
2.1 Datenerfassung in die Excel-Tabelle

2.1.1 Einführung

Durch Zählvorgänge oder durch Messungen, Befragungen und Beobachtungen entsteht in der Regel mehrdimensionales Datenmaterial. Als Voraussetzung für jegliche Aufbereitung und statistische Analyse muss dieses Datenmaterial zuerst in einer Excel-Tabelle *erfasst und gespeichert* werden. Zur Vorbereitung der Erfassung sollte man sich etwas Zeit nehmen; je mehr Gedanken man sich in dieser Phase macht, desto übersichtlicher werden später die Daten vorliegen. Das beginnt bereits bei den *Spaltenüberschriften*.

2.1.2 Spaltenüberschriften

2.1.2.1 Pro und contra erste Zeile

Viele Lehrende verlangen, dass ein Datenbestand in einer Excel-Tabelle stets links oben ab der Zelle A1 eingetragen wird. Das bedeutet also, dass die *Überschriften* in der *Zeile 1* und der eigentliche *Datenbestand* später darunter in den *Zeilen 2, 3, usw.* eingetragen wird.

Um sich die Vor- und Nachteile dieser Vorgehensweise deutlich zu machen, sollte man stets daran denken, dass *nach* der Erfassung der Daten mit Sicherheit *Auswertungen* vorgenommen werden müssen.

Reicht der Platz rechts neben dem Datenbestand oder darunter aus, dann können dort die auswertenden Formeln und Funktionen eingetragen werden und bleiben nach der Da-

Abteilung	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez	Jan
Abt_1	6,0	7,3	7,6	6,1	6,2	6,3	7,2	8,0	5,9	6,3	7,3	7,5	6,9
Abt_2	6,7	5,4	6,2	5,4	5,6	6,2	6,1	7,9	5,4	5,8	6,6	5,1	7,7
Abt_3	7,4	7,5	6,6	7,8	5,1	6,7	5,2	6,5	6,0	5,8	5,2	6,3	6,4
Abt_4	7,7	7,3	7,0	5,8	5,2	6,6	7,0	5,1	7,4	6,6	5,2	8,0	5,4
Abt_5	6,8	5,9	6,0	7,2	6,8	5,9	5,5	5,2	5,7	7,6	7,4	7,8	5,7
Abt_6	6,3	6,5	5,2	5,3	7,3	6,8	6,9	7,9	7,5	7,0	7,9	7,6	5,2
Abt_7	5,9	5,6	6,0	6,8	6,6	7,1	7,4	6,9	6,9	5,4	6,6	8,0	5,2
Abt_8	7,2	5,4	7,9	6,6	5,9	5,5	6,6	6,3	7,2	7,2	7,9	6,1	5,7
Summe:	54,0	50,9	52,5	51,0	48,7	51,1	51,9	53,8	52,0	51,7	54,1	56,4	48,2

Abb. 2.1 Monatsergebnisse mit Überschrift und unten angefügter Summenzeile

tenerfassung gemeinsam mit dem Datenbestand *auf einen Blick sichtbar*. Abbildung 2.1 zeigt beispielhaft die monatlichen Umsätze von lediglich *acht Abteilungen*, darunter ließ man, wie üblich, die *monatlichen Gesamtumsätze* berechnen und eintragen. Alles ist auf einen Blick verfügbar.

Wenn der kleine Bildschirm aber nicht ausreicht, dann müsste man zwischen Daten und Ergebnissen zeitraubend hin- und herwechseln, vielleicht müssten die Auswertungsergebnisse sogar in einem anderen Tabellenblatt angezeigt werden.

Denken wir beispielsweise an eine Anzahl von mehr als dreißig Abteilungen: Dann sieht man auf dem Bildschirm *entweder die Überschrift* und damit die Erläuterung zum Datenbestand – *oder man sieht die Summen*, die Analyseergebnisse. Hier empfiehlt es sich, die Überschrift erst in die *dritte Zeile* einzutragen, und die *Summen nach oben* zu holen (siehe Abb. 2.2).

Die leere zweite Zeile wurde zum einen aus optischen Gründen eingefügt, zum anderen dient sie der *Separierung des Datenbestandes*.

Summe:	54,0	50,9	52,5	51,0	48,7	51,1	51,9	53,8	52,0	51,7	54,1	56,4	48,2
Abteilung	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez	Jan
Abt_1	6,0	7,3	7,6	6,1	6,2	6,3	7,2	8,0	5,9	6,3	7,3	7,5	6,9
Abt_2	6,7	5,4	6,2	5,4	5,6	6,2	6,1	7,9	5,4	5,8	6,6	5,1	7,7
Abt_3	7,4	7,5	6,6	7,8	5,1	6,7	5,2	6,5	6,0	5,8	5,2	6,3	6,4
Abt_4	7,7	7,3	7,0	5,8	5,2	6,6	7,0	5,1	7,4	6,6	5,2	8,0	5,4
Abt_5	6,8	5,9	6,0	7,2	6,8	5,9	5,5	5,2	5,7	7,6	7,4	7,8	5,7
Abt_6	6,3	6,5	5,2	5,3	7,3	6,8	6,9	7,9	7,5	7,0	7,9	7,6	5,2
Abt_7	5,9	5,6	6,0	6,8	6,6	7,1	7,4	6,9	6,9	5,4	6,6	8,0	5,2
Abt_8	7,2	5,4	7,9	6,6	5,9	5,5	6,6	6,3	7,2	7,2	7,9	6,1	5,7

Abb. 2.2 Monatsergebnisse mit Bilanzen in der ersten Zeile

2.1.2.2 Separierung des Datenbestandes

Viele Anleitungen für Excel-Maßnahmen beginnen mit der Aufforderung, den Tabellenkursor in irgendeine Zelle des Datenbestandes zu setzen – und dann weitere Bedienhandlungen vorzunehmen.

- **Wichtiger Hinweis** Excel ist in der Lage, selbst die *Grenzen eines Datenbestandes* erkennen zu können. Das erspart zeitraubendes Markieren. Allerdings setzt das voraus, dass der Datenbestand *separiert* ist:

Ein Datenbestand ist *separiert*, wenn er durch *mindestens eine Leerzeile und Leerspalte* von anderen Inhalten des Tabellenblattes getrennt ist.

Ist der Datenbestand separiert, dann ist z. B. das Sortieren ganz einfach: Man setze den Tabellenkursor in *irgendeine Zelle derjenigen Spalte des Datenbestandes*, nach der sortiert werden soll (dabei aber *nichts markieren*). Dann ist in der Registerkarte **Start** und der Gruppe **Bearbeiten** die passende Schaltfläche zum Sortieren auszuwählen:

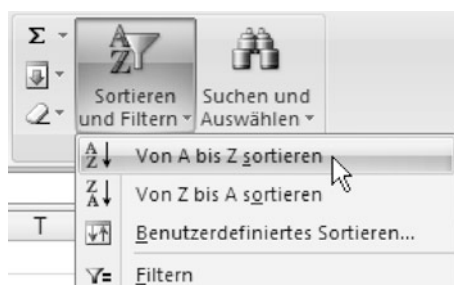


Abbildung 2.3a zeigt ein mögliches *brauchbares Sortierergebnis*, das wegen des offensichtlich *separierten Datenbestandes* (die Zeilen 2 und 12 sowie die Spalte O sind leer) problemlos erhalten werden konnte. Es entstand, indem der Tabellenkursor in irgendeine Zelle unter B3 gesetzt wurde und die aufsteigende Sortier-Reihenfolge verlangt wurde. Da der berühmte *Anfänger-Fehler*, nämlich die Markierung der Spalte A (siehe auch Kap. 1, Abschn. 1.7.1), *nicht* stattgefunden hat, wurde tatsächlich der *gesamte Datenbestand* so umgeordnet, dass in der Januar-Spalte die *aufsteigende Sortierung* erkennbar ist. Der Datenbestand wurde also durch diese Art der Sortierung *nicht zerstört*.

Lässt man aber die Leerzeile 2 jedoch weg, vergisst man also die Separierung der Summenzeile, dann kann es Konflikte geben: In Abbildung 2.3b wird ein dann denkbares „Sortier-Ergebnis“ gezeigt: Da Excel die *Grenzen des zu sortierenden Datenbestandes* nicht erkennen konnte, wurden Überschriften und – vor allem – die *Summenzeile* sinnlos mit in die Sortierung einbezogen. Dabei wurden auch die eingetragenen Summenformeln verändert, so dass die Summenzeile plötzlich nur noch Nullen enthält.

a

Summe:	54,0	50,9	52,5	51,0	48,7	51,1	51,9	53,8	52,0	51,7	54,1	56,4	48,2
Abteilung	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez	Jan
Abt 4	7,7	7,3	7,0	5,8	5,2	6,6	7,0	5,1	7,4	6,6	5,2	8,0	5,4
Abt 3	7,4	7,5	6,6	7,8	5,1	6,7	5,2	6,5	6,0	5,8	5,2	6,3	6,4
Abt 8	7,2	5,4	7,9	6,6	5,9	5,5	6,6	6,3	7,2	7,2	7,9	6,1	5,7
Abt 5	6,8	5,9	6,0	7,2	6,8	5,9	5,5	5,2	5,7	7,6	7,4	7,8	5,7
Abt 2	6,7	5,4	6,2	5,4	5,6	6,2	6,1	7,9	5,4	5,8	6,6	5,1	7,7
Abt 6	6,3	6,5	5,2	5,3	7,3	6,8	6,9	7,9	7,5	7,0	7,9	7,6	5,2
Abt 1	6,0	7,3	7,6	6,1	6,2	6,3	7,2	8,0	5,9	6,3	7,3	7,5	6,9
Abt 7	5,9	5,6	6,0	6,8	6,6	7,1	7,4	6,9	6,9	5,4	6,6	8,0	5,2

b

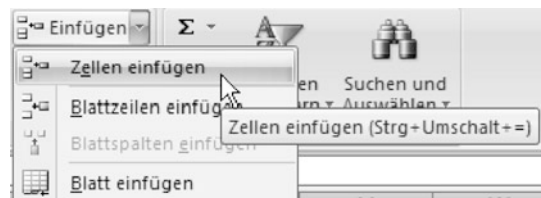
Abteilung	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez	Jan
Summe:	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Abt_4	7,7	7,3	7,0	5,8	5,2	6,6	7,0	5,1	7,4	6,6	5,2	8,0	5,4
Abt 3	7,4	7,5	6,6	7,8	5,1	6,7	5,2	6,5	6,0	5,8	5,2	6,3	6,4
Abt 8	7,2	5,4	7,9	6,6	5,9	5,5	6,6	6,3	7,2	7,2	7,9	6,1	5,7
Abt 5	6,8	5,9	6,0	7,2	6,8	5,9	5,5	5,2	5,7	7,6	7,4	7,8	5,7
Abt 2	6,7	5,4	6,2	5,4	5,6	6,2	6,1	7,9	5,4	5,8	6,6	5,1	7,7
Abt 6	6,3	6,5	5,2	5,3	7,3	6,8	6,9	7,9	7,5	7,0	7,9	7,6	5,2
Abt 1	6,0	7,3	7,6	6,1	6,2	6,3	7,2	8,0	5,9	6,3	7,3	7,5	6,9
Abt 7	5,9	5,6	6,0	6,8	6,6	7,1	7,4	6,9	6,9	5,4	6,6	8,0	5,2

Abb. 2.3 Sortierung (Januar-Ergebnisse absteigend) **a** bei separiertem Datenbestand, **b** bei nicht separiertem Datenbestand

Fassen wir zusammen:

- Natürlich spricht nichts dagegen, bei der Datenerfassung erst einmal die Überschrift in die Zeile 1 einzutragen.
- Ist der Datenbestand aber umfangreich, dann empfiehlt es sich, durch *Einfügen von leeren Zeilen* die Überschrift mit dem Datenbestand nach unten zu verschieben, um einerseits die *Separierung des Datenbestandes* zu erreichen und andererseits die *Analyseergebnisse* nicht erst lange suchen zu müssen.

Um eine *leere Zeile einzufügen*, markiert man die Zeile, die nach unten verschoben werden soll, zuerst mit Mausklick auf die *Zeilennummer*, danach muss in der Registerkarte *Start* in der Gruppe *Zellen* auf *Einfügen* geklickt werden:



Oder man kann mit Hilfe der Tastenkombination [↑] + [Leertaste] die Zeile markieren, die nach unten verschoben werden soll. Die Verschiebung (und damit das Einfügen einer leeren Zeile darüber) erfolgt dann mit [Strg] + [=].

2.1.2.3 Überschriften

Angenommen, ein größerer Datenbestand wird erfasst. Was trägt man für *Überschriften* ein? Sinnvoll wären möglichst *aussagekräftige Vokabeln* (manchmal *sprechende Überschriften* genannt), damit man erkennt, welchen sachlichen Inhalt die jeweilige Spalte hat.

Doch damit kann man *sehr viel Platz verschenken*.

