

Libreoffice Calc

Eine Einführung in das Tabellenkalkulationsprogramm

The screenshot displays the LibreOffice Calc interface with a data table, a bar chart, a pie chart, and a bar chart with error bars. The data table is as follows:

ID	Versuch	Probenummer	Messwert	Anteil an der Gesamtsumme (in %)	Messwert-Klassen	Häufigkeit
8	A	3	1.1	1		
9	A	4	0.6	0		
10	A	5	19.4	10		
11	B	1	20.1	11		
12	B	2	17.5	9		
13	B	3	12.3	7		
14	B	4	14.8	8		
15	B	5	9.6	5		
16	B	1	8.2	4		
17	B	2	11.4	6		
18	B	3	13.0	7		
19	B	4	7.8	4		
20	B	5	8.9	5		

Summary statistics:

- Summe: 188.2
- Mittelwert: 9.4
- Standardabweichung: 5.8
- Minimalwert: 0.6
- Maximalwert: 20.1
- Teilsomme: 109.7

Comparison of Versuch A and Versuch B:

Metric	Versuch A	Versuch B
Mittelwert Messwert	6.457	12.36
Standardabweichung	5.82672397	4.10018969750859

The bar chart 'Häufigkeit der Messwerte' shows the frequency of measurement values. The x-axis represents 'Messwerte' (4, 8, 12, 16, 20, 24) and the y-axis represents 'Häufigkeit' (0 to 6). The bars show frequencies of 4, 4, 5, 4, 2, and 1 respectively.

The pie chart shows the distribution of 20 data points across 20 categories, with a legend provided.

The bar chart 'Durchschnittliche Messwerte pro Versuch' shows the average measurement values for Versuch A and Versuch B. The x-axis represents 'Versuch A' and 'Versuch B', and the y-axis represents 'Messwert' (0 to 20). The bars show average values of approximately 6.5 and 12.4, with error bars indicating standard deviation.

Functions panel (Funktionen) is visible on the right, showing a list of functions including MITTELWERT, SUMME, ABRUNDEN, etc.

Libreoffice Calc

Eine Einführung in das Tabellenkalkulationsprogramm

Saskia Otto¹

1. Institut für Marine Ökosystem- und Fischereiwissenschaften, Fachbereich
Biologie, Universität Hamburg

Nov. 2022 (Version 4.0)



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR MATHEMATIK, INFORMATIK
UND NATURWISSENSCHAFTEN

Inhaltsverzeichnis

1	Allgemeines zu Calc	1
1.1	Installation von LibreOffice	1
1.2	Öffnen von Calc	1
1.3	Die grafische Oberfläche (GUI)	1
1.4	Aufbau eines Tabellenblatts	2
1.5	Wie sollten Daten idealerweise strukturiert sein?	3
2	Tabellen anlegen	5
2.1	Direkte Dateneingabe und Datentypen	5
2.1.1	Inhalte automatisch ausfüllen	5
2.1.2	Daten in Calc transponieren	7
2.2	Datenimport	7
2.2.1	CSV-Dateien direkt öffnen	8
2.2.2	TXT-Dateien und andere Dateitypen öffnen	10
2.2.3	Import einer Tabelle als weiteres Tabellenblatt	11
2.3	Dateispeicherung nicht vergessen	12
2.4	Tabellen übersichtlicher gestalten	12
2.4.1	Spaltenbreite und Zeilenhöhe ändern	12
2.4.2	Textausrichtung	13
2.4.3	Zeilen und Spalten fixieren	13
2.4.4	Daten filtern und sortieren	14
2.4.5	Zellen formatieren - Komma vs. Punkt	15
2.4.6	Umbenennen von Tabellenblättern	16
3	Rechnen in Calc	19
3.1	Wie geht das?	19
3.2	Warum nicht mit Zahlen rechnen?	20
3.3	Eingeben von Funktionen	21
3.3.1	Built-in Funktionen	21
3.3.2	Zellauswahl	23
3.3.3	Zellbezüge mit \$ fixieren	23
3.3.4	Beispiel einer typischen Rechenanwendung	25
3.4	Zellen verbinden mit der VERKETTEN() Funktion	26
3.5	Datenaggregation mit Pivot-Tabellen	27
3.6	Statistische Analysen automatisch generieren	29
3.7	Aktivieren von Add-Ins	30
4	Diagramme	34
4.1	Einfaches Streudiagramm mit Trendlinie	35
4.2	Datenaggregation - Beispiel Säulendiagramm	36
4.3	Datenaggregation - Beispiel Histogramm	37

5 Tabellenexport zur weiteren Verwendung	52
5.1 Vorbereitungen zum Export	52
5.2 Export von CSV Dateien	52
6 Problembehandlung	56
6.1 Wenn Calc Zahlen als Text erfasst	56

1 Allgemeines zu Calc

Calc ist wie alle anderen Tabellenkalkulationsprogramme eine Software für die **interaktive Eingabe** und **Verarbeitung** von numerischen und alphanumerischen **Daten** in Form einer Tabelle. Vielfach erlaubt Calc zusätzlich die **grafische Darstellung** der Ergebnisse in verschiedenen Anzeigeformen.

Mögliche Anwendungen sind

- die Digitalisierung von Daten,
- Betrachtung von Daten,
- Datenumstrukturierung,
- einfache Berechnungen und numerische Zusammenfassungen,
- grafische Darstellung der Daten zur Datenexploration und
- Export von Daten für die Nutzung in anderen Programmen (z.B. R).

Dieses Handbuch versucht, einen Einstieg und ersten Überblick der wichtigsten Funktionalitäten von Calc zu geben. Die komplette Bandbreite der Datenhandling- und Analysemöglichkeiten in Calc wird nicht aufgeführt, da wir in den Data Science Modulen hauptsächlich in der Programmiersprache R arbeiten werden. Falls Du Dich jedoch mehr mit dem Programm beschäftigen möchtest, empfehle ich Dir die ausführliche Dokumentation der [Libreoffice Hilfe-Website](#).

