



EXCEL 2016: DATENEINGABE UND - AUSWERTUNG

EINE HANDREICHUNG FÜR
STUDIERENDE DES FACHBEREICHES
SOZIAL- UND KULTURWISSENSCHAFTEN

STAND 03/2020



Excel 2016: Dateneingabe und -Auswertung. Eine Handreichung für die Studierenden des Fachbereichs Sozial- und Kulturwissenschaften.

Stand: März 2020

Arbeitsstelle WAS, Fachbereich Sozial- und Kulturwissenschaften, Hochschule Düsseldorf

Unter Mitarbeit von:

Katja Jepkens, Anne Ritschel, Thorsten Schlee



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz.

Vorwort

Über WAS

Die Arbeitsstelle *Wissenschaftliches Arbeiten und Sozialforschung* wurde im Wintersemester 2011/12 ins Leben gerufen, um einen Beitrag zur Verbesserung der Qualität von Studium und Lehre zu leisten.

Unser Aufgabenschwerpunkt ist die Weiterentwicklung und konzeptionelle Unterstützung der Propädeutik. Aus diesem Grund unterstützen wir den Prozess der Aneignung und Vertiefung sozialwissenschaftlicher Basiskompetenzen. Neben unseren Fortbildungsangeboten für Lehrende und Studierende, einer Lernplattform und einem Semesterapparat in der Fachbibliothek Sozialwesen der Hochschule Düsseldorf sollen auch unsere Reader zum wissenschaftlichen Arbeiten, zum Formatieren in Word und die hier vorliegenden Hinweise zur Datenauswertung mit Excel dazu beitragen.

Weitere Informationen über die Arbeitsstelle WAS sowie unser aktuelles Angebot finden Sie auf unserer Homepage: <https://soz-kult.hs-duesseldorf.de/studium/was>

Zur Auswertung quantitativer Daten mit Excel

Valides Datenmaterial über die Gegenstandsbereiche Sozialer Arbeit ist unerlässlich für eine reflexive Praxis Sozialer Arbeit. Valides Datenmaterial bedarf nicht nur kontrollierter Erhebungsmethoden, sondern auch effektiver Instrumente der Auswertung und Darstellung. MS Excel bietet gute Möglichkeiten auch umfangreiche Datensätze zu bündeln und grafisch darzustellen. Zu beachten ist freilich, dass Excel nicht in erster Linie eine Statistik- und Analysesoftware ist. Die Problemstellungen der schließenden Statistik, mit deren Hilfe Aussagen über das Verhältnis einer Stichprobe zur Grundgesamtheit getroffen werden können, sind sicherlich bei speziell dafür entworfenen Programmen besser aufgehoben. Beachten Sie dazu auch unsere regelmäßigen Schulungen zu SPSS. Zugleich ist Excel ein Kalkulationsprogramm mit umfangreichen Funktionsmöglichkeiten. Erschöpfende Einführungen in die Handhabung des Programms nehmen schnell mehrere hundert Buchseiten ein. Wir haben deshalb gezielt einige Funktionen gesammelt, die es Ihnen ermöglichen sollen, Ihre quantitativen Daten in MS Excel einzugeben und auszuwerten.

Eine solche Handreichung bleibt Entwurf und soll dynamisch bleiben. Deshalb freuen wir uns auf Ihre Rückmeldungen und Erweiterungsvorschläge.

Ihr Team WAS

Kontakt:

Hochschule Düsseldorf
Fachbereich SK Sozial- und Kulturwissenschaften
Arbeitsstelle WAS

Gebäude 3, Raum 03.2.009

Tel.: 0211-4351-2613

E-Mail: was.soz-kult@hs-duesseldorf.de

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einführung..... | 1 |
| 1.1 | <i>Zentrale Begriffe.....</i> | <i>1</i> |
| 1.2 | <i>Ablauf der Auswertung.....</i> | <i>2</i> |
| 2 | Erstellen einer Datenmatrix..... | 3 |
| 3 | Codierung und Eingabe der Daten | 4 |
| 4 | Berechnung deskriptiver Statistiken | 6 |
| 4.1 | <i>Lage- und Streuungsparameter.....</i> | <i>6</i> |
| 4.2 | <i>Häufigkeiten</i> | <i>8</i> |
| 4.3 | <i>Weitere Auswertungstools in Excel: PivotTabellen und Datenanalyse</i> | <i>9</i> |
| 5 | Sortieren und Filtern von Daten | 10 |
| 6 | Grafische Darstellung der Ergebnisse | 13 |

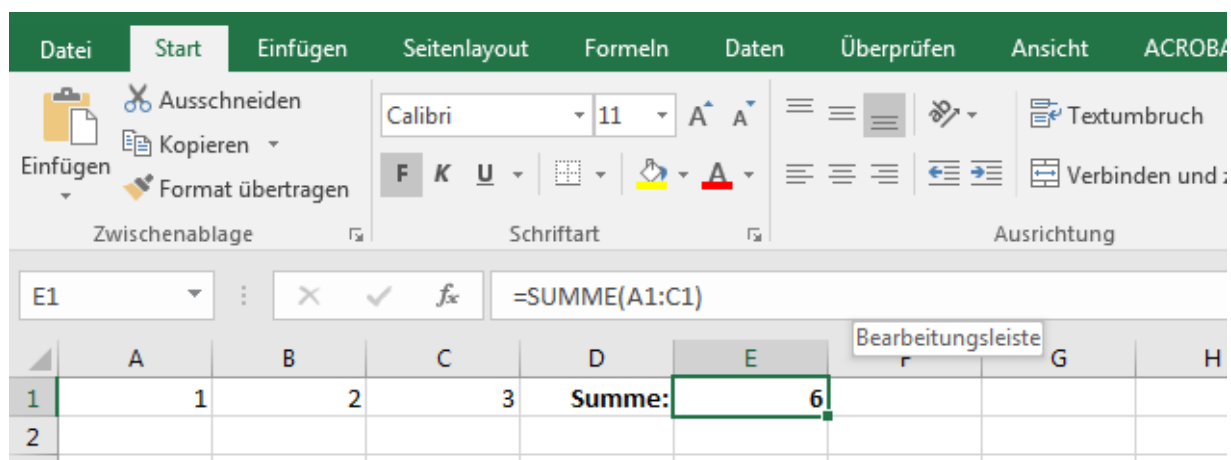
1 Einführung

Microsoft Excel 2016 führt eine Reihe von Bezeichnungen für seine Funktionen, die nicht unbedingt selbsterklärend sind. Wir stellen unseren Hinweisen deshalb eine kurze Erläuterung dieser Bezeichnungen vorneweg. Zudem möchten wir einführend erklären, wie der grundsätzliche Ablauf bei der Eingabe und Auswertung quantitativer Daten in Excel ist, an welchem sich der Aufbau dieses Readers orientiert.

1.1 Zentrale Begriffe

Im Reader bedienen wir uns in Bezug auf Excel folgender Bezeichnungen:

Wenn Sie das Programm öffnen, sehen Sie ein sog. **Menüband** (siehe Screenshot unten), das aus verschiedenen **Registerkarten** (*Start, Einfügen, Seitenlayout, Formeln* usw.) mit jeweils mehreren **Befehlen** besteht. Hier ist die Registerkarte *Start* geöffnet:



Die **Registerkarte** bezeichnet die wichtigste Gliederungs- und Sortierungsebene. Bei den Standard-Voreinstellungen haben Sie folgende Registerkarten zur Auswahl: *Datei, Start, Einfügen, Seitenlayout, Formeln, Daten, Überprüfen* und *Ansicht*. Jede Registerkarte unterteilt sich in **Gruppen** – im Falle der Registerkarte *Start* in die Gruppen *Zwischenablage, Schriftart, Ausrichtung, Zahl, Formatvorlagen, Zellen und Bearbeiten* (s.o.).