Betrachten wir zum Beispiel nur die ersten fünf Spalten einer mit Excel erfassten Auswertung einer Fragebogenaktion, die aber insgesamt nicht weniger als 168 Angaben enthält:

Fragebogennummer	Stadtgebiet	Eigentümer	Geschosszahl	Wohnungsgröße
1	1	2	1	1
2	1	2	1	1
3	1	2	1	1
43	1	2	1	3
46	1	2	1	3

Sicher sind die Spaltenüberschriften *sprechend* – aber auf dem kleinen Bildschirm ist weder von den restlichen Spalten noch vom Datenbestand etwas zu erkennen. Die erste Möglichkeit, diese *Platzvergeudung* zu verhindern, um möglichst viel vom Datenbestand zu sehen, besteht in der Verwendung von *Abkürzungen* anstelle aussagekräftiger Vokabeln. Dann müsste man allerdings dem Betrachter der Tabelle ein Verzeichnis der Abkürzungen mitliefern.

Da die Verwendung derartiger Abkürzungen nicht sonderlich zum *Verständnis des sachlichen Inhalts des Datenbestandes* beiträgt, sollen nun zwei andere Möglichkeiten geschildert werden, wie man trotz aussagekräftiger Vokabeln doch einen recht großen Ausschnitt des Datenbestandes auf dem Bildschirm erkennen kann. Man erreicht das entweder durch *Umbruch in der Überschriftenzeile* oder durch *schräge* oder sogar *senkrechte Ausrichtung der Überschriften*.

Ein *Umbruch* in der Überschriftenzeile findet dann statt, wenn der Zelleninhalt innerhalb der Zelle auf *mehrere Zeilen* verteilt wird.

Nach der *Markierung der gesamten Überschriftenzeile* (durch Anklicken der Zeilennummer mit der Maus) kann mit folgendem Vorgehen der *Umbruch* in der *gesamten Zeile* erlaubt werden:

Entweder wird das Fenster *Zellen formatieren* mit der Tastenkombination [Strg]+[1] angefordert, dann ist in der Registerkarte *Ausrichtung* in der Checkbox *Zeilenumbruch* der Haken zu setzen.

Oder es wird in der Gruppe *Ausrichtung* der Registerkarte *Start* das entsprechende Sinnbild angeklickt:



Hinweis Soll nur in einer *einzigsten Zelle* umgebrochen werden, dann kann das auch schnell mit der Tastenkombination [Alt]+[↵] erreicht werden. So könnte man zum Beispiel mittels Zeilen→[Alt]+[↵]→umbruch die Überschrift *Zeilenumbruch* zweizeilig eintragen.

- **Wichtiger Hinweis** Der *Umbruch sprechender Überschriften* trägt sehr viel zur besseren Lesbarkeit bei, und – nicht zu vergessen – man kann auch einen viel größeren Ausschnitt aus dem Datenbestand auf dem Bildschirm sehen.

Nach geschickt eingefügten Minuszeichen und Verkleinerung der Spaltenbreite zeigen sich die Überschriften im Umbruch:

Frage- bogen- nummer	Stadt- gebiet	Eigen- tümer	Ge- schoss- zahl	Woh- nungs- größe
1	1	2	1	1
2	1	2	1	1
3	1	2	1	1
43	1	2	1	3
46	1	2	1	3

Man kann eine *bessere Übersichtlichkeit* aber auch erreichen, indem man die Überschriftenzeile markiert (wie oben beschrieben) und die *Überschriften senkrecht oder schräg* stellt.

Dazu wird wieder mit der Tastenkombination [Strg]+[1] das Fensters Zellen formatieren angefordert – oder es wird im Registerblatt Start in der Gruppe Ausrichten das passende Sinnbild angeklickt. Lässt man anschließend wieder die Spalten auf *optimale Breite* einstellen, dann erhält man auf dem Bildschirm wohl den relativ besten Überblick über den Datenbestand.

Bei einer *Präsentation der Daten* sollte man allerdings wohl grundsätzlich darauf verzichten, die Überschriften *senkrecht* zu stellen. Denn kein Publikum wird begeistert sein, wenn es den Kopf auf die Schulter legen muss, um überhaupt etwas lesen zu können.

2.1.2.4 Speicherung und Möglichkeiten der Dateneingabe

Ist die Zeile der Überschriften entweder mit Abkürzungen oder mit Umbruch oder senkrechter Anordnung Platz sparend vorbereitet, sollte die Excel-Tabelle bereits das erste Mal gespeichert werden: Zur Speicherung kann wahlweise eine der drei Tastenkombinationen [Strg]+[F12] oder [Strg]+[s] oder [Alt]+([d]→[s]) oder der *Klick auf das Diskettensymbol* links oben genutzt werden. Anschließend erfolgt die *Erfassung der Daten*.

Die Daten können entweder

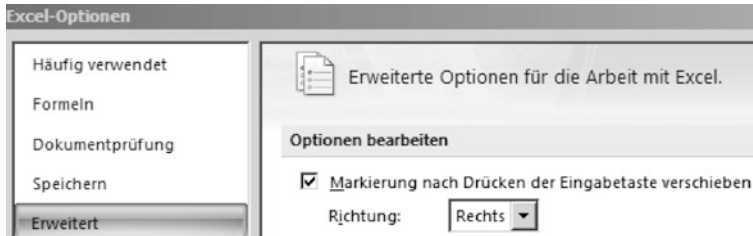
- unmittelbar in die Tabelle eingetragen werden oder
- mit Hilfe der *Maske* erfasst werden.

Letzteres ist allerdings nur möglich, wenn die Spaltenanzahl nicht zu groß und die Überschrift bereits separiert ist.

2.1.2.5 Dateneingabe direkt in die Tabelle

Vor Beginn der *tabellarischen Dateneingabe* sollte man zuerst dafür sorgen, dass bei Bestätigung der Eingabe mit der Eingabetaste [↵] der Tabellenkursor *automatisch nach rechts* weiterwandert:

Dazu wird zuerst – sofern vorhanden – die runde Schaltfläche Office ganz links oben angeklickt und Excel-Optionen gewählt. Gibt es diese runde Schaltfläche nicht, dann kann man über Datei→Optionen ebenfalls zum Fenster Excel-Optionen kommen. Wählt man dort Erweitert, dann öffnet sich ein neues Fenster und man kann sofort auswählen, wie sich die Markierung nach Drücken der Eingabetaste verschieben soll:



Nun kann man die Maus weglegen und schnell und sicher arbeiten, indem jede Eingabe mit der Taste [↵] bestätigt wird. Bei *reinen Zahlenerfassungen* sollte man dafür zweckmäßig den Ziffernblock rechts an der Tastatur und die dort vorhandene Taste [Enter] benutzen.

- **Wichtiger Hinweis** Wenn der Ziffernblock defekt zu sein scheint und keine Zahlenangaben annimmt, dann prüfe man, ob die Kontroll-Lampe NumLock leuchtet. Ist das nicht der Fall, dann muss die Taste [Num] (im Ziffernblock links oben) betätigt werden.

Fehlt der separate Ziffernblock, wie allgemein bei Laptops üblich, dann kann dort ein Teil der Tastatur mittels [Fn] + [Num] zum Ziffernblock gemacht werden.

Für *Datumseingaben* sollte man ebenfalls den Ziffernblock benutzen: Excel erkennt nämlich insbesondere Eingaben der Art 12/4/11 (der Schrägstrich befindet sich auf dem Ziffernblock zwischen den Tasten [Num] und [*]) als deutsches Datum und trägt es in der Tabelle entsprechend der eingestellten Darstellung ein.

Ist das Ende einer Zeile erreicht, so wird zuerst mit [↵] in die nächste Zeile gewechselt und mit [Strg] + [←] dort zum Zeilenanfang gesprungen.

- **Man beachte** Mit Hilfe der beiden Tastenkombinationen [Strg] + [F12] oder [Strg] + [s] kann (und sollte) jederzeit sichernd gespeichert werden.

2.1.2.6 Dateneingabe mit der Maske

Falls die *Anzahl der Spalten nicht zu groß* und der *Datenbestand separiert* ist, dann kann man die Daten mit Hilfe der *Maske* erfassen.

Die *Anforderung der Eingabemaske* erfolgt so, dass zuerst der Tabellenkursor in irgendeine Zelle der *Überschriftenzeile* des Datenbestandes gesetzt wird. Dann folgt die Tastenkombination $[Alt] + ([n] \rightarrow [m])$.

Abbildung 2.4 zeigt eine solche Eingabemaske, die rechts neben den Überschriften die Eingabefenster für die Daten enthält.

In äußerst übersichtlicher Form können nun die Belegungen eingetragen werden, wobei jetzt als Bestätigung der Eingabe die *Tabulator-Taste* (links an der Tastatur) verwendet werden sollte.

Für *Rechtshänder* ist das ideal: Die linke Hand auf der *Tabulator-Taste* bestätigt die Eingaben und veranlasst gleichzeitig das Fortschreiten zum nächsten Eingabefenster – die *Enter-Taste* wäre dafür nicht geeignet.

The screenshot shows a window titled 'Tabelle1' with a list of labels and input fields on the left, and a vertical scroll bar and a column of buttons on the right. The labels and their corresponding input fields are:

- Erhebungsnr.:
- Stadtgebiet:
- Eigentümer:
- Geschosszahl:
- Wohnungsgröße:
- Wahrnehmung der Wohnsituation:
- Seit wann in Wohnung:
- Wer beeinflusst Wohlempfinden:
- Was beeinflusst Wohlempfinden:
- Lebensqualität : zufrieden:
- Lebensqualität: beunruhigt:
- Lebensqualität: abgesichert:

On the right side, there is a vertical scroll bar and a column of buttons:

- Neuer Datensatz
- Neu
- Löschen
- Wiederherstellen
- Vorherigen suchen
- Weitersuchen
- Kriterien
- Schließen

Abb. 2.4 Form einer Eingabemaske

Kombiniert man die *Umschalttaste* [↑] mit der *Tabulatortaste*, dann bewegt man sich zurück zum vorigen Eingabefenster (nach „oben“) – das ist wichtig, wenn man sich vertippt hat und korrigieren muss.

Ist das letzte Fenster bearbeitet (entweder durch Eintrag oder durch Übergehen), so führt die Tabulatortaste zur Aktivierung der Schaltfläche [Neu]; man erkennt dies durch eine farbliche oder umrandete Hervorhebung dieser Schaltfläche:



Nun reicht ein Druck auf die Leertaste oder [Enter], damit wird [Neu] betätigt, und für die nächste Zeile der Tabelle steht dann wieder eine *leere Maske* bereit. Das erste Eingabefenster ist dann schon ausgewählt, mit Zahlenblock und Tabulatortaste kann weiter eingegeben werden. Natürlich hätte man für das alles auch die Maus benutzen können – aber damit verlangsamt man die Datenerfassung beträchtlich.

Will man mit Excel schnell arbeiten, dann verwende man, wenn möglich, die Tastatur.

2.1.3 Eindimensionales Datenmaterial

Eindimensionales Datenmaterial, d. h. eine einfache Datenreihe, sollte man *stets in der ersten Spalte* erfassen. Da nun rechts daneben genug Platz für spätere Analysewerte ist, kann in diesem Falle in der Zelle A1 mit dem Dateneintrag begonnen werden. Eine *Überschrift* ist in diesem Fall nicht zwingend nötig. Vor der Dateneingabe sollte jetzt jedoch die *Bewegung des Tabellenkursors nach unten* eingestellt werden.

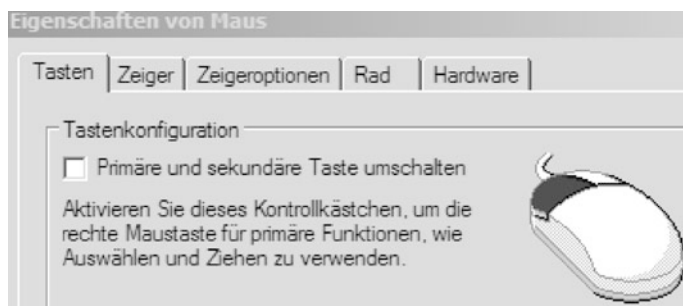
2.1.4 Umgang mit dem Datenmaterial

Es kann hier wohl vorausgesetzt werden, dass der Umgang mit Excel mit Hilfe der *Maus* oder anderen taktilen Bedienelementen allgemein bekannt ist. Muss man jedoch mit einem größeren Datenbestand arbeiten, dann wird schon die Markierung von Teilen des Datenbestandes oder auch nur die Feststellung, wo denn nun die Datenmenge endet, zu einem zeitlichen Geduldsspiel. Hier sei insbesondere auf den Abschn. 1.8.2 im Kap. 1 hingewiesen, in dem viele nützliche Tastenkombinationen aufgelistet sind.

Bevor aber einige ausgesprochen wertvolle Tastenkombinationen mitgeteilt werden, mit denen schnell in großen Datenbeständen gearbeitet werden kann, soll für alle *Linkshänder* beschrieben werden, wie die Funktion der beiden Maustasten vertauscht werden kann, so dass die Maus *links neben der Tastatur* liegen kann.

Diese Einstellung wird im Betriebssystem vorgenommen und gilt damit auch für alle anderen Anwendungen, ist dann nicht auf Excel beschränkt.

Man finde dafür die *Systemsteuerung* und mit ihrer Hilfe das *Einstellungsfenster für die Maus*. Anschließend braucht man nur mitzuteilen, ob die Maus links oder rechts von der Tastatur liegen soll. Manchmal wird auch gefragt, ob man Rechts- oder Linkshänder sei, oder ob – wie in folgendem Beispiel – die primäre und die sekundäre Maustaste vertauscht werden sollen:



Hat man sich als Linkshänder dafür entschieden, die Maus links neben die Tastatur zu legen, so wird dann die rechte Maustaste die Haupt-Maustaste (primär) und die linke Maustaste die Neben-Maustaste (sekundär). In vielen Büchern (sicher ist auch das vorliegende Buch nicht ganz frei davon) wird inkorrekt eigentlich immer nur die (Rechtshänder-)Standardsituation berücksichtigt:

- **Wichtiger Hinweis** Man spricht allgemein von der „linken Maustaste“ und meint die Haupttaste, und wenn man von der „rechten Maustaste“ spricht, denkt man an die Taste mit der Nebenbedeutung.