1.1 Installation von LibreOffice

Calc bzw. das ganze LibreOffice Paket lässt sich über die [Homepage von Libreoffice](#) herunterladen. Dazu wählt man in der Dropdown-Liste das entsprechende Betriebssystem aus, z.B. Windows (64-bit). Mit einem Klick auf "Download" und einer Bestätigung des Speicherortes wird die Installationsdatei herunter geladen. Mit einem Doppelklick auf die heruntergeladene Datei öffnet sich der Installationsassistent, dem man einfach nur zu folgen braucht. Fertig!

1.2 Öffnen von Calc

Beim Starten von Libreoffice öffnet sich jedesmal der Startcenter (Abb. 1.1), von dem man aus bestehende Dateien oder auch ein neues Calc Tabellendokument öffnen kann.

1.3 Die grafische Oberfläche (GUI)

Wie andere Tabellenkalkulationsprogramme auch, besteht die GUI von Calc hauptsächlich aus dem Tabellenblatt selbst (Abb. 1.2). Darüber hinaus gibt es die Menüleiste ganz oben, die Symbolleiste mit den Schnellzugriff-Schaltflächen für die wichtigsten Software-Funktionen (nicht zu verwechseln mit den Funktionen, die man auf die Daten anwenden kann) und Formatierungen.

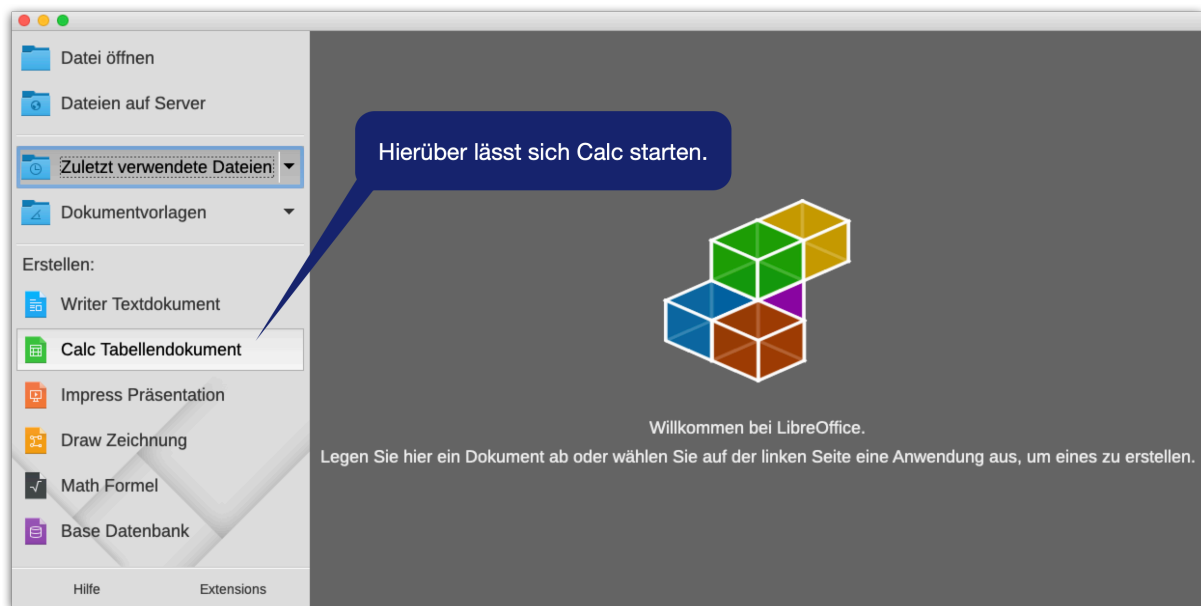


Abbildung 1.1: Das Startcenter beim Öffnen von Libreoffice.

Beides lässt sich auch über das Hauptmenü steuern, hier geht es nur einfach schneller. Darunter befindet sich die Rechenleiste mit der Bearbeitungszeile für die Formeleingabe. Alle beiden Symbolleisten und die Rechenleiste lassen sich auch über die *Ansicht* in der Menüleiste ausblenden.

Auf der rechten Seite befindet sich (anders als bei Excel) eine Seitenleiste, welche weitere Schnellzugriff-Schaltflächen zu den Bereichen *Eigenschaften* (damit ist die Formatierung gemeint), *Formatvorlagen*, *Gallery* (Bilder- und Icon-Galerie), *Navigator* und *Funktionen* bereit stellt. Letzteres sind genau die Funktionen, die Du auf die Daten anwendet wirst und auf die ich im Kapitel **Built-in Funktionen** näher eingehen werde.

Alle folgenden Abbildungen in diesem Handbuch beziehen sich auf die GUI von Calc, wenn es unter dem MacOS Betriebssystem läuft. Also nicht wundern, falls es leichte Abweichungen zu Deiner Windows oder Linux Version geben könnte.

1.4 Aufbau eines Tabellenblatts

Wie gesagt, primär besteht ein Tabellenkalkulationsprogramm aus Tabellen. Eine Tabelle wiederum besteht aus Spalten und Zeilen. Den Schnittpunkt einer Spalte und einer Zeile bezeichnet man als Zelle. Jede Zelle hat einen eindeutigen Namen, eine **Zelladresse** sozusagen (Abb. 1.3).

Wenn Du dich in einer Zelle befindest, wird diese durch einen dicken schwarzen Rahmen hervorgehoben (siehe Abb. 1.4) und im Namenfeld wird der Name der Zelle angezeigt. So kannst Du leicht erkennen, welche Zelle aktiv ist und bearbeitet werden kann. Außerdem wird der Buchstabe der Spalte sowie die Zahl der Zelle blau hervorgehoben.