Im oberen Teil dieser Gruppen befinden sich die **Befehle**. Mit ihrer Hilfe können Sie zügig auf einzelne Bereiche zugreifen und sie z.B. formatieren. Im unteren Teil finden Sie die Gruppenbezeichnung und teilweise am rechten unteren Rand kleine Pfeile, mit denen Sie das jeweils zugehörige **Dialogfeld** öffnen können. Das ist ein neues Fenster, in dem Sie alle Befehle finden, die Sie in der Registerkarte im Menüband haben, aber darüber hinaus auch weitere Möglichkeiten.

Unterhalb des Menübands sehen Sie zum einen das **Namenfeld**, in dem Ihnen angezeigt wird, in welcher **Zelle** Sie sich gerade befinden – hier die Zelle *E1*, die grün umrahmt ist. Daneben befindet sich die **Bearbeitungsleiste**. Hier sehen Sie die **Formel**, die für die jeweilige Zelle hinterlegt ist. Es sind verschiedene Rechenoperationen (+, -, *, /) sowie die Nutzung von **Funktionen** möglich (im o.g. Beispiel die *Summe* der Zellen *A1* bis *C1*). In der Zelle selbst wird währenddessen das Ergebnis der hinterlegten Funktion als Zahlenwert angezeigt (hier der Wert *6*). Wie Sie diese Funktionen für Ihre Arbeit nutzen können, erfahren Sie in Kapitel 4.

1.2 Ablauf der Auswertung

In Excel können Sie die von Ihnen gesammelten Daten systematisch einpflegen und anschließend auswerten lassen. Im Detail sind dafür folgende Schritte notwendig:

1. *Vergabe von IDs für alle Fälle.* Wenn Sie Ihre Daten gesammelt haben, ist der erste Schritt, die Fragebögen/Testbögen zu nummerieren; so wird jeder befragten Person eine ID-Nummer zugewiesen. Dieses Vorgehen dient dazu, Fragebögen nicht zu verwechseln, doppelt einzugeben oder die Dateneingabe später nachvollziehen zu können. Die Nummerierung kann aus einfachen Zahlen oder bestimmten Codes bestehen – beispielsweise Kürzel für die jeweilige Institution oder Berufsgruppe, wenn dies zutrifft. Diese Bezeichnung soll so gestaltet sein, dass sie nur einmal vergeben wird, also ein Code eindeutig einem Bogen zuzuordnen ist.
2. *Anlegen einer Datenmatrix.* Nachdem Sie das erledigt haben, ist der nächste Schritt die Erstellung einer Datenmatrix (siehe Kapitel 2 *Erstellen einer Datenmatrix*). Dabei legen Sie sozusagen eine Ihrem Fragebogen entsprechende Spalten- und Zeilenbeschriftungen in Excel an, innerhalb der später die einzelnen Fragebögen „eingetippt“ werden.
3. *Dateneingabe.* Es folgt die Codierung der Antwortmöglichkeiten bzw. Variablenausprägung und die Eingabe der Daten in die Datenmatrix – Fragebogen für Fragebogen (siehe Kapitel 3 *Codierung und Eingabe der Daten*)
4. *Datenauswertung.* Sobald alle Ihre gesammelten Daten in Excel eingepflegt worden sind, kann die Datenauswertung starten. Sie können zum Beispiel zentrale Kennwerte der deskriptiven Statistik berechnen (siehe Kapitel 4 *Wichtige Kennzahlen der deskriptiven Statistik*), ihre Datentabelle strukturieren (siehe Kapitel 5 *Sortieren und Filtern von Daten*) und ihre Rechenergebnisse grafisch, bspw. in Form von Diagrammen, darstellen (siehe Kapitel 6 *Grafische Darstellung der Ergebnisse*)

2 Erstellen einer Datenmatrix

Das Ziel der Dateneingabe ist, für jeden Fall einzutippen, was er oder sie im Fragebogen für jede einzelne Frage angegeben hat. Den Rahmen dafür stellt die sogenannte Datenmatrix dar. Diese erstellen Sie, bevor Sie die Daten tatsächlich in Excel einpflegen. Da dieser Rahmen das tabellarische Abbild der Fragen Ihres Fragebogens darstellt, ist es auch möglich, die Datenmatrix schon *vor* der Datenerhebung zu erstellen. Das kann hilfreich sein, um im Vorfeld schon mögliche Schwierigkeiten bei der Dateneingabe auszuräumen, die man sonst übersehen könnte.

Ihre Datenmatrix kann etwa so aussehen:

| | | B | C | D | E | F | G | |
|----|-----|-------|-----------------|------------|-----------|-------|-----------------|---|
| 1. | Nr. | Beruf | Berufserfahrung | Geschlecht | Schulform | Stadt | Stressbelastung | S |
| 2 | SP1 | | | | | | | |
| 3 | LK1 | | | | | | | |
| 4 | SP2 | | | | | | | |
| 5 | LK2 | | | | | | | |
| 6 | SP3 | | | | | | | |
| 7 | SP4 | | | | | | | |
| 8 | LK3 | | | | | | | |
| 9 | SP5 | | | | | | | |

1. **Variablen.** In der ersten Zeile der Matrix stehen Bezeichnungen für alle Variablen, die Sie in Ihrem Fragebogen abgefragt haben. Aus der Frage „*Wie viele Jahre Berufserfahrung haben Sie?*“ wird hier bspw. die Variable „*Berufserfahrung*“. Die Variablennamen sollten möglichst kurz sein, um den Überblick in Ihrer Datenmatrix zu behalten. Selbstverständlich sollten Sie sich an geeigneter Stelle notieren, was die Variablennamen genau bedeuten.
2. **Fälle.** In der ersten Spalte der Matrix stehen die jeweiligen ID-Nummern der Fragebögen (bzw. des Falls, der Untersuchungseinheit...), der in der entsprechenden Zeile eingetragen werden soll. In diesem Beispiel wurde ein zusätzliches Kürzel eingefügt („*SP*“ für Sozialpädagogische Fachkräfte und „*LK*“ für Lehrkräfte). Verwenden Sie hier möglichst die gleiche Codierung, die Sie auch auf den Fragebögen selbst verwendet haben, um die Zuordnung zu vereinfachen.

Ihre Matrix hat also jetzt pro Untersuchungseinheit (= Fall oder Person) eine Zeile, in der nun folgend seine Antworten je Frage eingetragen werden (siehe Kapitel 3 *Codierung und Eingabe der Daten*).

Tip: Bei großen Datenmengen kann es sinnvoll sein, für jede Gruppe (z.B. Stadt, Klasse, Schule, ...) ein eigenes Tabellenblatt anzulegen, da dies übersichtlicher ist als eine Gesamtaufstellung. In jedem Fall sollten Sie aber zusätzlich ein Tabellenblatt mit den Gesamtdaten haben.

3 Codierung und Eingabe der Daten

Die Datenmatrix wartet nun darauf, mit Ihren frisch gesammelten Daten gefüttert zu werden. Das bedeutet konkret, dass Sie zeilenweise für jeden einzelnen Fall eintragen, was in welcher Frage angegeben wurde. Dies erfordert einen Zwischenschritt – die **Codierung** von Antwortvorgaben. In der Regel schreiben Sie keine Begriffe in die Datenmatrix, sondern Zahlen. Das bedeutet, dass jede Antwortvorgabe bei Bedarf in eine Zahl „übersetzt“, d.h. codiert werden muss. So entsteht eine Datenmatrix, deren Inhalt hauptsächlich aus Zahlen besteht. Ähnlich wie bei der Vergabe von Variablennamen sollten Sie sich die Bedeutung der Werte irgendwo notieren.

Das Codierungsverfahren ist abhängig vom jeweiligen Fragetyp. Wir unterscheiden hier **vier Fragetypen**, die wir in den beiden Beispielen für ausgefüllte Datenmatrizen auch dementsprechend gekennzeichnet haben:

1. Frage mit offener Angabe in Form einer Zahl

Es gibt Variablen, deren Ausprägungen per se in Form einer Zahl vorliegen (z. B. Alter, Berufserfahrung in Jahren, Ergebnis einer Klausur als Note). In diesem Fall werden die Werte einfach direkt in die betreffende Zelle eingetragen, eine Codierung ist hier nicht notwendig.