2.2 Kontrollen des Datenbestandes

2.2.1 Nützliche Auskunftsfunktionen

Die Auskunftsfunktion `=ANZAHL2 (. . .)` teilt mit, wie viele Zellen in dem angegebenen Bereich überhaupt *eine Belegung* (welcher Art auch immer) besitzen. Die Auskunftsfunk-

6	7	8	6	6	6	7	8	6	6	7	8	7	6	6	7	=ANZAHL(A1:P8)
7	A	6	5	6	6	Z	8	5					5	6	7	=ANZAHL2(A1:P8)
7	8	C	8	5	7	Y	7	N	6				6	6	5	
8	7	7	6	Q	7	X	5	V	7				7	7	5	
7	6	6	7	7	6	6	5	6	R	7	8	6	6	8	7	
6	7	R	5	7	7	7	8	8	7	8	8	5	8			
6	6	6						5	7	8	5	7				
7	5	8	7	6	6	7	6	7	7	8	6	6	7			

6	7	8	6	6	6	7	8	6	6	7	8	7	6	6	7	96
7	A	6	5	6	6	Z	8	5					5	6	7	106
7	8	C	8	5	7	Y	7	N	6				6	6	5	
8	7	7	6	Q	7	X	5	V	7				7	7	5	
7	6	6	7	7	6	6	5	6	R	7	8	6	6	8	7	
6	7	R	5	7	7	7	8	8	7	8	8	5	8			
6	6	6						5	7	8	5	7				
7	5	8	7	6	6	7	6	7	7	8	6	6	7			

Abb. 2.5 Anwendung der beiden Auskunftsfunktionen

tion =ANZAHL (. . .) teilt dagegen nur speziell mit, wie viele der belegten Zellen in dem angegebenen Bereich einen *numerischen Inhalt* haben. Datums- und Uhrzeitangaben zählen dabei auch als numerische Inhalte.

Betrachten wir zum Beispiel in Abb. 2.5 einen Datenbestand, der mit Hilfe dieser beiden Auskunftsfunktionen grundsätzlich analysiert werden soll. Offensichtlich beginnt dieser Datenbestand in der Zelle A1 – aber wo endet er? Es könnte doch sein, dass sich in weit entfernten Zellen, die auf dem Bildschirm nicht erscheinen, noch Daten befinden.

Um das *tatsächliche Ende eines Datenbestandes* zu erfahren, setze man den Tabellenkursor in die Zelle A1 und wähle dann die Tastenkombination [Strg] + [Ende]. Dann springt der Tabellenkursor auf die tatsächlich letzte Zeile, die jetzt (oder auch irgendwann früher) einen Inhalt besitzt oder besessen hat.

Im Fall unseres Beispiels springt der Tabellenkursor tatsächlich in die Zelle P8, dort endet also der zu analysierende Datenbestand. Der Datenbestand befindet sich mit Sicherheit im Innern des Rechteckbereiches A1 : P8.

Dieser Bereich enthält insgesamt 16 · 8 = 128 Zellen (weil die Spalte P die 16-te Spalte ist). Damit ist eine sichere Obergrenze für die Anzahl der vorhandenen Daten gefunden. Weiter soll herausgefunden werden, wie viele der Zellen in dem Bereich überhaupt belegt sind, davon interessiert speziell die Anzahl der Zellen mit numerischen Einträgen. Außerdem soll mitgeteilt werden, wie viele Zellen leer sind.

Dazu werden in zwei Zellen außerhalb des Bereiches die beiden Auskunftsfunktionen eingetragen, in Abb. 2.5 erkennt man sie in den Zellen Q1 und Q2.

Die mitgeteilten Zahlen lassen erkennen: Von den oben genannten 128 Zellen sind nur 106 überhaupt belegt, also sind 22 Zellen leer. 96 der belegten Zellen enthalten numerische Werte. Also enthalten $106 - 96 = 10$ Zellen nichtnumerische Belegungen.

Übrigens hätte man sich selbst die kleine Rechnung sparen können, denn es gibt außerdem die Auskunftsfunction `=ANZAHLLEEREZELLEN(. . .)`, die – das sagt schon ihr Name – die Anzahl der nicht belegten Zellen im Bereich mitteilt.

2.2.2 Erfasste Merkmalswerte und das Filtern

Kein Mensch ist vollkommen. Fehler können immer auftreten, natürlich auch bei der Eingabe von Daten und ihrer Speicherung in einer Excel-Tabelle. Wie schnell vertippt man sich, insbesondere, wenn nichtnumerische Daten einzugeben sind und der bequeme Zahlenblock der Tastatur nicht benutzt werden kann (oder nicht vorhanden ist, wie bei vielen tragbaren Computern heutzutage).

Wie kann man Fehler in einem erfassten Datenbestand finden? Eine Möglichkeit dazu besteht in der Anwendung des Excel-Werkzeuges `Filter`.

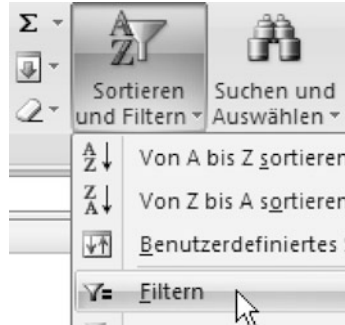
Nehmen wir folgende Situation an: In Auswertung einer Fragebogenaktion steht in Spalte C eines großen Datenbestandes das Geschlecht, eigentlich sollte dort nur „m“ für männlich und „w“ für weiblich stehen. Doch man kann sich bei der Erfassung sehr schnell vertippen, zumal auf jeder Tastatur m und n direkt nebeneinander stehen. Also kann es passieren, dass – wie in der folgenden Beispiel-Tabelle – anstelle vom „m“ auch einmal „n“ eingetippt wurde.

Name	Vorname	Geschlecht	Angabe_1	Angabe_2
Name_1	Vorname_1	w		
Name_2	Vorname_2	w		
Name_3	Vorname_3	m		
Name_4	Vorname_4	m		
Name_5	Vorname_5	n		
Name_6	Vorname_6	w		
Name_7	Vorname_7	w		

Hier ist dieser offensichtliche Tippfehler in Zeile 6 sofort zu sehen – aber wie fände man ihn, wenn er in Zeile 667 stehen würde?

Nehmen wir an, es seien mehr als tausend Fragebögen erfasst worden. Wie kann man schnell herausfinden, ob in dieser Spalte tatsächlich nur korrekt „m“ oder „w“ steht, ohne den Datenbestand mühsam prüfend durchsehen zu müssen?

Dazu benutzt man die Excel-Leistung FILTER. Sie wird – nachdem der Tabellenkursor in irgendeine Zelle des Datenbestandes gesetzt wurde – mit Hilfe der Gruppe Bearbeiten der Registerkarte Start angefordert:

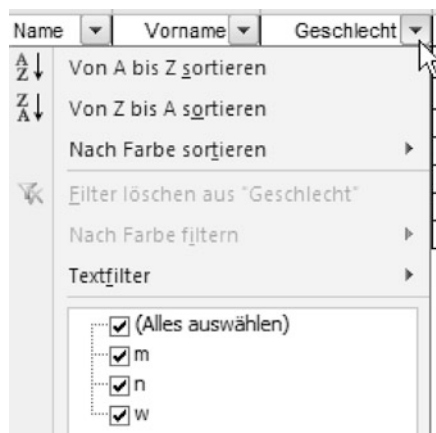


Ist der Filter eingeschaltet, dann erscheinen dann in jeder Überschriftenzelle rechts unten Schaltflächen mit kleinen Dreiecken:

	A	B	C	D	E
1	Name ▾	Vorname ▾	Geschlecht ▾	Angabe_1 ▾	Angabe_2 ▾
2	Name_1	Vorname_1	w		

Wird nun das Dreieck in der Spalte, in der man die Fehler finden möchte, angeklickt, dann öffnet sich ein Informationsfenster.

Im Informationsfenster werden in lexikografischer Folge alle in der betrachteten Spalte des Datenbestandes *überhaupt auftretenden Merkmalswerte* angezeigt.



Hier also unschwer zu erkennen, dass sich beim Eintippen *mindestens einmal ein* „n“ eingeschlichen hat – vielleicht auch öfter. Wie kann man das herausbekommen?

Ganz einfach: Sorgt man dafür, dass sich nur bei dem falschen Merkmalswert in der Merkmalswertliste ein Haken befindet, so blendet Excel alle Zeilen des Datenbestandes aus, in denen dieser Wert *nicht* steht – folglich werden damit nur die falschen Zeilen herausgefiltert.

Sehen wir uns das Ergebnis an: Nur in der Zeile 6 ist dieser Tippfehler aufgetreten. Gäbe es mehrere Zeilen mit „n“, dann wären auch sie angezeigt worden. So findet man alle Zeilen des Datenbestandes mit falschen Einträgen.

	A	B	C	D	E
1	Name ▾	Vorname ▾	Geschlecht ▾	Angabe_1 ▾	Angabe_2 ▾
6	Name_5	Vorname_5	n		

Der Filter wird über die Gruppe *Bearbeiten* wieder ausgeschaltet.

- **Wichtiger Hinweis** Mit Hilfe des Filters lässt sich auch die Frage, wie viele Merkmalswerte überhaupt vorhanden sind und wie sie heißen, schnell und elegant beantworten.

Es seien zum Beispiel in einem Datenbestand „Personal“ Hunderte von Namen und Vornamen und Wohnorte erfasst. Durch einfaches Filtern könnte sofort die Frage beantwortet werden, in welchen Orten die erfassten Personen wohnen:



Es gibt also nur fünf verschiedene Wohnorte diesem Datenbestand. Man erkennt auch wieder gut die lexikografische (Wörterbuch-)Anordnung der aufgelisteten Merkmalswerte.

Mit ein wenig Übung lernt man das Filtern in Excel-Tabellen schätzen.

Es muss aber noch einmal darauf hingewiesen werden, dass das Filtern in Excel-Tabellen nur unter *zwei wichtigen Voraussetzungen* erfolgreich durchführbar ist.

- Der Datenbestand muss, notfalls durch Einfügen von leeren Zeilen und/oder leeren Spalten, von anderen Einträgen in der Excel-Tabelle separiert sein.
- Der Tabellenkursor muss sich in einer Zelle im Inneren des Datenbestandes befinden, *bevor* der Filter eingeschaltet wird.

Betrachten wir noch ein kleines Beispiel für den Nutzen des Filters: Angenommen, in einer Mittelstands-Firma arbeiten ca. 100 Angestellte, und die Sekretärin des Direktors führt eine (Excel-)Liste mit Namen und Geburtsdaten. Sehen wir uns den Anfang dieser Liste an:

Name	Vorname	Geburtsdatum
Göxfek	Dew	25.12.1964
Däphin	Vube	22.10.1968
Röfköguv	Duk	03.11.1989

Jeweils am Monatsende möchte der Chef wissen, wer im Folgemonat Geburtstag hat. Die Sekretärin setzte den Tabellenkursor in irgendeine Zelle der Spalte C und startet den Filter. Doch damit kann sie den Auftrag des Chefs nicht erfüllen, denn der Filter bietet nur die Möglichkeit, die Geburtsjahre auszuwählen. Was ist zu tun?

Hier helfen die drei *Excel-Datumsfunktionen* =MONAT(...), =TAG(...) und =JAHR(...):

Name	Vorname	Geburtsdatum	Monat	Tag	Jahr
Göxfek	Dew	25.12.1964	=MONAT(C2)	=TAG(C2)	=JAHR(C2)
Däphin	Vube	22.10.1968	=MONAT(C3)	=TAG(C3)	=JAHR(C3)
Röfköguv	Duk	03.11.1989	=MONAT(C4)	=TAG(C4)	=JAHR(C4)

Sie lösen aus dem Datum die entsprechenden Angaben heraus:

Name	Vorname	Geburtsdatum	Monat	Tag	Jahr
Göxfek	Dew	25.12.1964	12	25	1964
Däphin	Vube	22.10.1968	10	22	1968
Röfköguv	Duk	03.11.1989	11	3	1989

Nun kann der Filter auf die Spalte D angewandt werden, der Haken in der Monatsliste wird dann auf den gewünschten Monat gesetzt.

Und wenn *anschließend* mit Hilfe des Filters in Spalte E die aufsteigende Sortierung angefordert wird, kann die Sekretärin ihrem Chef schnell die gewünschte Übersicht vorlegen.

Wenn sie noch die Differenz zwischen dem aktuellen Jahr und dem Geburtsjahr hinzufügen lässt, sieht der Chef auch sofort, wie alt sein Mitarbeiter werden wird ...