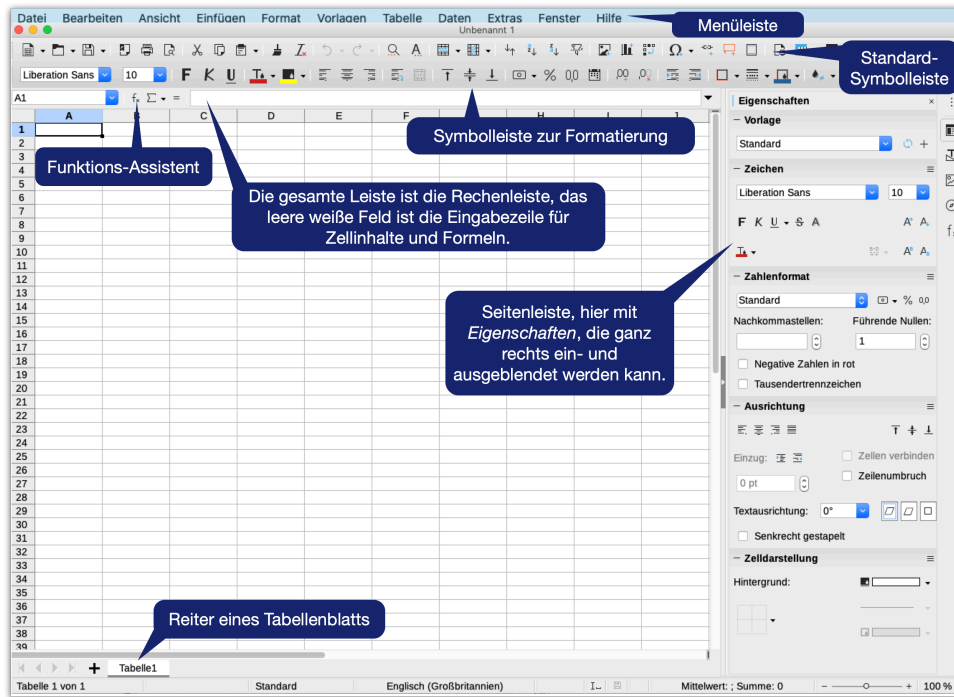


Abbildung 1.2: Übersicht der Hauptelemente der GUI.

1.5 Wie sollten Daten idealerweise strukturiert sein?

Es gibt einen gewissen Konsens, wie semi-strukturierte Daten in Tabellenkalkulationsprogrammen organisiert sein sollten, damit

- Daten ohne Probleme mit verschiedenen Programmen bzw. Programmiersprachen genutzt werden können,
- einfacher zu sortieren, filtern, aggregieren, visualisieren und modellieren sind und
- auch für den Betrachter schneller und leichter zu verstehen sind.

Diese Struktur wird in Abbildung Abb. 1.4 dargestellt. Variablen sollten grundsätzlich als Spalten angeordnet sein, Beobachtungs- oder Untersuchungseinheiten als Zeilen. Dieses Format nennt man auch das **lange** Format, welches im Kontrast steht zum **weiten** Format. Hierbei werden Variablen auf mehrere Spalten aufgeteilt. Auf unser Beispiel in Abb. 1.4 bezogen wäre das Format weit, wenn z.B. die Messwerte für Versuch A und B in getrennten Spalten sind. Dann wären die Zeilen 1 bis 11 in den Spalten A bis E gefüllt (Spalte D wäre Messwert-Versuch A und Spalte E Messwert-Versuch B).

Darüber hinaus sollte in der **ersten Zeile** jeder Spalte der **Variablennamen** angegeben sein. Wie Du die Variablen bezeichnest und ob Sonder- oder Leerzeichen verwendet werden ist innerhalb von Calc noch unwichtig. Spätestens wenn die Daten als CSV- oder TXT-Datei exportiert werden sollen, um diese anschließend in einem anderen Programm zu importieren, solltest Du eine bestimmte Namenskonvention einhalten. Aber dazu mehr im Kapitel **Vorbereitungen zum Export**.

Gut, aber wie kommt man nun von einem leeren Tabellenblatt zu dieser formatierten Tabelle ?

	A	B	C	D	E	F
1	ID	Versuch	Probenummer	Messwert		
2	1	A	1	10.25		
3	2	A	2	3.8		
4	3	A	3	4.9		
5	4	A	4	5.32		
6	5	A	5	9		
7	6	A	1	4.3		
8	7	A	2	2.8		
9	8	A	3	1.1		
10	9	A	Zelle	0.6	Zeilen	
11	10	A	5	19.4		
12	11	B	1	20.1		
13	12	B	2	17.5		
14	13	B	3	12.3		
15	14	B	4	14.8		
16	15	B	5	9.6		
17	16	B	1	8.2		
18	17	B	2	11.4		
19	18	B	3	13		
20	19	B	4	7.8		
21	20	B	5	8.9		
22			Spalten			
23						
24						

Abbildung 1.3: Zeilen- und Spaltenbezüge wie sie in allen Tabellenkalkulationsprogrammen gängig sind.

	A	B	C	D	E	F	G
1	ID	Versuch	Probenummer	Messwert			
2	1	A	1	10.25			
3	2	A	2	3.8			
4	3	A	3	4.9			
5	4	A	4	5.32			
6	5	A	5	9			
7	6	A	1	4.3			
8	7	A	2	2.8			
9	8	A	3	1.1			
10	9	A	4	0.6			
11	10	A	5	19.4			
12	11	B	1	20.1			
13	12	B	2	17.5			
14	13	B	3	12.3			
15	14	B	4	14.8			
16	15	B	5	9.6			
17	16	B	1	8.2			
18	17	B	2	11.4			
19	18	B	3	13			
20	19	B	4	7.8			
21	20	B	5	8.9			
22							
23							
24							
25							
26							
27							

Abbildung 1.4: Empfohlene Anordnung von Variablen im Tabellenblatt.