2. Geschlossene Fragen (Kategorien)

Wenn auf eine Frage im Fragebogen verschiedene festlegte Antwortkategorien folgen, ist eine Codierung dieser Kategorien notwendig (z. B. Geschlecht, Bewertungsfragen mit einer Skala von „sehr unzufrieden“ bis „sehr zufrieden“). Angenommen, in einem Fragebogen wurde bei der Variable *Geschlecht* die Ausprägung „männlich“ angekreuzt: Statt den Begriff „männlich“ in die Datenmatrix einzutippen, müssen Sie bspw. eine „1“ eintragen, für „weiblich“ dementsprechend eine „2“. Die Vergabe der Zahlen zu Kategorien erfolgt nach dem logischen Prinzip, dass a) eine Zahl eindeutig einer Kategorie zugeordnet ist und b), dass auf- bzw. absteigende Skalen durch auf- bzw. absteigende Zahlenfolgen ausgedrückt werden. **Achtung:** Dieser Fall tritt nur bei geschlossenen Fragen („Ankreuzfragen“) ein, bei denen nur eine Antwort ausgewählt werden soll. In anderen Fällen müssen Sie bei Typ 3 (Mehrfachnennung) oder 4 (offene Fragen) schauen!

Fragetyp: 2 1 2 2 2 2

| | A | B | C | D | E | F | G | |
|-----|-------|-----------------|------------|-----------|-------|-----------------|---|--|
| Nr. | Beruf | Berufserfahrung | Geschlecht | Schulform | Stadt | Stressbelastung | S | |
| SP1 | | 1 | 3 | 2 | 1 | 4 | 3 | |
| LK1 | | 2 | 7 | 2 | 2 | 3 | 5 | |
| SP2 | | 1 | 30 | 2 | 4 | 1 | 2 | |
| LK2 | | 2 | 6 | 1 | 1 | 4 | 5 | |
| SP3 | | 1 | 15 | 1 | 2 | 2 | 4 | |
| SP4 | | 1 | 10 | 1 | 2 | 3 | 3 | |
| LK3 | | 2 | 4 | 2 | 4 | 1 | 2 | |
| SP5 | | 1 | 9 | 1 | 3 | 4 | 3 | |

3. Mehrfachnennung

Manche Fragen im Fragebogen ermöglichen eine Mehrfachnennung von Antwortkategorien, d.h. die befragte Person darf mehrere Antworten pro Frage ankreuzen. In unserem Beispiel wird nach Stressfaktoren im Berufsalltag gefragt und der Befragte hat die Möglichkeit, aus verschiedenen Stressoren mehr als einen auszuwählen. Die Codierung des Antwortverhaltens nach dem gängigen Muster gestaltet sich hier als schwierig, weshalb ein Trick angewandt wird: Für jede Antwortmöglichkeit wird eine einzelne Variable (und entsprechend eine separate Spalte) angelegt. Die einzelnen Variablen werden weiterführend binär codiert: Es wird eine „1“ eingetragen, wenn sie angekreuzt wurde, und eine „0“, wenn nicht.

4. Offene Fragen

In vielen Fragebögen gibt es mindestens eine Frage, in der die befragte Person selbst eine Antwort formulieren soll (z.B. Frage nach Wünschen oder Gestaltungsvorschlägen). Grundsätzlich übernehmen Sie diese Antworten zunächst wortgetreu in Ihre Datenmatrix, so dass Sie das Rohmaterial möglichst vollständig vorliegen haben. In einem zweiten Schritt (falls nötig) müssen Sie dann selbst entscheiden, nach welchen Kriterien Sie die Antworten zusammenfassen möchten. Bei der Variable „Wünsche“ bietet es sich bspw. an, mehrere Antworten zu einer Kategorie zusammenfassen lassen. Wenn Sie den zweiten Schritt vollzogen haben, können Sie die entstandenen Gruppen codieren (wie bei Variablentyp 2 bzw. 3). Ob und wie Sie den zweiten Schritt vollziehen, hängt immer von Ihrer Fragestellung und von Ihrem Erkenntnisinteresse ab.

FrageTyp: 3 3 3 4

| A | H | I | J | K |
|-----|------------|------------|------------|--------------------------------------|
| Nr. | Stressor 1 | Stressor 2 | Stressor 3 | Wünsche |
| SP1 | 1 | 1 | 0 | kA |
| LK1 | 1 | 0 | 0 | weniger volle Lehrpläne |
| SP2 | 0 | 0 | 1 | kA |
| LK2 | 0 | 1 | 1 | Lehrpläne ausmisten |
| SP3 | 0 | 1 | 1 | klare Aufgabenverteilung |
| SP4 | 1 | 1 | 1 | stärkerer Einbezug regulärer Abläufe |
| LK3 | 0 | 1 | 1 | kA |
| SP5 | 0 | 0 | 1 | bessere Einarbeitung und Einbindung |

Bei fehlenden Werten (keine Antwort, ungültig oder unverständlich) wird in die entsprechende Zelle ein nicht-zahlenförmiges Kürzel wie „kA“ oder ein Wert eingetragen, den diese Variable normalerweise nicht annehmen kann (i.d.R. 999). **Aufgepasst: die Vergabe der Codes für fehlende Angaben muss bei der Auswertungsmethode ggf. berücksichtigt werden!** Eine „999“ darf bspw. nicht in die Mittelwertberechnung einfließen; in vielen Fällen ist daher ein nicht-zahlenförmiges Kürzel einfacher zu handhaben und daher zu bevorzugen.

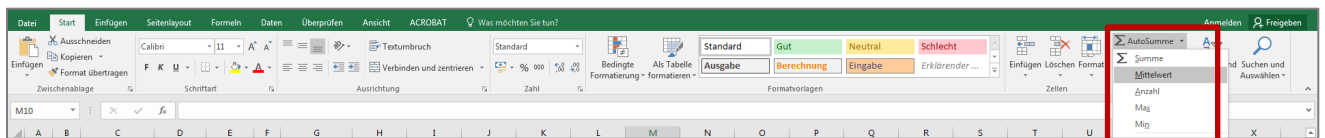
4 Berechnung deskriptiver Statistiken

Mit dem Einsatz eines Fragebogens werden zahlreiche Daten produziert, die sich dann unsortiert in Ihrer Datenmatrix finden. Deskriptive Statistik beschäftigt sich mit der Organisation, Zusammenfassung und Darstellung von Daten, um sie übersichtlich und für eine*n Betrachter*in leicht fassbar zu machen. Aus einer Vielzahl von Daten werden also einige wenige Kennzahlen einer Statistik gebildet, die zentrale Eigenschaften der Datenmenge anzeigen.

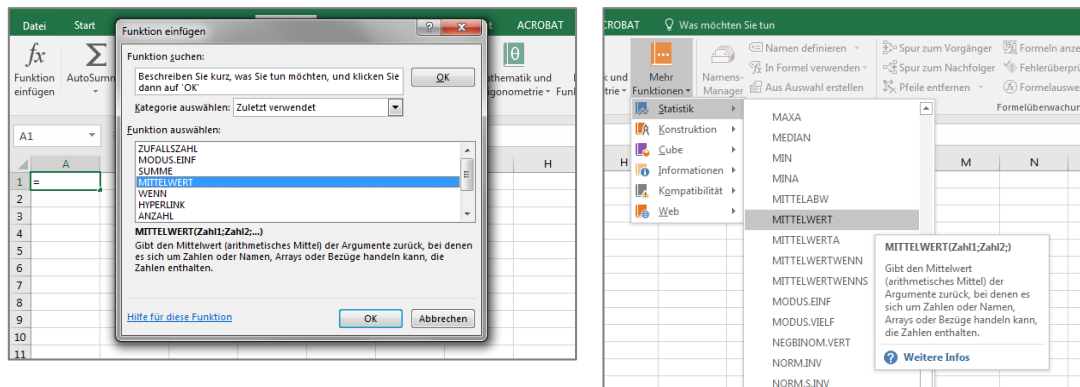
4.1 Lage- und Streuungsparameter

In Excel können Sie sich relativ einfach Lage- und Streuungsmaße wie z.B. das arithmetische Mittel, den Median oder die Standardabweichung ausgeben lassen. Das wichtigste Tool von Excel zur Auswertung von Daten sind die **Funktionen**. Diese gibt es für alle statistischen Kennwerte, die Sie interessieren. Sie gelangen auf unterschiedliche Weisen zu der riesigen Auswahl statistischer Funktionen in Excel. Dies wird im Folgenden an dem Beispiel des Mittelwerts deutlich.