2.2.3 Prüfen von Minimum und Maximum

Mit Hilfe der beiden Excel-Funktionen =MIN (. . .) bzw. =MAX (. . .) lässt sich für alle Spalten von erfassten Datenmengen, in denen sich *numerische Werte* befinden, eine weitere *Überprüfung der Dateneingabe* vornehmen. So sollten im folgenden Datenbestand nur die Zensuren 1 bis 6 enthalten sein – das kann man mit Hilfe dieser beiden Funktionen prüfen:

			=MIN(C5:C1000)
			=MAX(C5:C1000)
Name	Vorname	Zensur	
Modlöz	Vixu	4	

			1
			6
Name	Vorname	Zensur	
Modlöz	Vixu	4	

Da hier als Minimum 1 und als Maximum 6 angezeigt wird, scheinen Tippfehler mit unsinnigen Zahlenangaben nicht vorzuliegen. Der angegebene Bereich C5 : C1000 ist sehr großzügig bemessen; das schadet aber nicht und entlastet von der Notwendigkeit, die Formeln jedes Mal ändern zu müssen, wenn eine neue Zeile hinzugefügt worden ist. Man beachte bei der Interpretation der Ergebnisse der Funktionen =MIN (. . .) und =MAX (. . .) , dass beide Extremwerte aber genauso wären, wenn in einer oder mehreren Zellen versehentlich ein *Buchstabe*, z. B. „x“, stehen würde. Er würde dann aber wieder beim Filtern (wie oben beschrieben) entdeckt.

2.3 Absolute Häufigkeiten

Wenn im Datenbestand nur *einige wenige Merkmalswerte* auftreten (z.B. Zensuren), dann kann die Anzahl des Auftretens jedes Merkmalswertes mit Hilfe der leistungsfähigen Excel-Funktion =ZÄHLENWENN (. . . ; . . .) schnell ermittelt werden kann.

Dagegen ist bei *Messergebnissen*, bei denen eine unüberschaubare Vielzahl an Merkmalswerten auftreten kann (z. B. Körpergröße, Gewicht, Temperatur, Gehalt) das Auszählen erst nach erfolgter *Klassenbildung* (siehe Abschn. 2.3.2) sinnvoll. Hierfür stellt

Excel vielfältige Hilfsmittel bereit. Bei *mehrdimensionalem Datenmaterial* tritt darüber hinaus die Frage auf, wie oft gewisse *Paare von Merkmalswerten* (z. B. gewisse Zensurenkombinationen zweier Fächer) vorliegen. Diese Frage kann mit Hilfe von Pivot-Tabellen beantwortet werden, zu deren Erzeugung der Pivot-Tabellen-Assistent genutzt wird.

2.3.1 Einfaches Abzählen mit ZÄHLENWENN

Excel ermöglicht mit Hilfe der Funktion `=ZÄHLENWENN (. . . ; . . .)` das einfache Abzählen, wenn es nur einige wenige Merkmalswerte gibt.

Dabei ist es prinzipiell gleichgültig, ob die Merkmalswerte in *numerischer Form* (d. h. als Zahlen, Datums- oder Zeitangaben) oder in *nichtnumerischer Form* (d. h. als einzelne Zeichen oder Zeichenfolgen) vorliegen.

- **Wichtiger Hinweis** Es gibt mit den *Wahrheitswerten* noch eine dritte Form von Merkmalswerten, die jedoch in diesem Buch keine Rolle spielen, so dass auf sie nicht näher eingegangen werden soll.

2.3.1.1 Abzählen von numerischen Einträgen

Wiederholen wir: Man erkennt einen *numerischen Eintrag* daran, dass er unmittelbar nach seiner Eingabe, und bevor keine speziellen Formatierungen vorgenommen wurden, rechtsbündig (d. h. am rechten Zellenrand anliegend) eingetragen ist. Nicht-numerische Daten (Text, Zeichenfolgen) dagegen werden von Excel *linksbündig* eingetragen, und *Wahrheitswerte* mittig.

Sehen wir uns die Wirkung der Funktion `=ZÄHLENWENN (. . . ; . . .)` am Beispiel der Abb. 2.6 an: Im Bereich B2 : B10 sind *numerische Merkmalswerte* von neun Abteilungen eingetragen, darüber befindet sich die Überschrift. Rechts neben dem Datenbestand ist ausreichend Platz für das Eintragen von Auswertungsergebnissen; also kann man mit dem Erfassen des Datenbestandes in der Zeile 1 beginnen. Die Spalte C sollte aber unbedingt leer bleiben, damit der Datenbestand *separiert* ist.

Nun kann man Excel abzählen lassen, wie oft sich ein bestimmter Wert im Datenbestand befindet. Excel kann aber auch abzählen, wie oft ein kleinerer oder größerer Wert enthalten ist. In der Tabelle ist dann zweierlei zu beachten:

Abteilung	Wert		=ZÄHLENWENN(B2:B10;7777)	Abteilung	Wert		1
Abt_1	5424		=ZÄHLENWENN(B2:B10;"=7777")	Abt_1	5424		1
Abt_2	10214		=ZÄHLENWENN(B2:B10;"="+7777")	Abt_2	10214		0
Abt_3	12569		=ZÄHLENWENN(B2:B10;"&"7777")	Abt_3	12569		1
Abt_4	9632		=ZÄHLENWENN(B2:B10;">7777")	Abt_4	9632		7
Abt_5	11021		=ZÄHLENWENN(B2:B10;"<=7777")	Abt_5	11021		2
Abt_6	7777		=ZÄHLENWENN(B2:B10;D10)	Abt_6	7777		2
Abt_7	11414		=ZÄHLENWENN(B2:B1000;D10)	Abt_7	11414		2
Abt_8	7936		=ZÄHLENWENN(B:B;D10)	Abt_8	7936		2
Abt_9	9280		<=7777	Abt_9	9280		<=7777

Abb. 2.6 Abzählen im Datenbestand: Formeln und Werte

Wenn ein Kriterium (d. h. eine Bedingung) angegeben wird (zweite, dritte und vierte Formel), dann muss dieses in Anführungsstriche "... " gesetzt werden.

Das Kriterium ist folglich eine *Zeichenfolge*; es kann (siehe D4) sogar mittels des Verkettungsoperators & aus einzelnen Bestandteilen zusammengesetzt werden.

- **Man beachte** Der sonst gern genutzte *Verkettungsoperator* + darf hier nicht verwendet werden – die Formel in Zelle D3 liefert mit ihm ein falsches Ergebnis.

Es ist sogar möglich, das *Kriterium aus einer anderen Zelle* zu holen: Die Formeln in D7, D8 und D9 verwenden den Ausdruck aus der Zelle D10. Wird im Ausnahmefall nur die Vielfachheit eines bestimmten *numerischen* Wertes gesucht, dann kann dieser (siehe die erste Formel in der Zelle D1) auch ohne Kriterium, d. h. ohne die Anführungsstriche, eingetragen werden.

Befinden sich die Analyseergebnisse über oder neben dem Datenbestand, dann sind sie stets schnell zu finden mit [Strg] + [Pos1] . Die *Positionierung über dem Datenbestand* hat noch einen zweiten Vorteil: Man braucht in den Formeln, die Bereiche verwenden, oft nicht das konkrete Ende des Datenbestandes anzugeben, sondern kann durchaus großzügiger sein.

In der Tabelle von Abb. 2.6 ist in den Formeln in den Zellen D1 bis D7 der Zählbereich B2 : B10 sorgsam und genau eingetragen worden; die großzügige und weit reichende Angabe B2 : B1000 in D8 wäre aber nicht falsch, da leere Zellen übergangen werden. Und in D9 sieht man, dass man sogar *die komplette Spalte als Zählbereich* angeben kann – das wird viel zu selten benutzt. Dabei hat solche Großzügigkeit ihren nicht zu unterschätzenden Nutzen: Bei Eingabe weiterer Daten brauchen die Formeln nicht geändert zu werden.

Name	Geschlecht		=ZÄHLENWENN(B2:B10000;"m")	Name	Geschlecht		4
Person_1	m		=ZÄHLENWENN(B2:B10000;"w")	Person_1	m		5
Person_2	w		=ZÄHLENWENN(B2:B10000;"<"&"m")	Person_2	w		0
Person_3	m		=ZÄHLENWENN(B:B;">m")	Person_3	m		5
Person_4	m		=ZÄHLENWENN(B:B;"<m")	Person_4	m		1
Person_5	w		=ZÄHLENWENN(B:B;D10)	Person_5	w		1
Person_6	w			Person_6	w		
Person_7	w			Person_7	w		
Person_8	w			Person_8	w		
Person_9	m	<m		Person_9	m		<m

Abb. 2.7 Nichtnumerischer Datenbestand: Formeln und Werte

2.3.1.2 Abzählen von nichtnumerischen Einträgen

Die Tabelle in Abb. 2.7 zeigt, wie man *nichtnumerische Werte* (hier sind es einzelne Buchstaben) zählen lassen kann. Der Datenbestand befindet sich auch hier wieder im Bereich B2 : B10. Da aber die Eingabe diverser weiterer Zeilen zu erwarten ist, wird weitsichtig der Bereich B2 : B10000 eingetragen.

Man könnte sogar die ganze Spalte mittels B : B eintragen.

Dabei muss man nun aber beachten, dass dann die Überschrift mit untersucht wird. Die (falsche) 1 in den Zellen D5 und D6 entsteht folglich, weil die Überschrift mit dem großen Anfangsbuchstaben „G“ in der lexikografischen Anordnung vor dem kleinen „m“ kommt.

Schließlich zeigt die Tabelle in Abb. 2.8, wie die ZÄHLENWENN-Funktion benutzt werden kann, um nichtnumerische Merkmalswerte nach *gegebenem Textmuster* oder entsprechend ihrer lexikografischen Anordnung zu zählen.

Bei der Angabe eines *Textmusters* steht das *Fragezeichen* stets für genau ein (beliebiges) Zeichen, der *Stern* steht für eine beliebige Zeichenfolge beliebiger Länge.

Name	Anzahl aller "Meier"	=ZÄHLENWENN(A2:A9;"Meier")	Name	Anzahl aller "Meier"	2
Meier	Anzahl aller Einträge nach "Maier"	=ZÄHLENWENN(A2:A9;">"&"Maier")	Meier	Anzahl aller Einträge nach "Maier"	5
Maier	alle Einträge mit führendem "M"	=ZÄHLENWENN(A2:A9;"M*")	Maier	alle Einträge mit führendem "M"	6
Mayer	alle Einträge mit 4 Zeichen	=ZÄHLENWENN(A2:A9;"????")	Mayer	alle Einträge mit 4 Zeichen	2
Meier	alle Einträge mit M, N, ..., Z	=ZÄHLENWENN(A2:A9;">M*")	Meier	alle Einträge mit M, N, ..., Z	8
Maas	alle Einträge mit "e" an 3. Stelle	=ZÄHLENWENN(A2:A9;D9)	Maas	alle Einträge mit "e" an 3. Stelle	1
Mahs			Mahs		
Nurxs			Nurxs		
Nuernberg		??e*	Nuernberg		??e*

Abb. 2.8 Zählung von Texten und Textmustern: Formeln und Werte

Unter *lexikografischer Anordnung* versteht man die Anordnung, wie sie sich in jedem Telefon- oder Wörterbuch findet. Bei übereinstimmendem ersten Zeichen entscheidet das zweite Zeichen über die Position, stimmen die ersten beiden Zeichen überein, so entscheidet das dritte Zeichen usw.

Haben zwei Zeichenfolgen denselben Anfang („Mai“ und „Maier“), so steht die kürzere Zeichenfolge zuerst. In den Zellen D2 und D6 der Tabelle in Abb. 2.8 ist wieder demonstriert, dass man das *Kriterium aus Einzelteilen* zusammensetzen oder auch *aus einer anderen Zelle* holen kann.

2.3.1.3 Abzählen von Postleitzahlen usw.

- **Man beachte** Postleitzahlen, ebenso Personalnummern, Artikelnummern, Kennzahlen usw. heißen zwar „Zahlen“, sie sind natürlich *Ziffernfolgen*, dürfen aber in Excel-Tabellen nicht wie *Zahlen* eingegeben werden.

Postleitzahlen müssen wie *nichtnumerische Angaben* behandelt werden – schließlich soll mit ihnen ja nicht *gerechnet* werden, vielmehr geht es um das Abzählen bestimmter PLZ oder auch nur von PLZ-Bereichen (z. B. PLZ-Bereich 3).

Vor der Eingabe der hier genannten Merkmalswerte ist deshalb unbedingt die dafür vorgesehene Spalte als Textspalte festzulegen.

Oder man macht sich die Mühe und setzt vor die erste Ziffer jeder PLZ ein Hochkomma – damit wird nämlich die Typerkennung durch Excel ausgeschaltet und alles, was folgt, wird wie Text, also nichtnumerisch, abgespeichert (siehe auch Abschn. 1.1.3 im Kap. 1).

Zur Kontrolle trage man stets eine Postleitzahl ein, die mit einer Null beginnt, z. B. 09126. Verschwindet die Null beim Bestätigen der Eingabe und steht rechtsbündig nur 9126, dann war die Formatierung nicht erfolgreich, die Spalte muss gelöscht und neu formatiert werden. Oder man tippt eben mit vorgesetztem Hochkomma ein: '09126.

- **Man beachte** Eine nachträgliche Formatierung einer Spalte als Textspalte ist nicht möglich!

Die Tabellen in Abb. 2.9 beschreiben den *zählenden Umgang mit Postleitzahlen*.

Man kann zählen lassen, wie oft eine bestimmte PLZ im Datenbestand auftritt. Durch Eingabe von Textmustern lassen sich aber auch ganz bestimmte PLZ-Bereiche abzählen.