2 Tabellen anlegen

2.1 Direkte Dateneingabe und Datentypen

Wenn Du eine Zelle mit Inhalt füllen willst, geschieht das indirekt über die Eingabezeile in der Rechenleiste. Du kannst zwar eine Zelle anklicken, so dass das schwarze Viereck als Umrandung erscheint und dann einfach los schreiben, Du wirst aber sehen, dass das Geschriebene gleichzeitig auch in der Eingabezeile erscheint. Möchtest Du später Teile des Inhalts korrigieren, dann mach das am besten in der Eingabezeile direkt. Dafür markierst Du die Zelle, dessen Inhalt Du korrigieren willst und bewegst den Cursor in die Eingabezeile an die Stelle, die zu ändern ist. Insbesondere Formeln lassen sich leichter direkt in der Eingabezeile schreiben (dazu aber mehr in [Rechnen in Calc](#)).

Die Zelle eines Tabellenkalkulationsprogrammes kann unterschiedliche Datentypen aufnehmen. Neben Zahlen unterschiedlichsten Formats sind Texte (genauer Zeichenketten, im Englischen auch 'strings' genannt) sowie Datum und Uhrzeit die gängigsten Datentypen, die Calc erkennt. Du kannst allerdings auch Objekte wie Bilder, ClipArts oder Formen in eine Tabelle einfügen.

Um eine Zelle mit Daten zu füllen, markiert man sie, so dass der schwarze Rahmen um die Zelle erscheint, und tippt, was man darin haben will. Der Datentyp wird meist automatisch erkannt und entsprechend formatiert:

- **Text** wird immer linksbündig angeordnet aber nicht weiter formatiert.
- **Datum/Zeit** wird rechtsbündig angeordnet und in ein vordefiniertes Format umgewandelt. Wenn die Sprache Deutsch eingestellt ist, dann sind die Formate für
 - Datum: MM.DD.YY → gibt man z.B. in die Zelle '31.Dez.' ein, wird dieser Inhalt sofort in das Format '31.12.20' umgewandelt.
 - Uhrzeit: HH.MM.SS → gibt man '11:00' ein, wird der Eintrag zu '11:00:00' umgewandelt.
- **Zahlen** sind immer rechtsbündig angeordnet und führende Nullen werden abgeschnitten. Will man (zum Beispiel für Beschriftungszwecke) alle führenden Nullen behalten, tippt man vor der Eingabe der Zahl ein ' (ein Hochkomma; auf der Tastatur wäre das die Kombination <shift> und <#>). Dann wird die Zahl zwar mit der führenden Null angezeigt, der Datentyp wurde jedoch zu Text umgewandelt und erscheint linksbündig. Das bedeutet jetzt aber auch, dass man nicht mehr mit ihr rechnen kann.

2.1.1 Inhalte automatisch ausfüllen

Möchtest Du rasch Zellen mit ähnlichem Inhalt füllen, z. B. Zahlenreihen, Wochentage, Monate, Texte, Formeln und ähnliches, dann bietet sich die Option *Inhalt automatisch ausfüllen* an. Dabei gibt es 2 Arten: den Inhalt als logische Reihe ausfüllen oder mit EXAKT dem gleichen Inhalt wie eine andere Zelle füllen. Wie Du dabei vorgehen musst, siehst Du in den folgenden beiden Abschnitten:

2.1.1.1 Logische Reihen

Calc beherrscht einige logische Reihen. Will man beispielsweise Zahlen von 1 bis 20 oder Jahreszahlen von 1920 bis 2020 eingeben, markiert man eine Zelle, schreibt 1 hinein, markiert die Zelle darunter und gibt 2 ein. Nun markiert man beide Zellen und bewegt den Mauszeiger (Cursor) auf die schwarze Ecke rechts unten (Abb. 2.1). Der Cursor nimmt die Form eines schwarzen Kreuzes an. Nun klickt man mit der linken Maustaste darauf und zieht die Maus bis auf die gewünschte Endzahl herunter, ohne die Maustaste loszulassen. Voilà, eine logische Reihe entsteht (Abb. 2.2).

