.

Der IFO-Konjunkturtest beschränkt sich auf die Erhebung von Entwicklungstendenzen bestimmter betrieblicher Variablen, wie z.B. Produktion, Preise, Auftrags- und Lagerbestand bzw. die Beurteilung dieser Entwicklungen bezogen auf die Vergangenheit und/oder die Zukunft. Themen der Fragen sind für das verarbeitende Gewerbe neben der aktuellen Beurteilung von Geschäftslage, Auftrags- und Lagerbestand die Veränderung gegenüber dem Vormonat von Nachfrage, Auftragsbestand, Produktion und Preisen sowie die Erwartungen über die Entwicklung der Preise, der Produktion, des Exportgeschäfts und der allgemeinen Geschäftslage. Ein Ausschnitt des Fragebogens ist in Abbildung 1.4 dargestellt.

Die daraus erhaltenen Daten werden auch Mikrodaten genannt, da sie auf Unternehmensebene erhoben werden. Im Folgenden sind einige Fragen für das verarbeitende Gewerbe wiedergegeben, die von den teilnehmenden Unternehmen jeweils monatlich beantwortet werden. Die Daten des Konjunkturtests stellen eine Alternative zu Daten der amtlichen Statistik dar, mit dem Vorteil, dass sie einfacher erhoben und deutlich rascher veröffentlicht werden können. Durch geeignete Aufbereitung, wie etwa im monatlichen Konjunkturspiegel, und statistische Analysen wird eine kurzfristige Beurteilung der konjunkturrellen Entwicklung möglich. Abbildung 1.5 zeigt den Geschäftsklimaindex für die gewerbliche Wirtschaft. Basis der Berechnung des Indexes sind sogenannte Salden. Die Unternehmen können auf die Fragen zur Geschäftserwartung mit „günstiger“, „gleichbleibend“ oder „ungünstiger“ antworten. Für die gegenwärtige Geschäftslage stehen die Antworten „gut“, „befriedigend“ oder „schlecht“ zur Auswahl. Der Saldowert der gegenwärtigen Geschäftslage ist die Diffe-

renz der Prozentanteile der Antworten „gut“ und „schlecht“, der Saldowert der Erwartungen ist die Differenz der Prozentanteile der Antworten „günstiger“ und „ungünstiger“.

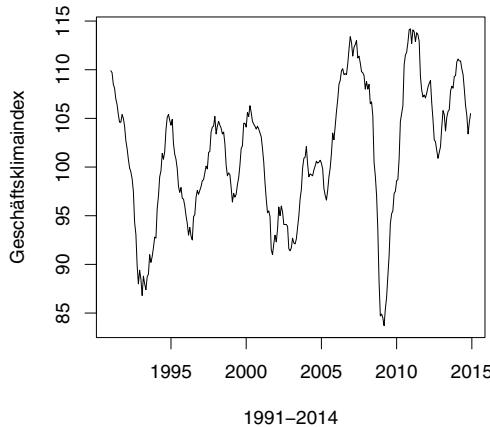


Abbildung 1.5: Geschäftsklimaindex für die gewerbliche Wirtschaft 01/1991-12/2014

Das Geschäftsklima ist ein (geometrischer) Mittelwert aus den Salden der Geschäftslage und Erwartungen. Durch Normierung entsteht daraus der Geschäftsklimaindex, dessen Werte wie in Abbildung 1.5 um den Wert 100 nach unten und oben schwanken.

□

## Das SHARE-Projekt

## Beispiel 1.9

Das SHARE-Projekt (Survey of Health, Ageing and Retirement in Europe) ist eine interdisziplinäre und länderübergreifende Studie mit Mikrodaten (also personenbezogenen Daten) zu den folgenden Merkmalen und Bereichen (nur exemplarische Auswahl).