Kunde	PLZ	Anzahl der PLZ 56552	=ZÄHLENWENN(B\$2:B\$9;F1)	56552	Kunde	PLZ	Anzahl der PLZ 56552	3	56552
Kunde 1	56552	Anzahl der PLZ 56552	=ZÄHLENWENN(B\$2:B\$9;F2)	=56552	Kunde 1	56552	Anzahl der PLZ 56552	3	=56552
Kunde 2	09126	Anzahl aller 3-er PLZ	=ZÄHLENWENN(B\$2:B\$9;F3)	3*	Kunde 2	09126	Anzahl aller 3-er PLZ	1	3*
Kunde 3	56552	Anzahl aller 56-er PLZ	=ZÄHLENWENN(B\$2:B\$9;F4)	56*	Kunde 3	56552	Anzahl aller 56-er PLZ	3	56*
Kunde 4	55090	Anzahl aller 7 bis 9er PLZ	=ZÄHLENWENN(B\$2:B\$9;F5)	>=7*	Kunde 4	55090	Anzahl aller 7 bis 9er PLZ	2	>=7*
Kunde 5	86217				Kunde 5	86217			
Kunde 6	79958				Kunde 6	79958			
Kunde 7	56552				Kunde 7	56552			
Kunde 8	38011				Kunde 8	38011			

Abb. 2.9 Zählen von Postleitzahlen und PLZ-Bereichen; Formeln und Werte

Dabei beachte man, dass die Inhalte der beiden Zellen F1 und F2 von Excel unbedingt linksbündig eingetragen werden müssen. Das erreicht man, wie schon erwähnt, durch Formatierung als Textspalte oder Vorstellen des Hochkommas.

2.3.1.4 Abzählen von Datumsangaben

Datumsangaben werden von Excel intern als *ganze Zahlen* gespeichert – anstelle des konkreten Datums merkt sich Excel den *Datumswert*, das ist die abgelaufene Zahl der Tage seit dem 31. Dezember 1899. Der Neujahrstag des Jahres 1900 hat deshalb den Datumswert 1.

Die Funktion =DATWERT (. . .) liefert, wie zu sehen ist, diesen Datumswert zu allen gängigen deutschen Datumsdarstellungen, von denen drei übliche eingetragen wurden:

=DATWERT("1.1.2011")
=DATWERT("2/1/1900")
=DATWERT("31. Dezember 2010")

40544
2
40543

Weil Excel den 1.1.1900 als „Tag 1“ bezeichnet, ist also der 1.1.2011 der „Tag 40544“.

Es sei nicht verschwiegen, dass die von Excel genannten Datumswerte leider (fast) alle falsch sind – die tatsächliche Anzahl der seit Silvester 1999 vergangenen Tage ist um Eins geringer. Das liegt daran, dass Excel in allen bisherigen Versionen einen 29. Februar 1900 berücksichtigt hat, doch diesen *Tag hat es nie gegeben*. Das Jahr 1900 war kein Schaltjahr. Nur die durch 400 teilbaren Jahrhunderte sind Schaltjahre: Auch 2100 wird kein Schaltjahr werden.

Rechnungsnummer	Bezahl am	Datums werte	Anzahl vom 4. Juni:	=ZÄHLENWENN(B2:B100;DATWERT("4/6/2007"))	Rechnungsnummer	Bezahl am	Datums werte	Anzahl vom 4. Juni:	
Rechnung_Nr_1	04.06.2007	39237	Anzahl vor dem 4. Juni:	=ZÄHLENWENN(B2:B100;"<39237")	Rechnung_Nr_1	04.06.2007	39237	Anzahl vor dem 4. Juni:	1
Rechnung_Nr_2	08.08.2007	39302	Anzahl nach dem 4. Juni	=ZÄHLENWENN(B2:B100;">39237")	Rechnung_Nr_2	08.08.2007	39302	Anzahl nach dem 4. Juni	4
Rechnung_Nr_3	05.06.2007	37412	Anzahl nach dem 4. Juni	=ZÄHLENWENN(B2:B100;">"&DATWERT("4/6/2007"))	Rechnung_Nr_3	05.06.2007	37412	Anzahl nach dem 4. Juni	4
Rechnung_Nr_4	04.06.2007	37411	Anzahl nach dem 4. Juni	=ZÄHLENWENN(B2:B100;">"&F9)	Rechnung_Nr_4	04.06.2007	37411	Anzahl nach dem 4. Juni	4
Rechnung_Nr_5	09.09.2007	37508			Rechnung_Nr_5	39334	37508		
Rechnung_Nr_6	02.02.2007	37289			Rechnung_Nr_6	39115	37289		
Rechnung_Nr_7	04.09.2007	37503			Rechnung_Nr_7	39329	37503		
Rechnung_Nr_8	04.06.2007	37411		04.06.2007	Rechnung_Nr_8	39237	37411		04.06.2007

Abb. 2.10 Auszählen von Datumsangaben: Formeln und Werte

Dieser (einer der wenigen) Excel-Fehler fällt jedoch bei den Anwendungen kaum ins Gewicht, weil heutzutage in der Regel Datumsdifferenzen nur zwischen Datumsangaben aus der zweiten Hälfte des 20. Jahrhunderts und/oder bereits aus dem 21. Jahrhundert gesucht sind.

Stehen in einer Spalte schon *formatierte Datumsangaben* (mit [Strg]+[1]→ZÄHLEN→DATUM kann man vielfältige Datumsformate einstellen), dann ist eine unmittelbare Anwendung der Funktion =DATWERT (. . .) nicht mehr möglich.

Denn eine *eingetragene Datumsangabe* ist eine *Zahl*, die Funktion =DATWERT (. . .) kann aber nur eine Zeichenkette verarbeiten.

Es ist jedoch recht einfach, die zu einem *Datum*, das bereits in *üblicher deutscher Notation* in einer Zelle eingetragen wurde, den zugehörnden *Datumswert* zu bekommen: Man kopiert die Inhalte der Datumszellen in eine andere Spalte und formatiert sie mittels [Strg]+[1]→ZELLEN→ZÄHL zu *ganzen Zahlen ohne Kommastellen* um. Nun sind alle notwendigen Vorbetrachtungen abgeschlossen. In der Tabelle von Abb. 2.10 wird beispielhaft vorgeführt, wie *Datumsangaben ausgezählt* werden können.

Man beachte aber: Da zwischen Semikolon und schließender Klammer der Funktion =ZÄHLENWENN (. . . ; . . .) nur eine so genannte *Textkonstante* stehen darf, gibt es eine Fehlermeldung, wenn dort als Eintrag

=ZÄHLENWENN (B10:B100; ">DATWERT ("4/6/2007") "))

vorgenommen wird. Setzt man dagegen das Kriterium in der folgenden Weise aus zwei Zeichenfolgen zusammen

ZÄHLENWENN(B10:B100; ">"&DATWERT("4/6/2007")),

dann gibt es keine Konflikte.

Zur Ermittlung der Anzahl von Einträgen *vor oder nach einem bestimmten Datum* sollte man folglich das Kriterium wie in den Zellen F4 oder F5 zusammensetzen.

2.3.2 **Klassenbildung**

2.3.2.1 **Einführung**

Eingangs soll beispielhaft eine Excel-Tabelle betrachtet werden, die zu mehr als eintausend untersuchten Personen das jeweilige Körpergewicht (in Kilogramm) enthält. Wenn nun versucht wird, unter Verwendung des *Filters* die Anzahl und Bezeichnung der Merkmalswerte zu erfahren, dann erleidet man Schiffbruch. Und das ist vollkommen verständlich – es wäre im Gegenteil sehr unnatürlich, wenn diese vielen Individuen nur einige wenige Gewichtswerte annehmen würden. An eine Ermittlung *absoluter Häufigkeiten*, d. h. also an *einfaches Abzählen* unter Verwendung dieser sehr vielen Merkmalswerte, ist nur schwer zu denken.

Vielmehr erweist es sich nun als unbedingt notwendig, zuerst *Klassen* festzulegen. Als Klassen bezeichnet man aneinander anschließende Intervalle, die den gesamten Merkmalbereich überdecken, das heißt, die erste Klasse muss links vom Minimum beginnen, und die letzte Klasse muss rechts vom Maximum enden.

So wären zum Beispiel für einen Gewichts-Datenbestand folgende Klassen sinnvoll:

		unter	60	Anzahl:	
ab einschließlich	60	bis unter	80	Anzahl:	
ab einschließlich	80	bis unter	100	Anzahl:	
ab einschließlich	100	bis unter	120	Anzahl:	
ab einschließlich	120	bis unter	140	Anzahl:	
ab einschließlich	140			Anzahl:	

- **Man beachte** Welche *Klassengrenzen* im Detail zu wählen sind, das hängt jeweils von der betrachteten Thematik ab. Ebenso ist es nicht zwingend, dass stets gleiche Klassenbreiten gewählt werden müssen.

Generell gilt, dass *nicht zu viele und nicht zu schmale* Klassen gebildet werden sollten.

Nach der Festlegung der Klassengrenzen muss für jeden Wert aus dem Datenbestand ermittelt werden, zu welcher Klasse er gehört. Dafür gibt es mehrere Methoden.

Zuerst soll der Umgang mit der Funktion =WENN (. . . ; . . . ; . . .) vorgestellt werden.

Anschließend erfolgt die Nutzung der Funktion =VERWEIS (. . . ; . . . ; . . .), dann folgt ein Hinweis auf die Abzählfunktion =HÄUFIGKEIT (. . . ; . . .) und schließlich die Arbeit mit dem Werkzeug HISTOGRAMM. Letzteres ist am bequemsten – wer es eilig hat, sollte zuerst den Abschn. 2.3.2.6 lesen.

2.3.2.2 Klassenzuordnung mit der Funktion WENN

► **Man beachte** Die Excel-Funktion =WENN (. . . ; . . . ; . . .) verlangt stets *drei Angaben*: =WENN (Test;Ja_Eintrag;Nein_Eintrag). Sie sind durch Semikolon zu trennen.

An die erste Position ist die *Bedingung* zu schreiben, die zu prüfen ist. In der Mitte ist anzugeben, was in der Zelle, die diese Funktion enthält, bei *erfülltem Test* einzutragen ist. Rechts vom zweiten Semikolon ist anzugeben, welchen Inhalt die Zelle bei *nicht erfülltem Test* bekommen soll. Sehen wir uns an, wie falsch eine vorschnelle Anwendung der WENN-Funktion im Hinblick auf die Klassen-Zuordnung sein kann:

Name	Gewicht (kg)	Klasse
Person_1	62,33	=WENN(B2<60;1;2)
Person_2	45,86	=WENN(B3<60;1;2)
Person_3	91,37	=WENN(B4<60;1;2)
Person_4	115,76	=WENN(B5<60;1;2)
Person_5	58,02	=WENN(B6<60;1;2)
Person_6	77,16	=WENN(B7<60;1;2)
Person_7	72,36	=WENN(B8<60;1;2)
Person_8	56,32	=WENN(B9<60;1;2)

Name	Gewicht (kg)	Klasse
Person_1	62,33	2
Person_2	45,86	1
Person_3	91,37	2
Person_4	115,76	2
Person_5	58,02	1
Person_6	77,16	2
Person_7	72,36	2
Person_8	56,32	1

Offenbar ist die Zuordnung aller Werte unterhalb von 60 kg korrekt, aber die größeren Gewichtswerte werden noch nicht differenziert. Deshalb muss im NEIN-Eintrag differenziert werden, was durch eine erneute, interne Verwendung der WENN-Funktion möglich ist:

Name	Gewicht (kg)	Klasse
Person_1	62,33	=WENN(B2<60;1;WENN(B2<80;2;3))
Person_2	45,86	=WENN(B3<60;1;WENN(B3<80;2;3))
Person_3	91,37	=WENN(B4<60;1;WENN(B4<80;2;3))
Person_4	115,76	=WENN(B5<60;1;WENN(B5<80;2;3))
Person_5	58,02	=WENN(B6<60;1;WENN(B6<80;2;3))
Person_6	77,16	=WENN(B7<60;1;WENN(B7<80;2;3))
Person_7	72,36	=WENN(B8<60;1;WENN(B8<80;2;3))
Person_8	56,32	=WENN(B9<60;1;WENN(B9<80;2;3))

Name	Gewicht (kg)	Klasse
Person_1	62,33	2
Person_2	45,86	1
Person_3	91,37	3
Person_4	115,76	3
Person_5	58,02	1
Person_6	77,16	2
Person_7	72,36	2
Person_8	56,32	1

Ist der erste Test erfüllt (Gewicht unter 60 kg), dann wird eine 1 eingetragen. Andernfalls folgt der innere Test (Gewicht unter 80 kg?). Ist dieser erfüllt, dann wird eine 2 eingetragen. Der Nein-Fall des inneren Tests führt schließlich zum Eintrag 3. Nun wird zwischen 40 und 80 kg bereits richtig in die Klassen eingeordnet, allerdings wird, wie zu erkennen, der über-100-Wert 115,76 noch falsch zugeordnet.

Folglich muss im Nein-Eintrag der inneren WENN-Funktion weiter differenziert werden. Man überlege sich, dass zur differenzierten Einordnung der Werte über 120 und 140 kg schließlich noch zwei weitere innere WENN-Funktionen verwendet werden müssen.