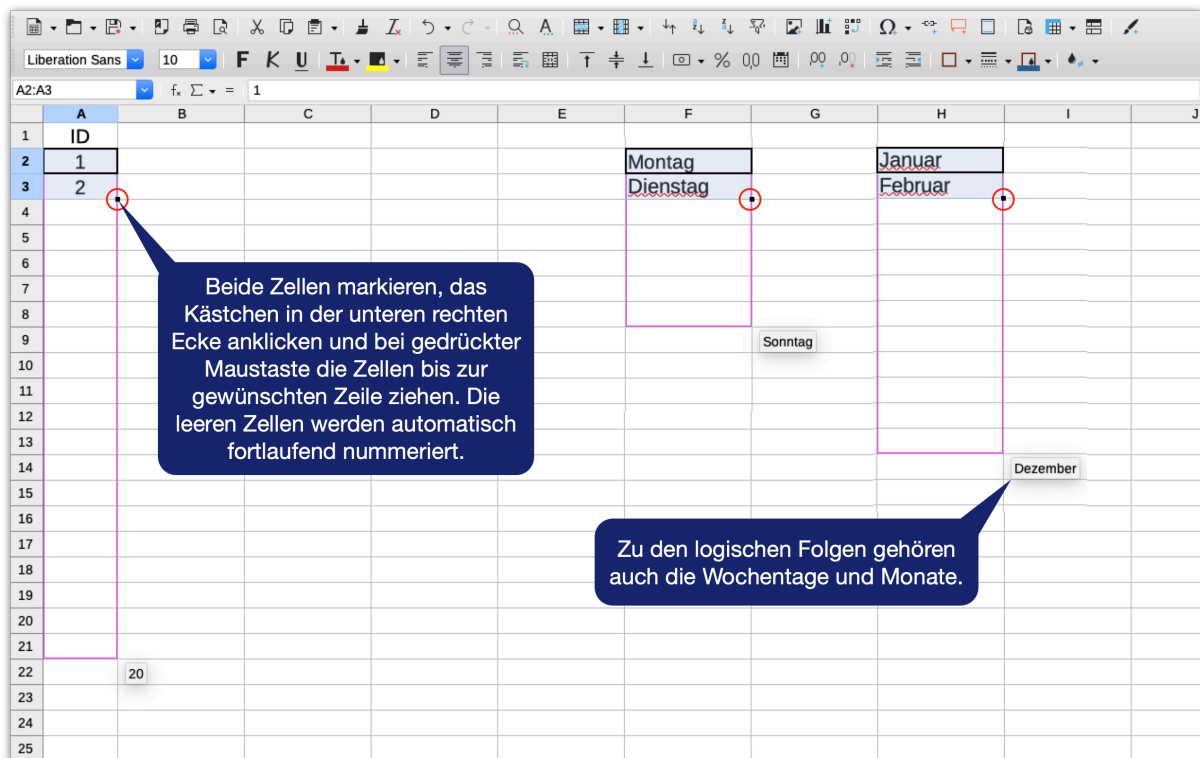


Abbildung 2.1: Eine logische Reihe entsteht.

Genauso lassen sich auch logische Reihen für Wochentage oder Monate erstellen. Wenn die deutsche Software Version von LibreOffice bei der Installation gewählt wurde (wird zumindest standardmäßig auf der deutschen LibreOffice Website angeboten), dann funktionieren logische Reihen bei Wochentagen und Monaten anderer Sprachen wie Englisch allerdings nicht.

2.1.1.2 Inhalte kopieren

Hierzu reicht es aus, die erste Zelle am 'Anfasser' (schwarze Ecke) anzuklicken und soweit nach unten zu ziehen, bis alle gewünschten Zellen mit dem gleichen Inhalt gefüllt sind. Wenn bereits die anliegenden Spalten gefüllt, beispielsweise Spalte A in Abb. 2.2, und Du möchtest die gleichen Zeilen füllen nur in Spalte B, dann kannst Du einen praktischen Kurzbefehl nutzen, indem Du zweimal auf den Anfasser klickst. Im Beispiel von Abb. 2.2 werden dadurch alle Zellen von B13 bis B21 mit 'B' gefüllt.

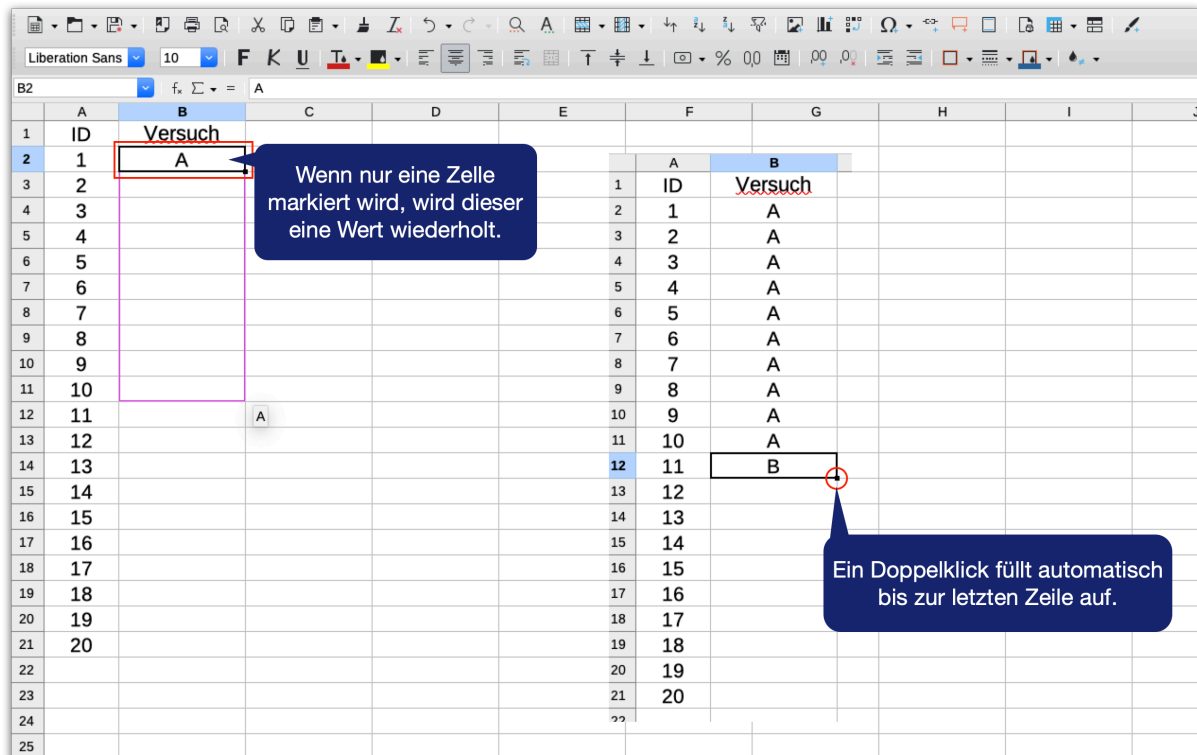


Abbildung 2.2: Automatisches Auffüllen

2.1.2 Daten in Calc transponieren

Falls man doch feststellt, dass die eingetragenen Daten nicht im langen sondern weiten Format vorliegen sollen (d.h. die Werte einer Variable werden auf einzelne Spalten und nicht Zeilen verteilt), dann ist die Option *Transponieren* sehr nützlich. Umgekehrt lassen sich damit natürlich auch Daten vom weiten Format ins lange umwandeln. Dazu werden alle Zellen markiert, die transponiert werden sollen, und mit der Tastenkombination <strg> und <c> bzw. <cmd> und <c> unter MacOS kopiert. Anschließend wird mit dem Cursor die Zelle aktiviert, von der aus die Inhalte nach rechts fortführend eingefügt werden sollen. Wichtig ist jetzt, nicht einfach den Inhalt mit der Tastenkombination <strg> und <v> bzw. <cmd> und <v> einzufügen, sondern, wie in Abb. 2.3 beschrieben, das Kontextmenü über die rechte Maustaste aufzurufen und *Inhalte einfügen* zweimal auszuwählen. Es wird sich anschließend ein Dialogfenster öffnen, wo Du auswählen kannst, wie der Inhalt eingefügt werden soll (ob als Text, Formel, mit Leerzeichen, etc.). Hier die Standardauswahl so lassen, außer bei *Optionen*, wo *Transponieren* ausgewählt werden muss!