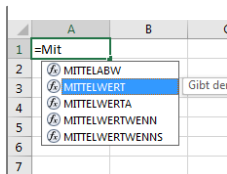
1: über die Schnellwahlleiste (Registerkarte *Start* --> *Autosumme*)



2: über die Registerkarte *Formeln* („Funktion einfügen“ oder „mehr Funktionen – Statistik“)



3: in der Zelle selbst



Am Beispiel des Mittelwerts kann auch gezeigt werden, wie Funktionen funktionieren: Es soll der Mittelwert der Variable *Berufserfahrung* in der Zelle C13 angegeben werden. Diese Zelle wird durch die Funktion „=MITTELWERT(C2:C9)“ definiert. Ausgesprochen bedeutet diese Funktion „Berechne den Mittelwert von der Zelle C2 bis C9“. Den gewünschten Bereich markiert Excel farblich, damit der Zell-

bezug einfach zu überprüfen ist. Diese Funktionsbestimmung können Sie oben in der Bearbeitungsleiste anpassen. Verlässt man die Bearbeitungsleiste, erscheint das Ergebnis in der entsprechenden Zelle.

| Datei Start Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht | | | | | | |
|--|------------------------------|--------------------|------------|-----------|-------|-----------------|
| ZUFALLSZ... | | | | | | |
| =MITTELWERT(C2:C9) | | | | | | |
| Nr. | Beruf | Berufserfahrung | Geschlecht | Schulform | Stadt | Stressbelastung |
| 1 | SP1 | 1 | 3 | 2 | 1 | 4 |
| 2 | LK1 | 2 | 7 | 2 | 2 | 3 |
| 3 | SP2 | 1 | 30 | 2 | 4 | 1 |
| 4 | LK2 | 2 | 6 | 1 | 1 | 4 |
| 5 | SP3 | 1 | 15 | 1 | 2 | 2 |
| 6 | SP4 | 1 | 10 | 1 | 2 | 3 |
| 7 | LK3 | 2 | 4 | 2 | 4 | 1 |
| 8 | SP5 | 1 | 9 | 1 | 3 | 4 |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | statistische Datenauswertung | | | | | |
| 12 | | | | | | |
| 13 | Mittelwert | =MITTELWERT(C2:C9) | | | | |
| 14 | | | | | | |

| Datei Start Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht | | | | | | |
|--|------------------------------|-----------------|------------|-----------|-------|-----------------|
| C13 | | | | | | |
| =MITTELWERT(C2:C9) | | | | | | |
| Nr. | Beruf | Berufserfahrung | Geschlecht | Schulform | Stadt | Stressbelastung |
| 1 | SP1 | 1 | 3 | 2 | 1 | 4 |
| 2 | LK1 | 2 | 7 | 2 | 2 | 3 |
| 3 | SP2 | 1 | 30 | 2 | 4 | 1 |
| 4 | LK2 | 2 | 6 | 1 | 1 | 4 |
| 5 | SP3 | 1 | 15 | 1 | 2 | 2 |
| 6 | SP4 | 1 | 10 | 1 | 2 | 3 |
| 7 | LK3 | 2 | 4 | 2 | 4 | 1 |
| 8 | SP5 | 1 | 9 | 1 | 3 | 4 |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | statistische Datenauswertung | | | | | |
| 12 | | | | | | |
| 13 | Mittelwert | 10,5 | | | | |
| 14 | | | | | | |

| Datei Start Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht | | | | | | |
|--|------------------------------|-----------------|------------|-----------|-------|-----------------|
| C14 | | | | | | |
| =MAX(C2:C9)-MIN(C2:C9) | | | | | | |
| Nr. | Beruf | Berufserfahrung | Geschlecht | Schulform | Stadt | Stressbelastung |
| 1 | SP1 | 1 | 3 | 2 | 1 | 4 |
| 2 | LK1 | 2 | 7 | 2 | 2 | 3 |
| 3 | SP2 | 1 | 30 | 2 | 4 | 1 |
| 4 | LK2 | 2 | 6 | 1 | 1 | 4 |
| 5 | SP3 | 1 | 15 | 1 | 2 | 2 |
| 6 | SP4 | 1 | 10 | 1 | 2 | 3 |
| 7 | LK3 | 2 | 4 | 2 | 4 | 1 |
| 8 | SP5 | 1 | 9 | 1 | 3 | 4 |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | statistische Datenauswertung | | | | | |
| 12 | | | | | | |
| 13 | Mittelwert | 10,5 | | | | |
| 14 | Spannweite | 27 | | | | |
| 15 | | | | | | |

Funktionen können auch miteinander kombiniert werden, was am Beispiel der Berechnung der Spannweite/Range gezeigt werden kann. Excel bietet für dieses Streuungsmaß leider keine eigene Funktion an.

Mit der entsprechenden Funktion können Sie genauso den Median (=MEDIAN), Modus/Modalwert (=MODUS.EINF), die Standardabweichung (=STABW.S) und weitere Lage- und Streuungsmaße einer Datenreihe berechnen.

Eine Besonderheit gibt es bei der Berechnung der **Standardabweichung**. Excel bietet hierfür zwei unterschiedliche Funktionen an: „STABW.N“ und „STABW.S“. Dahinter steckt die Unterscheidung von Stichprobe und Grundgesamtheit bzw. geschätzter/korrigierter und nicht-korrigierter Standardabweichung. An dieser Stelle gehen wir nicht auf die mathematischen Hintergründe dieser Unterscheidung ein. Bei der Berechnung der **Standardabweichung** merken Sie sich Folgendes: Die Funktion „**STABW.N**“ nutzen Sie, wenn die befragten Personen die Grundgesamtheit bilden, wenn Sie also keine Stichprobe gezogen, sondern alle Personen der interessierenden Zielgruppe befragt haben; hier ist keine Korrektur der Formel notwendig. Mit der Funktion „**STABW.S**“ schätzen Sie die Standardabweichung aus der Ihnen vorliegenden Stichprobe. Die Formel zur Berechnung der Standardabweichung wird korrigiert. Da Sie für ein sozialwissenschaftliches Forschungsprojekt i.d.R. eine Stichprobe aus der Grundgesamtheit ziehen, verwenden Sie meistens die geschätzte Standardabweichung STABW.S.

4.2 Häufigkeiten

Excel kann Ihnen auch die absolute und relative Häufigkeit der Nennungen einzelner Werte einer Variable anzeigen – also berechnen, wie oft eine bestimmte Antwort gegeben wurde. Dazu muss zunächst eine neue Tabelle angelegt werden. Wir zeigen Ihnen hier, wie Sie eine Häufigkeitstabelle manuell erstellen können.