Sehen wir uns die endgültige, ziemlich verschachtelte WENN-Funktion an, die aber alle Gewichtswerte richtig in die zugehörige Klasse einordnet.

```
=WENN (B2<60 ; 1 ; WENN (B2<80 ; 2 ; WENN (B2<100 ; 3 ; WENN (B2<120 ; 4 ; 5) ) ) )
```

Nun könnte die bekannte Anwendung der Funktion =ZÄHLENWENN(...; ...) folgen, mit deren Hilfe schließlich festgestellt wird, wie viele Merkmalswerte des Datenbestandes in die jeweilige Klasse fallen. Es wäre übrigens auch möglich gewesen, die Klassen anstelle mit den Zahlen 1 bis 6 auch nichtnumerisch zu bezeichnen.

Eine solche Variante, die allerdings wegen der gewählten Vokabeln nicht unproblematisch ist, soll mit folgender Formel angegeben werden, wobei die Schreibweise die Verschachtelung etwas deutlicher machen soll:

```
=WENN(B2<60; "dünn";  
      WENN(B2<80; "normal";  
            WENN(B2<100; "kräftig";  
                  WENN(B2<120; "korpulent"; "Jumbo"))))
```

Für alle Leserinnen und Leser, denen der Umgang mit der WENN-Funktion in dieser komplizierten Form unsympathisch sein sollte (was durchaus verständlich ist), sei hier bereits darauf hingewiesen, dass die Klassenzuordnung oft auch wesentlich einfacher vorgenommen werden kann. Dazu stellt Excel das Werkzeug HISTOGRAMM zur Verfügung, das in Abschn. 2.3.2.6 beschrieben wird. Trotzdem soll erst noch ein weiterer kleiner Abschnitt der leistungsfähigen WENN-Funktion gewidmet werden.

2.3.2.3 Kodierung von nichtnumerischen Daten mit WENN

Wie schon mehrfach erwähnt, bietet Excel im Gegensatz zu einigen reinen Statistik-Programmen die uneingeschränkte Möglichkeit, Daten in ihrer *ursprünglichen numerischen oder nichtnumerischen Form* in einer Tabelle zu speichern.

Mit der ZÄHLENWENN-Funktion kann man problemlos auch ihre Häufigkeiten analysieren.

Es hat allerdings Tradition, dass insbesondere Auswertungen von Befragungen grundsätzlich *kodiert* erfasst werden, d. h. jeder Antwort wird eine Zahl zugeordnet, die dann eingegeben wird.

Nicht wenige Bücher zur Marktforschung und Datenanalyse widmen sich ausführlich den Fragen einer sinnvollen Kodierung. Es gibt jedoch Grenzen der Verarbeitung nichtnumerischer Daten.

Für den Fall, dass *nichtnumerischen Merkmalsausprägungen* nachträglich *numerische Merkmalswerte* zugeordnet werden müssen, kann man die Funktion =WENN(. . . ; . . . ; . . .) sehr gut verwenden.

Sehen wir uns zum Beispiel an, wie die Zuordnung „w“ \rightarrow 1 und „m“ \rightarrow 0 realisiert wird.

Name	Geschlecht	Kodewert
Person_1	m	=WENN(B2="m";0;1)
Person_2	m	=WENN(B3="m";0;1)
Person_3	w	=WENN(B4="m";0;1)

Name	Geschlecht	Kodewert
Person_1	m	0
Person_2	m	0
Person_3	w	1

2.3.2.4 Die Funktion VERWEIS

Mit der Excel-Funktion =VERWEIS (...;...;...) lassen sich ebenfalls Klassenzuordnungen realisieren.

Diese Funktion verlangt *drei Angaben*:

- Zuerst ist die Zelle einzutragen, deren Inhalt klassifiziert werden soll.
- Anschließend ist ein Spaltenbereich einzutragen, in dem sich die sortierten Klassengrenzen befinden.
- Schließlich wird noch ein gleich langer Spaltenbereich verlangt, in dem sich die jeweiligen Klassenbezeichnungen befinden.

Die Funktion =VERWEIS (...;...;...) durchsucht für jeden Merkmalswert den Bereich der Klassengrenzen, von oben beginnend. Ist die nachfolgende Klassengrenze zu groß, dann wird der nebenstehende Wert als Klassenbezeichnung eingetragen.

2.3.2.5 Die Funktion HÄUFIGKEIT

Die Funktion =HÄUFIGKEIT (...;...) ist eine so genannte Matrix- oder Vektorfunktion. Sie liefert nicht nur einen einzigen Ergebniswert, wie wir es bei den bisher vorgestellten Funktionen durchweg erlebten, sondern füllt simultan einen *ganzen vorgegebenen Bereich*.

Eine erste Anwendung einer solchen Matrixfunktion wird im Abschn. 4.4.3 des Kap. 4 zu erleben sein; interessierte Leser seien zusätzlich auf weiterführende Literatur, z. B. [1], [10] und [15] verwiesen.

Hier wurden diesmal die Verbalbezeichnungen benutzt, um zu zeigen, dass es keinesfalls immer Zahlen sein müssen, mit denen man eine Klasse bezeichnet. Denn die Funktion =ZÄHLENWENN (...;...;...) kann anschließend auch problemlos abzählen, wie oft die Klasse Korpulent aufgetreten ist.

Name	Gewicht (kg)	Klasse	Klassen- grenzen	Klassen- bezeichnung
Person_1	62,33	=VERWEIS(B2:D\$2:D\$6;E\$2:E\$6)	40	Dünn
Person_2	45,86	=VERWEIS(B3:D\$2:D\$6;E\$2:E\$6)	60	Normal
Person_3	91,37	=VERWEIS(B4:D\$2:D\$6;E\$2:E\$6)	80	Kräftig
Person_4	115,76	=VERWEIS(B5:D\$2:D\$6;E\$2:E\$6)	100	Korpulent
Person_5	58,02	=VERWEIS(B6:D\$2:D\$6;E\$2:E\$6)	120	Schwer
Person_6	77,16	=VERWEIS(B7:D\$2:D\$6;E\$2:E\$6)		
Person_7	72,36	=VERWEIS(B8:D\$2:D\$6;E\$2:E\$6)		
Person_8	56,32	=VERWEIS(B9:D\$2:D\$6;E\$2:E\$6)		

Name	Gewicht (kg)	Klasse	Klassen- grenzen	Klassen- bezeichnung
Person_1	62,33	Normal	40	Dünn
Person_2	45,86	Dünn	60	Normal
Person_3	91,37	Kräftig	80	Kräftig
Person_4	115,76	Korpulent	100	Korpulent
Person_5	58,02	Dünn	120	Schwer
Person_6	77,16	Normal		
Person_7	72,36	Normal		
Person_8	56,32	Dünn		

2.3.2.6 Das Werkzeug HISTOGRAMM

Das Werkzeug HISTOGRAMM kombiniert die Leistungen der beiden bereits vorgestellten Funktionen =WENN (...;...;...) und =ZÄHLENWENN (...;...;...).

- **Wichtiger Hinweis** Mit Hilfe des Werkzeugs HISTOGRAMM kann nach Festlegung der oberen Klassengrenzen sofort das *Ergebnis der Auszählung* erhalten werden.

Es soll derselbe Datenbestand wie im vorigen Abschnitt verwendet werden, in Spalte D sind die oberen Klassengrenzen eingetragen:

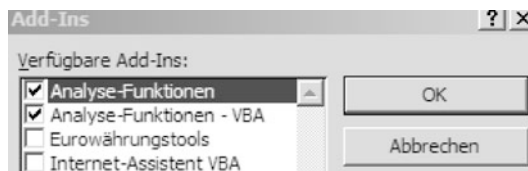
Name	Gewicht (kg)	Klassengrenzen
Person_1	62,33	60
Person_2	45,86	80
Person_3	91,37	100
Person_4	115,76	120
Person_5	58,02	
Person_6	77,16	
Person_7	72,36	
Person_8	56,32	

Nun muss das Werkzeug HISTOGRAMM angefordert werden. Wenn es bereits verfügbar ist, dann findet man es in der Registerkarte Daten in der Gruppe Analyse die Werkzeugsammlung Datenanalyse:



Nach Klick auf Datenanalyse öffnet sich ein Fenster Analyse-Funktionen, dort findet man schließlich das Werkzeug HISTOGRAMM.

Oft aber ist die Werkzeugsammlung Datenanalyse anfangs nicht verfügbar. Dann muss sie in folgender Weise angefordert werden: Man muss zuerst Excel-Optionen über die runde Schaltfläche Office (oder Optionen über Datei) finden, dann sind Add-Ins auszuwählen und die Schaltfläche [Gehe zu] ist anzuklicken. Aus dem dann erscheinenden Angebot sind Analyse-Funktionen und Analyse-Funktionen VBA auszuwählen:



Hat man das Werkzeug HISTOGRAMM erst einmal gestartet, dann ist seine Handhabung recht einfach. Man braucht nur, wie in Abb. 2.11 zu sehen ist, zwei Angaben in die beiden oberen Eingabefenster einzutragen. Das Fenster Eingabebereich bekommt Anfangs- und Endezelle des Bereiches mit den zu untersuchenden *Daten*, wie üblich getrennt durch einen Doppelpunkt. Hier war also für unser Beispiel der Bereich mit den *Gewichtsangaben* einzutragen. In das Fenster Klassenbereich ist einzutragen, wo sich die festgelegten *Klassengrenzen* befinden. Schließlich kann man wählen, ob die Ergebnisse des Werkzeugs auf ein *neues Tabellenblatt* oder neben den Datenbestand ausgegeben werden sollen.

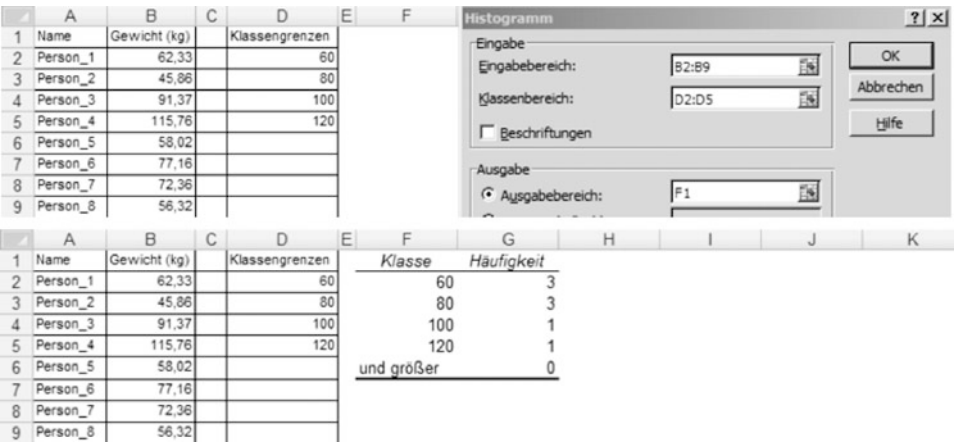


Abb. 2.11 Eingaben für das Werkzeug HISTOGRAMM und Ergebnis

Sehen wir uns nun die Erklärung zur Ergebnisausgabe des Werkzeugs an:

Klasse	Häufigkeit	<--	Erklärung
60	3	<--	unter 60
80	3	<--	über 60 bis einschließlich 80
100	1	<--	über 80 bis einschließlich 100
120	1	<--	über 100 bis einschließlich 120
und größer	0	<--	über 120

An dieser Stelle zeigt sich, weshalb im vorigen Abschnitt die WENN-Funktion beschrieben wurde. Sie muss man benutzen, wenn man spezielle Wünsche hat, beispielsweise, dass die angegebenen Klassengrenzen jeweils die *unteren Klassenränder* beschreiben sollen. Benutzt man dagegen das bequeme Werkzeug HISTOGRAMM, zählt man für diese Bequemlichkeit natürlich einen Preis:

- **Man beachte** HISTOGRAMM betrachtet grundsätzlich die angegebenen Werte im Klassenbereich stets als *obere Klassengrenzen*.

Wie gesagt, damit muss man leben. Man kann leider keine Einstellung in dem Werkzeug HISTOGRAMM vornehmen, so dass die angegebenen Grenzen als untere Ränder der Klassen angesehen werden. Individuelle Wünsche müssen tatsächlich mit der WENN-Funktion umgesetzt werden. Das ist zwar aufwändiger, hat aber zusätzlich den großen Vorteil, dass bei Änderung der Daten automatisch alle Ergebnisse neu berechnet werden.

Abb. 2.12 Umsatzmeldungen
von Mitarbeitern

Mitarbeiter	Monat	Umsatz in \$
Meier_1	1	964,19
Meier_2	1	1351,5
Meier_3	1	1358,32
Meier_3	1	986,38
Meier_2	1	1254,79
Meier_2	1	1399,48
Meier_2	2	1225,75
Meier_3	2	886,97
Meier_2	2	988,71
Meier_4	2	1195,53

Alle Excel-Werkzeuge liefern dagegen starre (statische) Ergebnisse; d.h. werden die Daten geändert, dann muss das Werkzeug wieder neu angewandt werden.