2.2 Datenimport

Wenn Daten bereits digital vorliegen, müssen diese natürlich nicht in Calc neu eingetippt werden, sondern können importiert werden. Allerdings funktioniert das nicht mit jedem Dateiformat und -typ. Zur Erinnerung: das **Dateiformat** beschreibt, wie die Datei **interpretiert** werden soll, um die darin abgelegten Informationen visuell oder akustisch darstellen, bzw. wiedergeben zu können. Dazu zählen Text-, Bild-, Ton- und Video-Formate. Diese Formate lassen sich weiter in verschiedene Dateitypen unterscheiden. Der **Dateityp**, bzw. die Dateierendungen (Suffix) dient

2 Tabellen anlegen

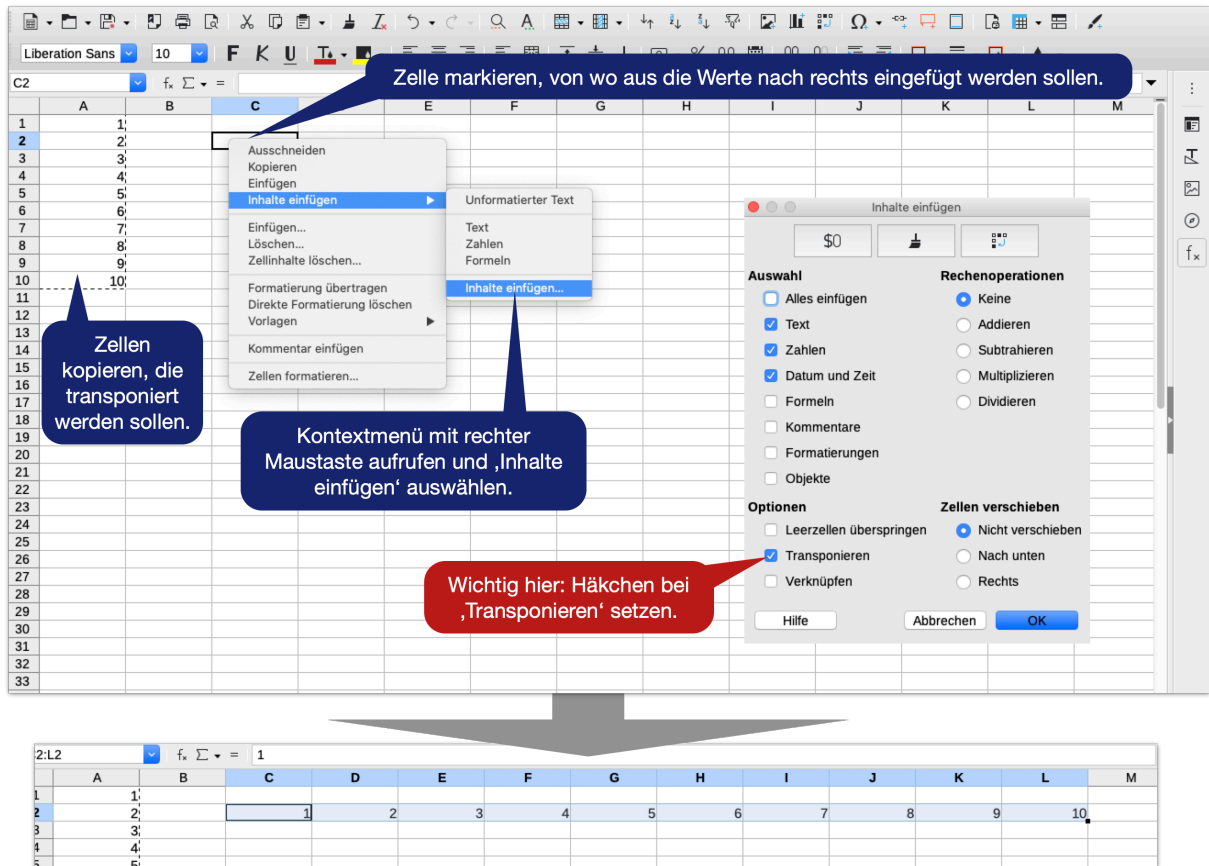


Abbildung 2.3: Zellen über 'Inhalte einfügen' transponieren.

dazu, dass Betriebssysteme eine Zuweisung treffen können, **mit welchem Programm** die Datei **geöffnet** werden soll bzw. kann. Das Dateiformat wird dann von der jeweiligen Anwendung gelesen und interpretiert. Zu den Bilddateien gehören beispielsweise die Dateitypen JPG, JPEG, BMP oder PNG, zu den Textdateien die Typen TXT, RTF und CSV.

Welche Dateien von Calc nun gelesen werden können, wird Dir beim Öffnen von Dateien aus Calc heraus in einer Dropdown-Liste angezeigt. Textdateien (meistens ASCII oder UTF-8 codiert) stellen dabei das gängigste und unkomplizierteste Format dar, welches von sehr vielen Programmen gelesen werden. Die verbreitetsten Typen sind hier die CSV-Datei (.csv), was für **Comma Separated Value** steht, und die TXT-Datei (.txt), bei dem Werte statt eines Kommas mit einem Tabulator getrennt sind. Beide Formate lassen sich aber nicht immer auf die gleiche Art in Calc öffnen.