1. **Häufigkeitstabelle anlegen:** Legen Sie fest, wie Ihre neue Tabelle strukturiert werden soll. Dazu müssen Sie alle möglichen Werte zeilenweise untereinander notieren und bestenfalls in einer Spalte daneben die entsprechende Wertebeschriftung (die Codierung) einfügen. In der oberen Zeile geben Sie ein, was berechnet wird (absolute und relative Häufigkeiten). Die Tabelle kann nun aussehen wie in der folgenden Abbildung. Hier soll die Häufigkeit der männlichen und weiblichen Befragten berechnet werden.
2. **Berechnung der absoluten Häufigkeit:** Markieren Sie die Zelle, in der die absolute Häufigkeit angegeben werden soll. Nutzen Sie die Funktion „ZÄHLENWENN(Bereich;Suchkriterien)“.¹ Mit dieser Funktion suchen Sie in einem von Ihnen definierten Bereich nach einem bestimmten Ausdruck. In diesem Beispiel suchen Sie in dem Bereich D2 bis D9 (Spalte der Variable „Geschlecht“) nach der Zahl in der Zelle B13 (also nach der „1“ und damit der Kategorie „männlich“). Die gleiche Funktion nutzen Sie bei der Kategorie „weiblich“.

| Nr. | Beruf | Berufserfahr | Geschlecht | Schulform | Stadt | Stressbelastung |
|--------------------------------------|-------|--------------|------------|-----------|-------|-----------------|
| SP1 | | 1 | 3 | 2 | 1 | 4 |
| LK1 | | 2 | 7 | 2 | 2 | 3 |
| SP2 | | 1 | 30 | 2 | 4 | 1 |
| LK2 | | 2 | 6 | 1 | 1 | 4 |
| SP3 | | 1 | 15 | 1 | 2 | 2 |
| SP4 | | 1 | 10 | 1 | 2 | 3 |
| LK3 | | 2 | 4 | 2 | 4 | 1 |
| SP5 | | 1 | 9 | 1 | 3 | 4 |
| Häufigkeitsverteilung für Geschlecht | | | | | | |
| | | | absolut | relativ | | |
| männlich | | | 1 | | | |
| weiblich | | | 2 | | | |
| Gesamt | | | | | | |

| Nr. | Beruf | Berufserfahr | Geschlecht | Schulform | Stadt | Stressbelastung |
|--------------------------------------|-------|--------------|------------------------|-----------|-------|-----------------|
| SP1 | | 1 | 3 | 2 | 1 | 4 |
| LK1 | | 2 | 7 | 2 | 2 | 3 |
| SP2 | | 1 | 30 | 2 | 4 | 1 |
| LK2 | | 2 | 6 | 1 | 1 | 4 |
| SP3 | | 1 | 15 | 1 | 2 | 2 |
| SP4 | | 1 | 10 | 1 | 2 | 3 |
| LK3 | | 2 | 4 | 2 | 4 | 1 |
| SP5 | | 1 | 9 | 1 | 3 | 4 |
| Häufigkeitsverteilung für Geschlecht | | | | | | |
| | | | absolut | relativ | | |
| männlich | | | =ZÄHLENWENN(D2:D9;B13) | | | |
| weiblich | | | | | | |
| Gesamt | | | | | | |

3. **Berechnung der relativen Häufigkeit in Prozent:** Berechnen Sie die relative Häufigkeit, indem Sie

| Nr. | Beruf | Berufserfahr | Geschlecht | Schulform | Stadt | Stressbelastung |
|--------------------------------------|-------|--------------|------------|-----------|-------|-----------------|
| SP1 | | 1 | 3 | 2 | 1 | 4 |
| LK1 | | 2 | 7 | 2 | 2 | 3 |
| SP2 | | 1 | 30 | 2 | 4 | 1 |
| LK2 | | 2 | 6 | 1 | 1 | 4 |
| SP3 | | 1 | 15 | 1 | 2 | 2 |
| SP4 | | 1 | 10 | 1 | 2 | 3 |
| LK3 | | 2 | 4 | 2 | 4 | 1 |
| SP5 | | 1 | 9 | 1 | 3 | 4 |
| Häufigkeitsverteilung für Geschlecht | | | | | | |
| | | | absolut | relativ | | |
| männlich | | | 1 | 50% | | |
| weiblich | | | 2 | | | |
| Gesamt | | | 8 | | | |

die absolute Häufigkeit in Relation zu *allen* Fällen (n) setzen (hier: C13/C15). Die relative Häufigkeit wird zunächst als Dezimalstelle ausgegeben. Mit der Option *Zellen formatieren* können Sie sich das Ergebnis auch als Prozentzahl ausgeben lassen. Die gleiche Prozedur wiederholen Sie bei allen anderen Kategorien.

Tipp: Bei sehr vielen verschiedenen Kategorien kann es sinnvoll sein, die Funktionen in die entsprechenden Zellen zu kopieren, bevor Sie jede Funktion immer wieder eintippen. Da sich die Zellbezüge mit dem Kopieren in andere Zellen mitverschieben, sollten Sie Zeilen, Spalten oder Zellen, die „fest“ stehen (wie etwa C15, die Gesamtzahl der Fälle n) entsprechend markieren. Dies machen Sie mit Dollarzeichen, die Sie der betroffenen Zeile und Spalte voranstellen („\$C\$15“). Nun können Sie die Funktion kopieren, ohne dass sich der Zellbezug ändert.

¹ Alternativ können Sie die [Funktion „HÄUFIGKEIT“](#) benutzen.

4.3 Weitere Auswertungstools in Excel: Pivot-Tabellen und Datenanalyse

Bei komplexeren Analysen lohnt es sich, sich mit alternativen Auswertungsmöglichkeiten in Excel auseinanderzusetzen. Hier schauen wir uns kurz die Pivot-Tabellen und das Datenanalyse-Tool an.

Mit [Pivot-Tabellen](#) können Sie benutzerdefinierte Tabellen für verschiedene Variablen und Variablenkombinationen anlegen (z.B. bivariate Statistik). Markieren Sie dazu Ihre Datenmatrix und klicken Sie unter der Registerkarte *Einfügen* auf *Pivot-Tabelle*. Es öffnet sich ein neues Tabellenblatt, welches erstmal leer ist. Mit Hilfe der Tafel auf der rechten Seite des Tabellenblatts definieren Sie, welche Variablen sie miteinander kombinieren und welche Kenngröße Sie berechnen wollen. Beispielsweise können Sie eine Tabelle generieren, in der Mittelwerte in einzelnen Gruppen dargestellt werden.

Mit dem [Datenanalyse-Tool](#) lassen sich verschiedene Berechnungen, darunter deskriptive Datenauswertungen und statistische Tests, durchführen. Sie finden es unter der Registerkarte *Daten*. Es werden Ihnen eine ganze Reihe von Analyse-Funktionen eröffnet. Um z.B. an eine Zusammenfassung üblicher Lage- und Streuungsparameter zu gelangen, wählen Sie *Populationskenngrößen* aus. Beachten Sie: Das Datenanalyse-Tool muss ggf. als Add-In installiert werden.

Wie Sie sehen, ist es auch ohne Funktionen möglich, statistische Kennwerte zu berechnen. Der Nachteil dieser Variante ist jedoch, dass die so erstellten Auswertungstabellen **nicht dynamisch** sind. Das heißt, sobald Sie etwas in Ihrer Datenmatrix ändern, müssen Sie die Datenanalyse erneut durchführen. Bei der manuellen Auswertung haben Sie dieses Problem nicht.

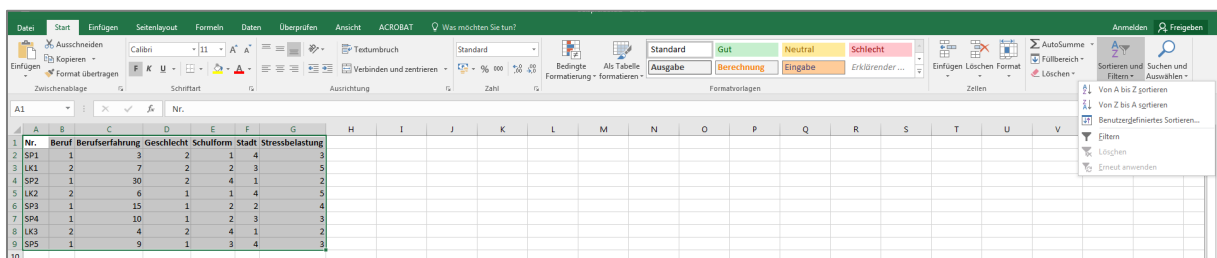
Tipp: Für weitere Informationen über Excel-Funktionen nutzen Sie die [Office-Support-Website](#) oder das Hilfe-Menü im Programm selbst („Hilfe für diese Funktion“).