2.3.3 Häufigkeiten von Paaren: Pivot-Tabellen

2.3.3.1 Aufgabenstellung und Bezeichnungen

Gehen wir von folgendem, sehr vereinfachendem Beispiel aus. Angenommen, eine Vertriebsfirma habe vier Außendienst-Mitarbeiter, nennen wir sie der Einfachheit halber Meier_1 bis Meier_4. Sie verkaufen Waren, machen Umsatz. Und jedes Mal, wenn sie einen Umsatz erzielt haben, melden sie sich in der Zentrale und teilen drei Dinge mit: ihren Namen, den aktuellen Monat, und ihren erzielten Umsatz. Alles wird in eine Excel-Datei eingetragen (siehe Abb. 2.12).

Am Jahresende soll Bilanz gezogen werden: Wer hat am besten gearbeitet? Welcher Gesamtumsatz wurde erzielt? Wie sieht es in den einzelnen Monaten aus?

Doch alle diese Fragen lassen sich mit der – in ihrer Aussagekraft kaum zu unterbietenden – Original-Excel-Datei nicht beantworten. es ist ein *unübertriffen nichts sagender Datenbestand*. Er ist in dieser Form für Auswertungen *völlig unbrauchbar*: Weder ist zu erkennen, wie viele Meldungen überhaupt vorliegen, noch lässt sich ablesen, welchen Umsatz welcher Mitarbeiter in welchem Monat tatsächlich erzielt hat. Benötigt wird eine Aufbereitung der Daten, mit deren Hilfe alle Fragen beantwortet werden können. Eine derartige Aufbereitung wird durch die *Pivot-Tabelle* geliefert.

Sehen wir uns schon einmal – bevor wir zur Herstellung kommen – in Abb. 2.13 das anzustrebende Ergebnis an: Diese *Pivot-Tabelle* (früher auch als *Kreuztabelle* bezeichnet) beantwortet alle Fragen. Sie interpretiert sich gewissermaßen von selbst.

Leicht zu erkennen sind die *Monats-Umsatz-Summen* der einzelnen Mitarbeiter, auch die Monate ohne Meldungen erkennt man sofort. Die *Jahres-Summen* der Mitarbeiter können in der Fußzeile abgelesen werden, die *Monatsumsätze* in der Bilanz-Spalte rechts, und der der *Gesamt-Jahresumsatz* in Höhe von 44085,60 steht ganz rechts unten.

Summe von Umsatz in \$	Mitarbeiter				
Monat	Meier_1	Meier_2	Meier_3	Meier_4	Gesamtergebnis
1	964,19	4005,77	2344,7		7314,66
2		2214,46	886,97	2235,17	5336,6
3		859,15	928,41		1787,56
4	1337,44		1296,71	1124,39	3758,54
5		1316,83	1131,47		2448,3
6			4442,26		4442,26
7		2208,24	4316,31		6524,55
8		2576,62		1009,09	3585,71
9		3337,23			3337,23
10		1109,06		1368,95	2478,01
11	1946,37				1946,37
12				1125,81	1125,81
Gesamtergebnis	4248	17627,36	15346,83	6863,41	44085,6

Abb. 2.13 Pivot-Tabelle

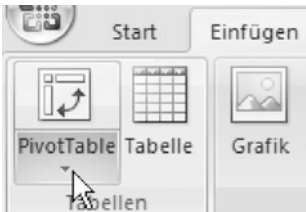
Eine Pivot-Tabelle liefert Bilanzen in außerordentlich übersichtlicher Form.

Kommen wir nun zur Technologie – wie wird eine Pivot-Tabelle hergestellt?

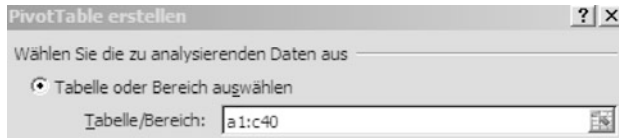
2.3.3.2 Herstellung einer Pivot-Tabelle

Wenn der *Datenbestand*, wie schon häufig gefordert, durch *mindestens eine Leerzeile* und durch *mindestens eine Leerspalte* von möglicherweise anderen Einträgen in der Tabelle *separiert* ist, dann reicht es zur Erzeugung der Pivot-Tabelle aus, den Tabellenkursor *in irgendeine Zelle des Datenbestandes* (z. B. in eine Zelle der Überschriften) zu setzen. Anschließend wird der Pivot-Tabellen-Assistent benötigt.

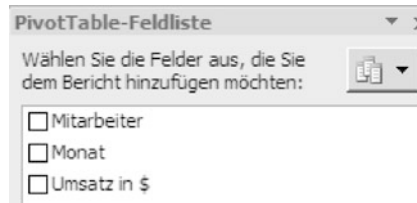
Man findet den Pivot-Tabellen-Assistent in Excel in der in Registerkarte Einfügen in der Gruppe Tabellen:



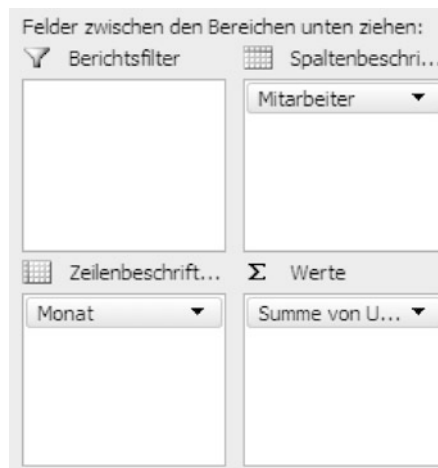
Nach dem Start des Pivot-Tabellen-Assistenten muss man zuerst die Frage beantworten, in welchem Bereich sich der *gesamte Datenbestand* befindet:



Danach ist es sehr einfach: Der Assistent bietet einerseits die drei Überschriften des Datenbestandes an:



und andererseits kommt das Angebot, die passenden Überschriften in die gewünschte Position zu ziehen:

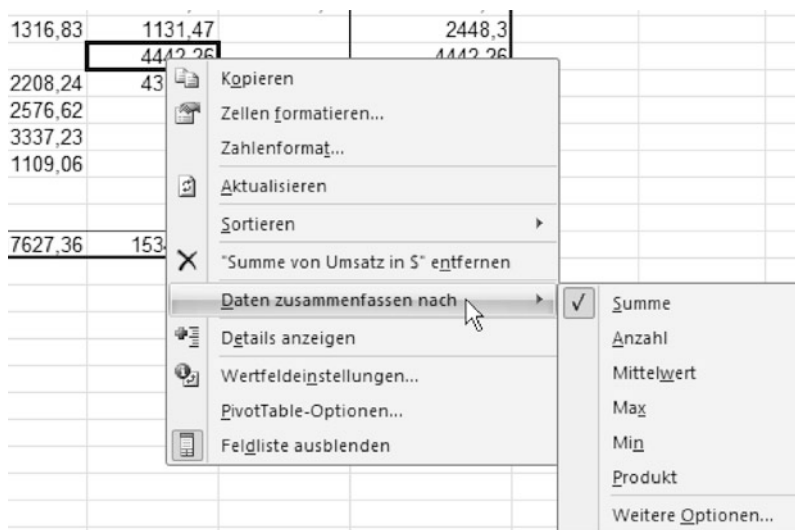


Hier wurde folglich festgelegt, dass die Namen der vier Mitarbeiter in der Kopfzeile (d. h. als Spaltenüberschriften) der Pivot-Tabelle erscheinen sollen, die Monatsnummern dagegen als Zeilenbeschriftungen. Und in das Feld Werte kommen die Umsätze, sie werden grundsätzlich aufsummiert (Änderungen sind aber möglich, siehe unten). Sollte die angezeigte Pivot-Tabelle den optischen Anforderungen nicht genügen, lassen sich problemlos die Überschriften vertauschen. Überhaupt kann man durch geschicktes Probieren schnell den Umgang mit dem Assistenten lernen.

2.3.3.3 Feldeinstellungen

Durch einfachste Bedienhandlungen kann die *Anzeige der Pivot-Tabelle* so umgestaltet werden, dass eine Fülle weiterer Aussagen ablesbar ist.

Dazu setzt man den Tabellenkursor in irgendeine Zelle, die sich aber im Inneren der Pivot-Tabelle befinden muss, klickt auf die Neben-Maustaste (für Rechtshänder die rechte Maustaste, für Linkshänder die linke Maustaste) und wählt die Zeile *Daten zusammenfassen nach* → (auch „Werte zusammenfassen nach →“) aus:



Die jetzt angezeigte Liste gibt die *vielfältigen Möglichkeiten* an, welche Informationen neben der Umsatzsumme pro Mitarbeiter und Monat noch einstellbar sind:

- Man kann sich die Anzahl der Meldungen pro Monat und pro Mitarbeiter anzeigen lassen.
- Weiter gibt es die Möglichkeit, sich die größte der jeweiligen Monats-Meldungen mittels Maximum anzeigen zu lassen.
- Analog dazu darf es auch die niedrigste Meldung im Monat (mittels Minimum) sein.
- Und schließlich kann auch das Mittel aller Meldungen im Monat angezeigt werden.

Da der Pivot-Tabellen-Assistent außerordentlich bedienfreundlich gestaltet ist und insbesondere die Möglichkeit bietet, jede Bedienhandlung sofort wieder rückgängig machen zu können, kann der Umgang mit diesem Excel-Werkzeug tatsächlich leicht erlernt werden.

2.3.3.4 Abzählen von Paaren mittels Pivot-Tabellen

Wieder soll von einer ganz einfachen Aufgabe ausgegangen werden: Für eine Menge von 39 Schülern sind Namen und die Zensuren in zwei Fächern in einer Excel-Tabelle erfasst. Zu beantworten ist die Frage, wie oft jede der 36 Zensurenkombinationen (1,1) bis (6,6) im Datenbestand auftritt.

Name	Fach_1	Fach_2
Schüler_1	5	4
Schüler_2	5	6
Schüler_3	4	6
Schüler_4	2	6
Schüler_5	3	2
Schüler_6	2	3
Schüler_7	4	6

Sehen wir uns zuerst die Antwort an, die wieder die hochinformativ Form einer Pivot-Tabelle hat:

Anzahl von Fach_1		Fach_2						
Fach_1		1	2	3	4	5	6	Gesamtergebnis
	1	2		1	2	2		7
	2	1	1	3	2		1	8
	3		1		1			2
	4	3	1	1		1	2	8
	5	1	1	1	3	2	2	10
	6	3		1				4
Gesamtergebnis		10	4	7	8	5	5	39

39 Schüler wurden erfasst, und man sieht eine relativ gleichmäßige Verteilung der Zensurenkombinationen (ein Zeichen dafür, dass die Zahlen offensichtlich nicht real sind, sondern nur zu Demonstrationszwecken zufällig erzeugt wurden).

Bei der Herstellung der Tabelle wird es aber ein Problem geben: *Eine Überschrift ist zu wenig vorhanden.* Was soll denn nun in das Feld WERTE eingetragen werden?

Die Lösung Man zieht zuerst in das Feld WERTE noch einmal eine der beiden Überschriften Fach_1 oder Fach_2. Dann entsteht vorerst eine sehr unsinnige Pivot-Tabelle:

Summe von Fach_1		Fach_2					
Fach_1		1	2	3	4	5	6
	1	2		1	2	2	
	2	2	2	6	4		2
	3		3		3		
	4	12	4	4		4	8
	5	5	5	5	15	10	10
	6	18		6			
Gesamtergebnis		39	14	22	24	16	20

Warum ist sie unsinnig? Man kann das in der Pivot-Tabelle in der Zelle A3 erkennen – dort steht *Summe-Fach_1*. Das bedeutet, es fand kein *Abzählen*, sondern ein (hier sinnloses) Aufsummieren von Zensuren statt.

Mit der Erinnerung an Abschn. 2.3.3.3 lässt sich der Übergang zur *Angabe der Vielfachheiten* beschreiben: Tabellenkursor in irgendeine Zelle der Pivot-Tabelle setzen → Neben-Mousetaste klicken → Daten/Werte zusammenfassen nach → auswählen → Anzahl einstellen. Dann liefert die Pivot-Tabelle in der Tat die erwünschten *absoluten Häufigkeiten*, so dass man die bereits dargestellte korrekte Übersicht erhält.

2.3.3.5 Variable Pivot-Tabellen

In den Datenbestand sei nun von den bereits betrachteten 39 Schülern neben den Zensuren in zwei Fächern zusätzlich in Spalte D unter der Überschrift Art das Geschlecht der Schüler aufgenommen worden:

Name	Fach_1	Fach_2	Art
Schüler_1	5	4	w
Schüler_2	5	6	w
Schüler_3	4	6	w
Schüler_4	2	6	w
Schüler_5	3	2	w
Schüler_6	2	3	w
Schüler_7	4	6	m
Schüler_8	2	2	w
Schüler_9	6	3	w
Schüler_10	4	1	m

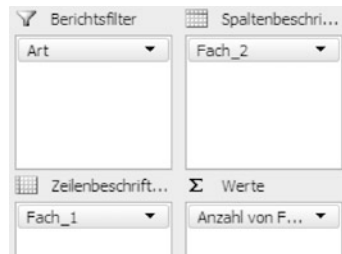
Drei Fragen sollen nun mit Hilfe von drei Pivot-Tabellen schnell beantwortet werden:

- Wie ist die Situation aller Schüler,
- wie speziell der Schülerinnen,
- wie speziell der Schüler?

Eine recht umständliche Vorgehensweise wäre, den Datenbestand zuerst einmal nach dem Geschlecht sortieren zu lassen (Tabellenkursor in Spalte D, dann die Schaltfläche zum Sortieren drücken, dabei nichts markieren). Zwischen Schülern und Schülerinnen würde

anschließend eine Leerzeile eingefügt – damit wären beide Datenbestände separiert. Anschließend könnten damit zwei verschiedene Pivot-Tabellen hergestellt werden.