2.2.1 CSV-Dateien direkt öffnen

CSV-Dateien lassen sich wie das 'hauseigene' Calc Format **ODS** auf 2 Wegen direkt öffnen: einmal aus dem Startcenter heraus, bevor Calc geöffnet wurde, und aus Calc direkt.

2.2.1.1 Öffnen einer CSV-Datei vom LibreOffice Startcenter

1. Wähle oben links *Datei öffnen* aus (siehe Abb. 2.4).
2. Wähle im erscheinenden Dialogfenster den entsprechenden Ordner aus, indem die Datei liegt.

3. Wähle als *Dateityp* ggf. 'Tabellendokumente' (damit sich wirklich Calc öffnet), klicke die entsprechende CSV Datei an und bestätige mit *Öffnen*.
4. Ein Hilfsfenster öffnet sich und eine Dateivorschau wird angezeigt (Abb. 2.5). Hier ist es wichtig folgendes zu prüfen:
 - Der **Zeichensatztyp** sollte '**Unicode (UTF-8)**' sein. Wenn ein anderer automatisch ausgewählt wurde, kann es passieren, dass die Daten als Sonderzeichen oder andere Sprachzeichen erscheinen.
 - Je nachdem in welchem Format **Dezimalzahlen** in der Datei vorliegen, ist es wichtig die entsprechende **Sprache** auszuwählen. Im vorliegenden Beispiel sind Dezimalzeichen mit einem Punkt angegeben, was dem englischen Format entspricht (und auch der Sprache des Computers!). Daher wird jetzt hier als Sprache 'Standard - Englisch (Großbritannien)' ausgewählt - das Land ist dabei zweitrangig.
 - Als Trennoptionen wird bei Calc automatisch der Tabulator, das Komma und das Semikolon als **Trennzeichen** ausgewählt. Das deckt somit das CSV Format ab (zur Erinnerung: CSV steht für Komma-separiert).
 - Wenn alles richtig eingestellt ist, sollte in der Datenvorschau die Anzahl der Spalten zum Datensatz passen.
5. Wenn alles soweit in Ordnung ist, bestätige den Import mit *OK*.

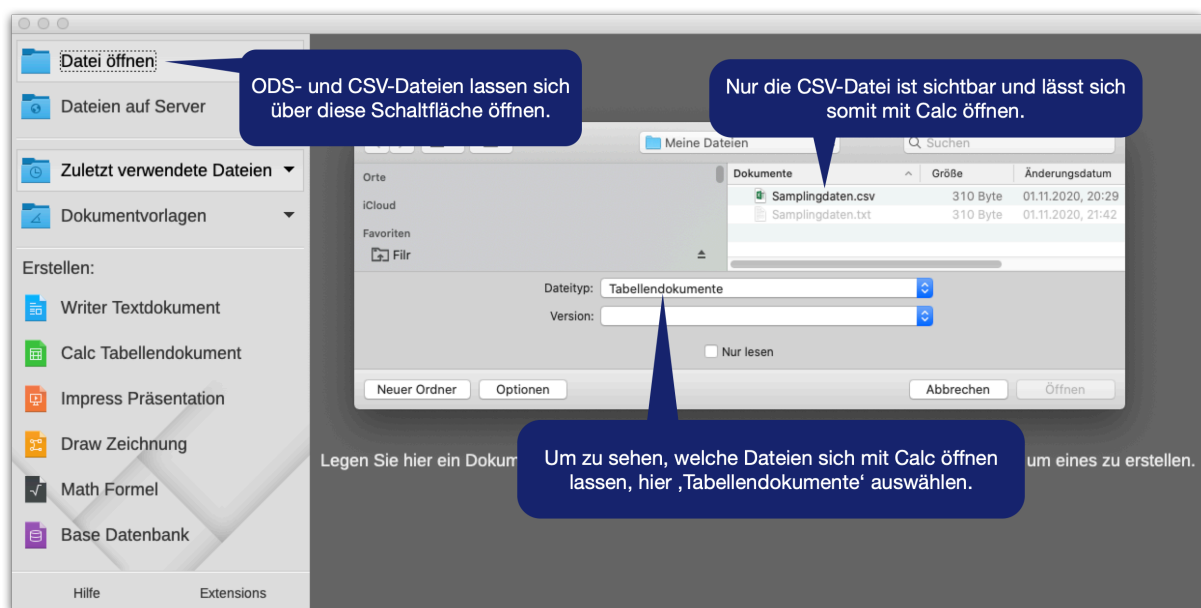


Abbildung 2.4: Öffnen von CSV-Dateien im Startcenter.

Auf die gleiche Art lassen sich im übrigen auch die Calc-spezifischen ODS-Dateien öffnen! Allerdings entfallen die Schritte 4 und 5.

2.2.1.2 Öffnen einer CSV-Datei aus Calc

1. Wenn Calc bereits geöffnet ist, gehe oben links im Hauptmenü auf *Datei* > *Öffnen...* oder wähle die Schaltfläche *Öffnen...* in der Standard-Symbolleiste (das ist das Icon mit dem Ordner).
2. Wähle im erscheinenden Dialogfenster den entsprechenden Ordner aus, indem die Datei liegt.