Komplexere Datenanalysen gestalten sich in anderer Statistik-Software wie SPSS wesentlich benutzerfreundlicher.

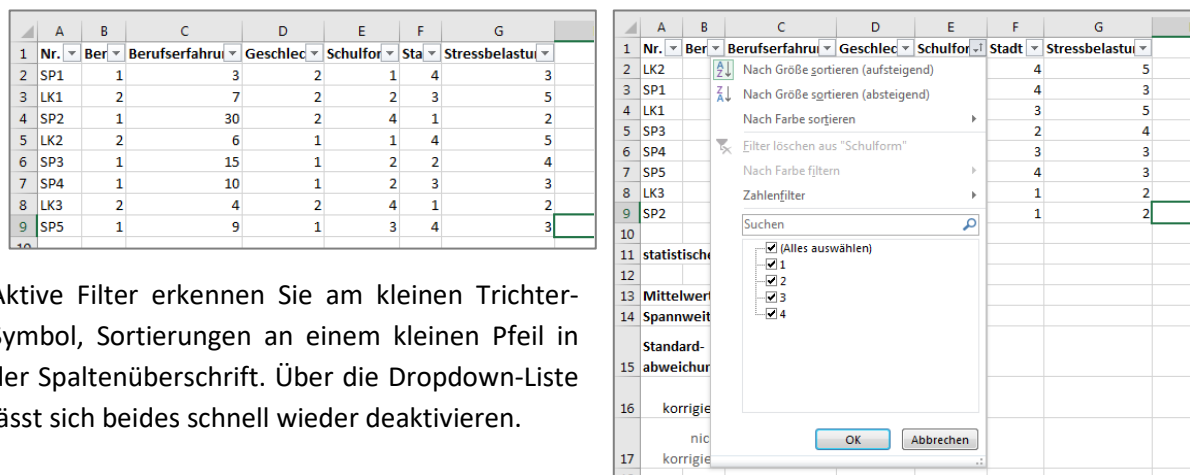
5 Sortieren und Filtern von Daten

Nachdem Sie die Daten eingegeben haben, sind die Datensätze zuerst in der Reihenfolge sortiert, in der Sie diese eingegeben haben. Excel bietet Ihnen die Möglichkeit, Ihre Daten nach verschiedenen Kriterien zu sortieren oder auch zu filtern, so dass nur bestimmte Datensätze angezeigt werden. Zudem können Sie bestimmte Daten optisch hervorheben. Das kann Ihnen dabei helfen, einen ersten Überblick über Ihre Daten zu bekommen.

Am einfachsten gelangen Sie zu den Funktionen Sortieren und Filtern, indem Sie Ihre gesamte Tabelle markieren, anschließend auf das Feld *Sortieren und Filtern* und dann auf *Filtern* klicken. Achten Sie darauf, dass Sie die gesamte Datenmatrix markieren. Nicht ausgewählte Spalten oder Zeilen werden ansonsten ggf. nicht mitsortiert – Datenchaos droht.



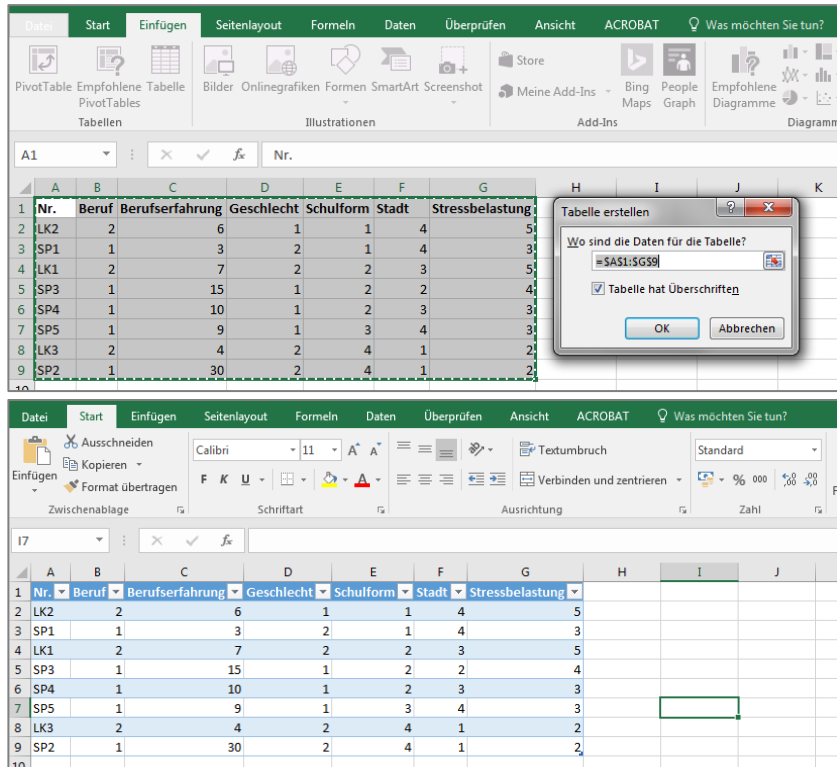
Excel erkennt selbst, dass die erste Zeile Ihrer Datenmatrix Spaltenüberschriften enthält, so sollte hinter jeder Überschrift ein kleiner Pfeil erscheinen. Diesen klicken Sie an und es öffnet sich eine Dialogbox, durch die Sie verschiedene Optionen haben, die Datenmatrix nach der ausgewählten Variable zu sortieren oder durch Auswahl bestimmter Variablenausprägungen zu filtern. Excel zählt hier alle Zellinhalte auf, die unter der betroffenen Spalte von Ihnen eingetragen worden sind. Sie können Daten filtern, indem Sie die Häkchen im unteren Teil des Feldes nur bei den Daten setzen, die angezeigt werden sollen (z.B. möchten Sie nur die Fälle anschauen, die an Gymnasien arbeiten und setzen deshalb nur bei 1 ein Häkchen).



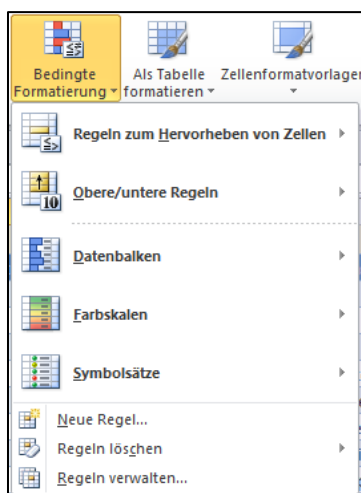
Aktive Filter erkennen Sie am kleinen Trichter-Symbol, Sortierungen an einem kleinen Pfeil in der Spaltenüberschrift. Über die Dropdown-Liste lässt sich beides schnell wieder deaktivieren.

Tip: Über die Schaltfläche *benutzerdefiniertes Sortieren* können Sie mehrere Sortierebenen einfügen und sogar eine neue, selbstdefinierte Reihenfolge einstellen, nach der sortiert werden soll. Excel sortiert schließlich per Default-Einstellung nur von A bis Z bzw. auf- oder absteigend. Bei Variablenkategorien, die nicht als Text vorliegen, aber nicht nur alphabetisch sortiert werden könnten (wie *hoch – mittel – niedrig*), ist diese Funktion sehr hilfreich.

Es kann mitunter hilfreich sein, seine **Datenmatrix als Tabelle zu formatieren**. Dazu markieren Sie die gesamte Datenmatrix und klicken unter *Einfügen* auf *Tabelle*. In dem Dialogfeld wird angezeigt, welcher Zellenbereich nun als Tabelle angezeigt werden soll. Da jede Datenmatrix Spaltenüberschriften enthält, aktivieren Sie das Feld *Tabelle hat Überschriften*. Die so erzeugte Tabelle enthält die oben genannten Funktionen automatisch und ist zudem übersichtlich gestaltet.