Viel einfacher lässt sich das alles durchführen, wenn man nach der Herstellung der einfachen Pivot-Tabelle (eine Überschrift zusätzlich in das Feld *Werte* ziehen, auf *Anzahl* umstellen) die Überschrift *Art* in das bisher nicht beachtete Feld *Berichtsfilter* (oder kurz *Filter*) zieht.



Dann erhält man eine Pivot-Tabelle mit einer *zusätzlichen Wahlmöglichkeit*, die sich auf alle Merkmalswerte bezieht, mit der das Feld *Berichtsfilter* belegt wurde. Denn nun erscheint ganz oben in der Pivot-Tabelle die Überschrift *Art* mit allen Auswahlmöglichkeiten:



Je nach Einstellung erhält man jetzt per Mausklick die drei gewünschten Pivot-Tabellen: die Gesamtübersicht, die Übersicht für die Schülerinnen, die Übersicht für die Schüler.

2.3.3.6 Pivot-Tabellen mit Datumsangaben

In den vorigen Abschnitten wurde ersichtlich, dass der Pivot-Tabellen-Assistent sowohl Datenbestände mit *numerischen* als auch mit *nichtnumerischen* Merkmalswerten verarbeiten kann.

- **Wichtiger Hinweis** Falls *nichtnumerische* Merkmalswerte vorliegen, wird *grundsätzlich abgezählt*, bei *numerischen Werten* ist standardmäßig anfangs die *Summenbildung* eingestellt.

Falls im Datenbestand *Datumsangaben* auftreten, dann muss man stets daran denken, dass Excel nicht das Angezeigte speichert, sondern den *Datumswert* – wie schon im Abschn. 1.1.2 des Kap. 1 ausführlich beschrieben wurde. Ein ausführliches Beispiel mag das verdeutlichen. Eine Excel-Tabelle enthalte die Geburtsdaten von 39 Ehepaaren:

Er	Sie
13.01.1968	01.07.1968
26.01.1978	18.12.1978
12.03.1977	02.05.1978
07.02.1974	24.04.1974

Gesucht ist eine Zusammenstellung, aus der ersichtlich ist, wie viele Paare es gibt, bei denen beide Partner im Januar Geburtstag haben, er im Januar und sie im Februar usw. Insgesamt gibt es also 144 Kombinationen der Geburtsmonate.

Was tun? Nun, die Idee ist einleuchtend, da wird eben die Menge aller Geburtsdaten markiert und mittels `[Strg] + [1] → ZAHLEN → BENUTZERDEFINIERT → MMM` hinsichtlich des Anzeigemodus so umgestellt, dass *nur noch der Monatsname* angezeigt wird:

Er	Sie
Jan	Jul
Jan	Dez
Mrz	Mai
Feb	Apr

Das Ergebnis kann sich sehen lassen – die Tabelle zeigt tatsächlich nur noch die Monatsnamen.

Was spricht dagegen, nun den Pivot-Tabellen-Assistenten zu starten?

Doch dieser liefert ein sehr, sehr überraschendes und konfuse Bild, ein ganz kleiner Teil davon ist in Abb. 2.14 zu sehen! Diese „Pivot-Tabelle“ besteht aus unsagbar vielen leeren Zellen, bisweilen findet sich eine verlorene Eins, und von sortierter Reihenfolge der Monate kann wohl auch nicht die Rede sein. Wie kann man diesen Misserfolg erklären?

Erklärung Der Pivot-Tabellenassistent hat *nicht* die *angezeigten Monatsnamen*, sondern die tatsächlich *gespeicherten Datumswerte* der ursprünglich eingetragenen Geburtsdaten analysiert. Und sieht man sich diese Datumswerte an (herzustellen über `[Strg] + [1] → ZAHLEN → ZAHL`), dann ist es klar, dass nur verschwindend wenige der Kombinationen dieser vielen, vielen Datumswerte eine Belegung haben werden.

Der Pivot-Tabellen-Assistent weiß das aber nicht, er stellt immer *alle möglichen Kombinationen der Merkmalswerte* zusammen. Was also ist zu tun?

Anzahl von Sie	Sie														
Er	Jul	Feb	Mrz	Mrz	Apr	Jul	Sep	Jan	Aug	Dez	Feb	Jul	Aug	Jan	Jul
Jan				1											
Feb		1													
Mai			1												
Sep	1														
Mai					1										
Sep								1							
Mrz						1									
Jul							1								
Nov											1				
Jul									1						
Dez										1					
Jan												1			
Sep													1		
Jun															1
Aug														1	

Abb. 2.14 Ausschnitt (links oben) der großen Pivot-Tabelle

Es muss dafür gesorgt werden, dass sich *nicht die Datumswerte*, sondern die *Monatsnamen* oder *-nummern* im Datenbestand befinden. Das erreicht man aber nicht durch *Veränderung der Anzeige*, sondern durch *Anwendung der Funktion* =MONAT (. . .).

Abbildung 2.15 illustriert das Vorgehen. In den Spalten A und B stehen noch einmal die *Geburtsdaten* in klassisch deutscher Anzeige. In den Spalten D und E finden sich die *zugehörigen Datumswerte*, die man nach Kopie und Formatumstellung erhielt. In den Spalten G und H schließlich werden mit Hilfe der Datumsfunktion =MONAT (. . .) aus den Datumswerten die *Monatsnummern* berechnet.

Nun kann man auf den neuen numerischen Datenbestand in den Spalten G und H (er ist durch die Leerspalte auch *separiert*) den Pivot-Tabellen-Assistenten anwenden. Wieder

Er	Sie	Er	Sie	Er	Sie
13.01.1968	01.07.1968	24850	25020	=MONAT(D2)	=MONAT(E2)
26.01.1978	18.12.1978	28516	28842	=MONAT(D3)	=MONAT(E3)
12.03.1977	02.05.1978	28196	28612	=MONAT(D4)	=MONAT(E4)

Er	Sie	Er	Sie	Er	Sie
13.01.1968	01.07.1968	24850	25020	1	7
26.01.1978	18.12.1978	28516	28842	1	12
12.03.1977	02.05.1978	28196	28612	3	5

Abb. 2.15 Anwendung der MONAT-Funktion: Formeln und Werte

wird, da eine *Abzählaufgabe* vorliegt, anfangs eine der beiden Überschriften in das Feld Werte gezogen und dann die Anzeige, wie oben beschrieben, auf Anzahl umgestellt. Sehen wir uns nun die entstandene, richtige Pivot-Tabelle an:

Anzahl von Sie		Sie												Gesamtergebnis
Er		1	2	3	4	5	6	7	8	9	10	11	12	
1				2		1		1		1			1	6
2			1		1	1	1	1						5
3						1	1	2						4
4						1				1				2
5				1	1			1						3
6							1	1			1			3
7									1	1	1			3
8	1						1					1		3
9	1							2	1					4
10						1								1
11			1				1							2
12			1										2	3
Gesamtergebnis		2	3	3	2	5	5	8	2	3	2	1	3	39

Sie zeigt uns die Verteilung der Geburtsdaten der Ehepartner der erfassten 39 Ehepaare. Man kann zum Beispiel daraus erkennen, dass es dabei lediglich vier Paare gibt, deren Ehepartner im selben Monat geboren sind: Je ein Paar im Februar und Juni, und bei zwei Paaren haben „Er“ und „Sie“ gemeinsam im Dezember Geburtstag.

2.3.4 Absolute Summenhäufigkeiten

Die folgende Tabelle zeigt, wie man vorgehen muss, wenn man

- die *Summe von Erträgen* (Umsätzen o. ä.) angezeigt bekommen möchte und dazu
- *weiter* sehen will, wie sich diese Ertragssumme *schrittweise aus den Umsätzen der einzelnen Abteilungen* entwickelt.

Abteilung	Umsatz	
Abt_1	2562,81	=SUMME(B\$2:B2)
Abt_2	2899,33	=SUMME(B\$2:B3)
Abt_3	2459,80	=SUMME(B\$2:B4)
Abt_4	2228,62	=SUMME(B\$2:B5)
Abt_5	2291,93	=SUMME(B\$2:B6)
Abt_6	2302,63	=SUMME(B\$2:B7)
Abt_7	2671,63	=SUMME(B\$2:B8)
Abt_8	2011,65	=SUMME(B\$2:B9)
Abt_9	2109,74	=SUMME(B\$2:B10)
Abt_10	2494,34	=SUMME(B\$2:B11)
Summe	24032,48	

Abteilung	Umsatz	
Abt_1	2562,81	2562,81
Abt_2	2899,33	5462,14
Abt_3	2459,80	7921,94
Abt_4	2228,62	10150,56
Abt_5	2291,93	12442,49
Abt_6	2302,63	14745,12
Abt_7	2671,63	17416,75
Abt_8	2011,65	19428,40
Abt_9	2109,74	21538,14
Abt_10	2494,34	24032,48
Summe	24032,48	

Die Daten in Spalte C bezeichnet man dann als *absolute Summenhäufigkeiten*.

Sonderlich aussagekräftig sind diese Zahlen allerdings noch nicht. Das ändert sich schon ein wenig, wenn anfangs der Tabellenkursor in eine Zelle des Bereiches B2:B11 gesetzt wird (aber dabei nichts markieren) und die *Schaltfläche zum Sortieren* angeklickt wird. Das führt zum Umordnen des *gesamten* Datenbestandes, so dass die Einträge unter der Überschrift *Umsatz* aufsteigend sortiert erscheinen.

Abteilung	Umsatz	
Abt_8	2011,65	2011,65
Abt_9	2109,74	4121,39
Abt_4	2228,62	6350,01
Abt_5	2291,93	8641,94
Abt_6	2302,63	10944,57
Abt_3	2459,80	13404,37
Abt_10	2494,34	15898,71
Abt_1	2562,81	18461,52
Abt_7	2671,63	21133,15
Abt_2	2899,33	24032,48
Summe	24032,48	

Wäre der Fehler gemacht worden, die Spalte B vor dem Sortieren zu markieren, dann wäre nur der Inhalt dieser Spalte in eine neue Reihenfolge gekommen – und der Datenbestand wäre zerstört.

Es sei hier ein weiteres Mal darauf hingewiesen, dass die Sortierung statisch ist – d. h. *bei Änderungen im Datenbestand muss stets neu sortiert werden*.

Wird nach richtigem Sortieren eine passende Grafik hergestellt, dann erkennt der Betrachter auf einen Blick den *Beitrag der einzelnen Abteilungen am Gesamtumsatz*.

Weil die absoluten Werte im Regelfall aber meist weniger interessieren und oft auch von geringer Aussagekraft sind, wird später – im Kap. 3 in den Abschn. 3.5.1 und 3.5.2 – ausführlich dargelegt, wie man zu den *relativen Häufigkeiten* und den *relativen Summenhäufigkeiten* kommen kann.

Vorher soll aber den Möglichkeiten, die Excel zur schnellen und attraktiven *grafischen Darstellung von Datenbeständen* bietet, ein besonderes Kapitel gewidmet werden.

Literatur

1. Benker, H.: Wirtschaftsmathematik – Problemlösungen mit EXCEL. Vieweg-Verlag, Wiesbaden (2007)
2. Bourier, G.: Beschreibende Statistik. Gabler-Verlag, Wiesbaden (2013)
3. Duller, C.: Einführung in die Statistik mit EXCEL und SPSS: Ein anwendungsorientiertes Lehr- und Arbeitsbuch. Physica-Verlag, Heidelberg (2010)
4. Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G.: Statistik. Der Weg zur Datenanalyse. Springer-Verlag, Berlin (1997)
5. Fleischhauer, C.: EXCEL in Naturwissenschaft und Technik. Addison Wesley Verlag, München (1998)
6. Jeschke, E., Pfeifer, E., Reinke, H., Unverhau, S., Fienitz, B.: Excel: Formeln und Funktionen. O'Reilly-Verlag, Köln (2014)
7. Leiner, B.: Grundlagen statistischer Methoden. Oldenbourg-Verlag, München Wien (1995)
8. Monka, M., Schöneck, N., Voß, W.: Statistik am PC – Lösungen mit Excel. Hanser Fachbuchverlag, München (2008)
9. Sauerbier, T., Voss, W.: Kleine Formelsammlung Statistik. Carl Hanser Verlag, München (2008)
10. Schells, I.: Excel im Allgemeinen. O'Reilly-Verlag, Köln (2005)
11. Schira, J.: Statistische Methoden der VWL und BWL. Pearson-Verlag, München (2005)
12. Schnell, R.: Graphisch gestützte Datenanalyse. Oldenbourg- Verlag, München, Wien (1994)
13. Schuster, H.: Professionelle Datenauswertung mit Pivot. VNR-Verlag für die deutsche Wirtschaft, Bonn (2009)
14. Untersteiner, H.: Statistik – Datenauswertung mit Excel und SPSS. Verlag UTB, Stuttgart (2007)
15. Zwerenz, K.: Statistik verstehen mit EXCEL. Oldenbourg-Verlag, München Wien (2001)

Statistik und Excel

Elementarer Umgang mit Daten

Matthäus, H.; Matthäus, W.-G.

2016, XXII, 523 S. 107 Abb., Softcover

ISBN: 978-3-658-07688-7