- Demografie: Alter, Geschlecht, Geburtsland, Alter und Geschlecht des Partners
- Haushaltszusammensetzung: Haushaltsgröße, im Haushalt lebende Kinder
- Soziales Netzwerk: Anzahl Kinder, Anzahl Enkelkinder, Anzahl lebender Geschwister
- Bedingungen im Kindesalter: Anzahl Bücher im Alter von 10 Jahren, relative mathematische Kenntnisse und Sprachkenntnisse im Alter von 10 Jahren
- Gesundheit und Gesundheitsverhalten: chronische Krankheiten, Depressionsskala (EURO-D), CASP-12 Index für Lebensqualität, Body Mass Index (BMI), Rauch- und Trinkgewohnheiten, Sportaktivitäten
- Funktionale Einschränkungen, gemessen mit verschiedenen Indizes: Mobilitätsindex, grob-motorische und fein-motorische Fähigkeiten, kognitive Funktionen
- Arbeit und Geld: momentane Arbeitssituation, Art des Hauptjobs, Arbeitsstunden pro Woche, Zufriedenheit mit dem Hauptjob, Einkommensperzentil des monatlichen Haushaltseinkommens, Pläne für vorzeitigen Ruhestand, Ausgaben für Nahrungsmittel

Die Studie ist als Panelstudie angelegt, das heißt die Befragung wird über mehrere Jahre in sogenannten Wellen durchgeführt, und enthält mittlerweile über 123.000 Personen im Alter von 50 Jahren oder darüber aus 20 europäischen Ländern und Israel. Insgesamt wurden bis jetzt etwa 293.000 Interviews durchgeführt. Jede eingeschlossene Person erhielt eine eindeutige Identifikationsnummer und kann mehrfach interviewt worden sein.  $\square$

### Beispiel 1.10 Marktanteile bei Smartphones

In 2014 wurden etwa 1,3 Milliarden Smartphones verkauft. Die größten Marktanteile besitzen Samsung (314,2 Millionen) und Apple (192,7 Millionen). (Quelle: <http://communities-dominate.blogs.com/>, abgerufen am 25.2.2016).

Samsung	314.2	LG	59.1
Apple	192.7	Coolpad	50.4
Lenovo	95.0	ZTE	46.1
Huawei	75.0	TCL-Alcatel	41.4
Xiaomi	61.1	Sony	40.2

Tabelle 1.5: Anzahl verkaufter Smartphones in Millionen der zehn Hersteller mit den meisten Verkäufen

In Tabelle 1.5 sind die Anteile der zehn Smartphone-Hersteller mit den meisten Verkäufen angegeben, auf die wir uns im Folgenden beschränken. Diese machen etwa 75% aller Verkäufe aus. Kann man hier bereits von einer hohen Marktkonzentration sprechen, d.h. wenige Hersteller dominieren den Smartphone-Markt? In Kapitel 2.3 werden wir geeignete Maßzahlen zur Messung der Konzentration kennenlernen.  $\square$

## 1.2 Was macht man mit Statistik?

Wie an den vorangehenden Beispielen deutlich wurde, lassen sich allgemein bei wissenschaftlichen Untersuchungen oder Problemstellungen der Praxis verschiedene Stufen unterscheiden, die von der Planung einer Untersuchung bis hin zur Auswertung der Ergebnisse und Formulierung neuer Fragestellungen reichen. Auf jeder Stufe kommen Methoden entsprechender Teilgebiete der Statistik zum Einsatz. So interessieren im Rahmen der Planung einer Erhebung oder eines Versuches statistische Methoden der Stichprobenziehung und der Versuchsplanung. Aber auch zur Überprüfung der Reliabilität und Validität von Fragebögen werden statistische Verfahren eingesetzt. Speziell bei der Datenanalyse lassen sich drei Grundaufgaben der Statistik angeben: Beschreiben (*Deskription*), Suchen (*Exploration*) und Schließen (*Induktion*). Jeder dieser Aufgaben entspricht ein Teilgebiet der Statistik. So widmet sich die deskriptive Statistik der Beschreibung und Darstellung von Daten. Die explorative Statistik befasst sich mit dem Auffinden von Strukturen, Fragestellungen und Hypothesen, während die induktive Statistik Methoden bereitstellt, um statistische Schlüsse mittels stochastischer Modelle ziehen zu können. Diese drei Gebiete werden im Folgenden kurz beschrieben und in entsprechenden Kapiteln dieses Buches ausführlich diskutiert.

Die *deskriptive Statistik* dient zunächst zur beschreibenden und grafischen Aufbereitung und Komprimierung von Daten (vgl. Kapitel 2 und 3). Dies ist vor allem zur Präsentation

Statistik

Der Weg zur Datenanalyse

Fahrmeir, L.; Heumann, C.; Künstler, R.; Pigeot, I.; Tutz, G.

2016, XVI, 581 S. 162 Abb., Softcover

ISBN: 978-3-662-50371-3