Bedingte Formatierung



Unter der Option *Filter* wurde oben die Möglichkeit beschrieben, sich nur bestimmte Werte, wie z.B. jene, die unter oder über dem Durchschnitt der Variable liegen oder die 10 höchsten Werte insgesamt, anzeigen zu lassen. Sie können diese Werte aber auch durch eine sogenannte bedingte Formatierung hervorheben, anstatt Sie herauszufiltern. Im Unterschied zum oben beschriebenen Filtern werden dann weiterhin alle Werte angezeigt; es findet lediglich eine optische Hervorhebung der für Sie relevanten Daten durch die bedingte Formatierung statt.

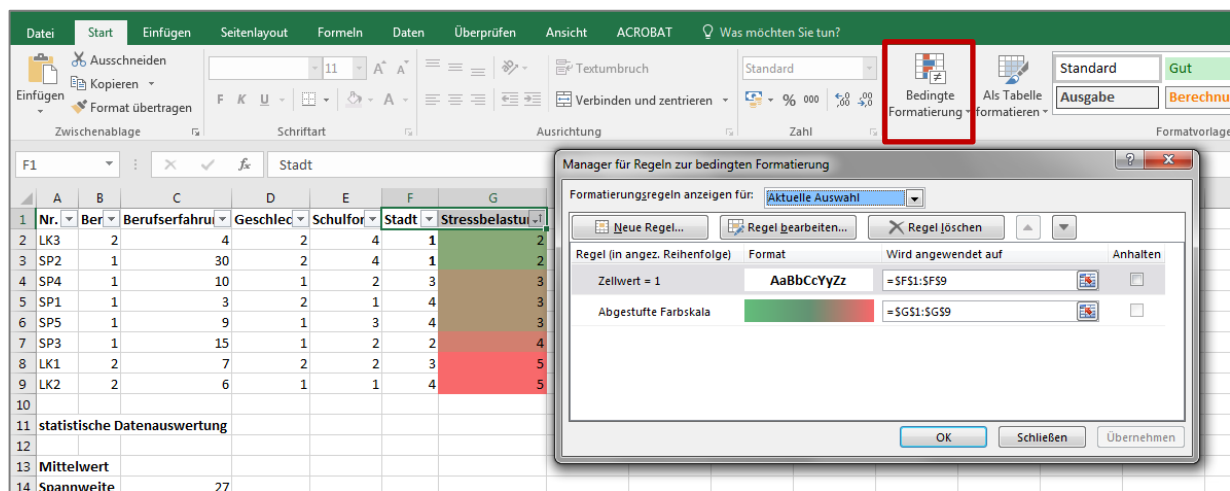
Die Möglichkeit zur optischen Kennzeichnung finden Sie in der Registerkarte *Start* in der Gruppe *Formatvorlagen*. Wenn Sie dort auf *Bedingte Formatierung* klicken, öffnet sich ein Dialogfeld, das Ihnen eine ganze Reihe von voreingestellten Gestaltungsmöglichkeiten bietet.

Regeln zum Hervorheben von Zellen sind sinnvoll, wenn Sie Ihre Daten mit einem bestimmten Wert abgleichen wollen. Diesen können Sie selbst festsetzen oder auch einen Zellbezug nutzen. So wird jede Zelle in einer von Ihnen ausgewählten Form markiert (z.B. hellrot), die bspw. über 5 oder dem Wert in

der Zelle C15 liegt. *Obere/untere Regeln* eignen sich, um Zellen hervorzuheben, die in Bezug zur gesamten Datenreihe auffällig sind. Dies sind beispielsweise Extremwerte oder Werte über bzw. unter dem Durchschnitt.

Sie können neben dieser reinen Hervorhebung auch *Datenbalken*, *Farbskalen* oder *Symbolsätze* einfügen. Farbskalen eignen sich beispielweise bei Benotungen, um auf dem ersten Blick besonders gute, mittlere und schlechte Bewertungen zu identifizieren. Hier gilt: je nach Variable bieten sich unterschiedliche Darstellungen an.

Unterhalb der von Excel vorgeschlagenen Formatierungen können Sie in allen Fällen auf *Weitere Regeln* klicken. Dort haben Sie die Möglichkeit, eine Formatierung ganz nach Ihrem Bedarf anzupassen. Probieren Sie hier einfach aus und klicken Sie sich durch die verschiedenen Möglichkeiten durch!



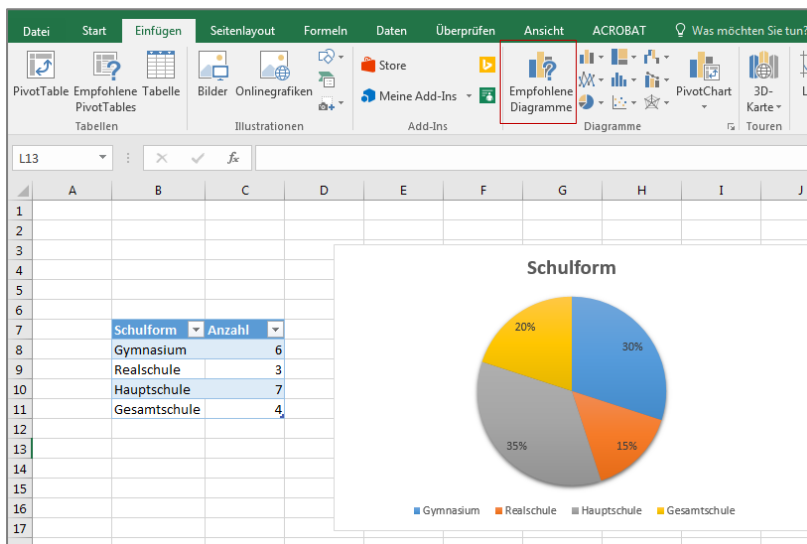
In diesem Beispiel haben wir für die Variable „Stadt“ und „Stressbelastung“ jeweils eine Regel definiert. Immer, wenn bei Stadt eine 1 eingetragen wird, wird diese automatisch **fett** angezeigt; bei der Stressbelastung entsteht eine Farbskala, die die Stärke der Ausprägung von 1 bis 5 widerspiegeln soll. Klicken Sie unter *Bedingte Formatierung* auf *Regeln verwalten*, um sich wie hier alle Regeln anzeigen zu lassen, zu bearbeiten und zu löschen.

6 Grafische Darstellung der Ergebnisse

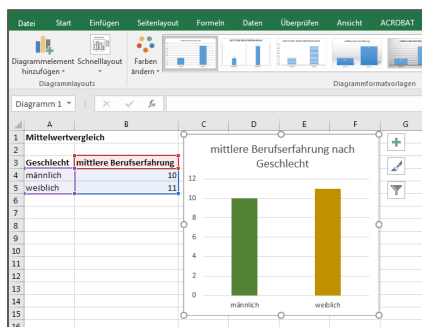
Für Ihren Forschungsbericht werden Sie regelmäßig grafische Darstellungen Ihrer Ergebnisse in Form von Diagrammen einfügen. Abhängig davon, welche Aussage Sie mit Ihren Daten treffen wollen, bieten sich je unterschiedliche Diagrammtypen an. Entsprechend bietet Excel eine große Auswahl an grafischen Darstellungsweisen Ihrer Daten. Die Diagrammtypen finden Sie in der Registerkarte *Einfügen* in der Gruppe *Diagramme*. Unter *Empfohlene Diagramme* können Sie sich bei vorheriger Markierung der Tabelle eine Vorschau verschiedener Diagrammtypen anzeigen lassen – sehr praktisch!

Darstellung von Häufigkeiten

I.d.R. generieren Sie Kreis-, Balken- und Säulendiagramme nicht aus Ihren Rohdaten, sprich direkt aus der Datenmatrix. Sie müssen die Daten zuvor schon weiterverarbeitet haben, indem Sie z.B. die Häufigkeiten der einzelnen Werte oder die Durchschnittswerte mehrerer Variablen errechnen lassen. Für die Darstellung von Häufigkeiten muss zunächst also eine entsprechende Tabelle vorliegen, die Sie, wie oben beschrieben, erstellen können. Markieren Sie dann die relevanten Teile Ihrer Tabelle inkl. der Spaltenbeschriftung (Spalte 1: Variable mit Ausprägungen und Spalte 2: entsprechende Häufigkeit/absolute Anzahl). Wählen Sie nun einen für Sie passenden Diagrammtyp aus. Zum besseren Verständnis bietet es sich übrigens an, die in Zahlen codierten Kategorien wieder in Texte zu „übersetzen“.



Darstellung von Mittelwerten

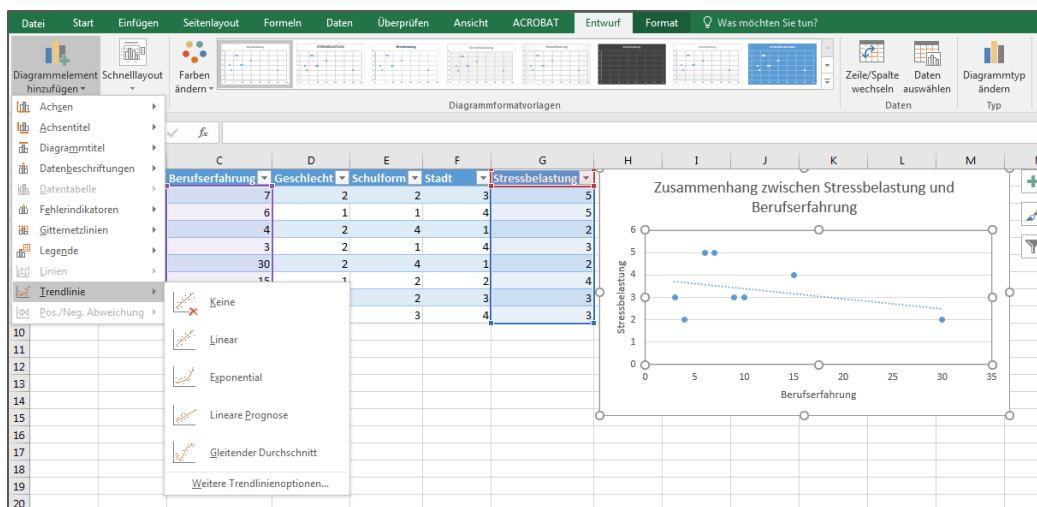


Möchten Sie Mittelwerte für verschiedene Gruppen als Diagramm darstellen lassen, müssen Sie auch dafür zunächst eine eigene Tabelle anlegen. In der ersten Spalte sollten die Kategorien bzw. Gruppen liegen (z.B. männlich und weiblich), in der zweiten Spalte der entsprechende Mittelwert einer anderen Variable (z.B. die mittlere Berufserfahrung). Lassen Sie sich nun bspw. ein Säulendiagramm ausgeben.

Darstellung von Zusammenhängen

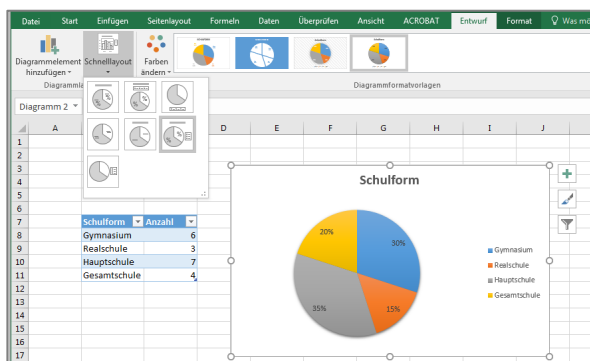
Um zwei metrische Variablen grafisch einander in Beziehung zu setzen, bietet sich i.d.R. ein Punktdiagramm (auch Streudiagramm genannt) an. In diesem Fall ist es nicht nötig, Ihre Daten tabellarisch neu zu ordnen, sie können direkt auf die Rohntabelle zurückgreifen.

Markieren Sie dazu die beiden Variablenspalten, für die Sie ein Diagramm erstellen möchten, und lassen Sie sich ein Punktdiagramm erstellen. Das Einfügen und die Auswahl der Bezugsdaten folgt der gleichen Systematik, wie bei der Erstellung von Kreisdiagrammen. Über das Feld „Diagrammelement hinzufügen“ können Sie sich zusätzlich eine Trendlinie zum Zusammenhang der beiden Variablen ausgeben lassen. Hier zeichnet sich ein Zusammenhang ab: Je höher die Berufserfahrung, desto geringer die Stressbelastung.



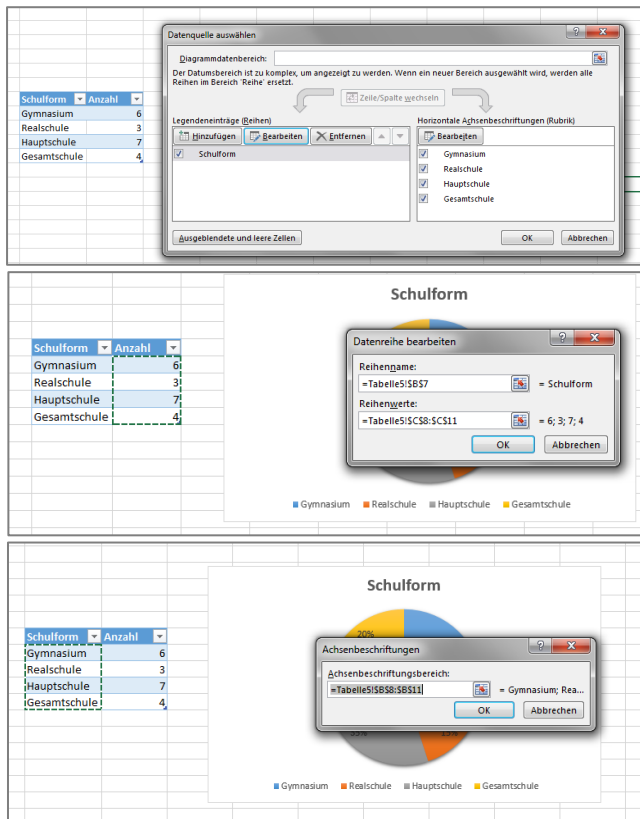
Tip: Über die Schaltfläche *Pivot-Charts* können Sie alle Diagramme direkt aus der Rohntabelle erstellen. Dies erfordert ein wenig Einarbeitungszeit, kann sich aber lohnen.

Anpassen und Bearbeitung von Diagrammen



Sie können die konkrete Gestaltung eines Diagramms nach dem Einfügen verändern. Die obige Darstellungsvariante erhalten Sie, indem Sie das Diagramm anklicken und unter *Diagrammtools* in der Registerkarte *Entwurf* in die Gruppe *Layouts* gehen. Dort haben Sie verschiedene Optionen Ihr Diagramm zu gestalten, einzelne Bereiche der Grafik hervorzuheben usw. Auch hier gilt: Probieren Sie die verschiedenen Darstellungsvarianten einfach einmal aus.

Mit dem Befehl *Daten auswählen* können Sie bei Bedarf nachträglich auch die Datenbezüge Ihres Diagramms verändern. Es öffnet sich folgendes Dialogfeld:



Im *Diagrammdatenbereich* können Sie den Datenbezug insgesamt anpassen und die Teile der Tabelle markieren, die Sie in Ihr Diagramm berücksichtigen wollen.

Unter *Legendeneinträge (Reihen)* bearbeiten Sie den Namen („Schulform“) und die Werte bzw. den Bereich der Daten, der hier abgebildet werden soll (die einzelnen Häufigkeiten) für die darzustellende Datenreihe. In diesem Beispiel legen Sie fest, dass die Werte aller 4 Schultypen in das Diagramm aufgenommen werden soll.

Die Benennungen Ihrer zu den Werten gehörenden Kategorien (hier die Schulformen 1-4) verändern Sie, indem Sie die *Horizontale Achsenbeschriftung* bearbeiten und analog die Bezugsreihe verändern.