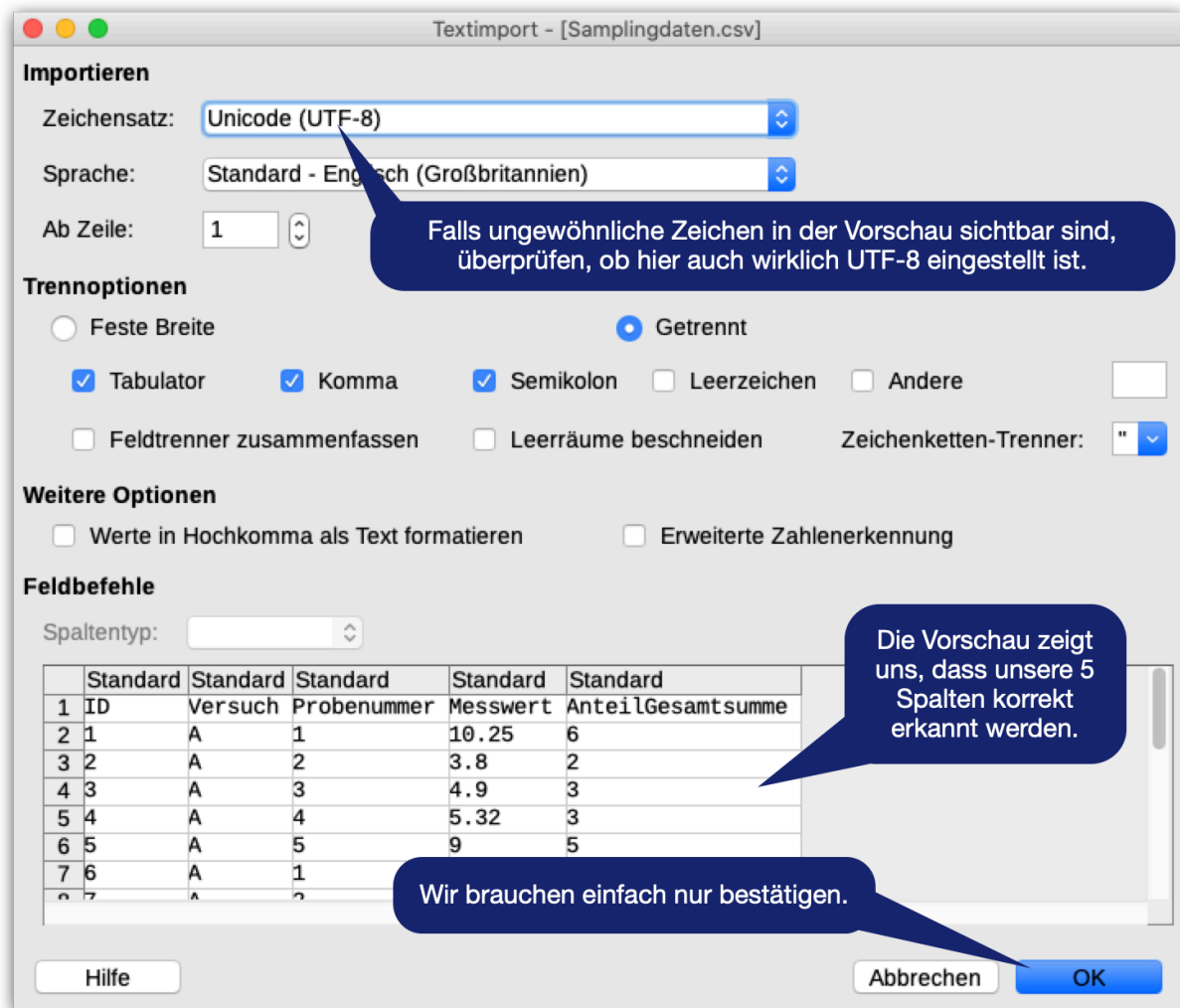


Abbildung 2.5: Einstellungen der Importoptionen zum korrekten Einlesen der Daten.

3. Falls die CSV Datei nicht sichtbar sein sollte, wähle ggf. unter *Dateityp* 'Alle Dateien'. Jetzt sollte die Datei sichtbar sein. Wähle sie aus und bestätige mit *Öffnen*.
4. Wie oben bereits beschrieben, öffnet sich ein Hilfsfenster mit der Datenvorschau. Führe wie oben beschrieben eine Überprüfung durch und bestätige anschl. mit *OK*.

2.2.2 TXT-Dateien und andere Dateitypen öffnen

Dateien, die nicht dem CSV- oder ODS-Typen entsprechen, lassen sich **nicht** über den **Startcenter** in Calc öffnen, da LibreOffice hier meist ein anderes seiner Programme standardmäßig auswählt. So öffnet sich beispielsweise bei einer TXT-Datei *Writer* statt Calc. Stattdessen müssen diese Dateien direkt aus Calc geöffnet oder als Tabelle importiert werden.

2.2.2.1 Direktes Öffnen aus Calc

Die Schritte zum Öffnen von TXT-Dateien und anderen Dateitypen sind die gleichen wie sie für CSV-Dateien oben beschrieben wurden (siehe **Öffnen einer CSV-Datei aus Calc**). Welche Dateitypen

sich alle öffnen lassen, kannst Du der Dropdown-Liste unter *Dateityp* in Schritt 3 entnehmen. Stattdessen kannst Du aber auch 'Alle Dateitypen' auswählen, dann sind alle Dateitypen die sich öffnen lassen sichtbar und auswählbar. In Schritt 4 kann es allerdings passieren, dass die Daten in der Vorschau nicht korrekt aussehen, was oft daran liegt, dass das Trennzeichen ein anderes ist. Mehr dazu aber im nächsten Abschnitt.

2.2.3 Import einer Tabelle als weiteres Tabellenblatt

Manchmal sollen externe Datensätze als weiteres Tabellenblatt in eine bestehende Calc Datei hinzugefügt werden. Dies geht vom Hauptmenü aus über *Tabelle > Tabelle aus Datei einfügen...* (siehe Abb. 2.6). Es öffnet sich wieder das Dialogfenster, in dem die zu importierende Datei ausgewählt werden kann. Auch hier wieder ggf. bei *Dateityp* 'Alle Dateien' auswählen. Wie beim direkten Öffnen einer Datei, muss im darauf folgenden Hilfsfenster überprüft werden, ob Zeichensatz und Sprache korrekt sind, und ob die Trennoptionen so gewählt sind, dass die Variablen im Datensatz auch wirklich als getrennte Spalten eingelesen werden (Abb. 2.7).

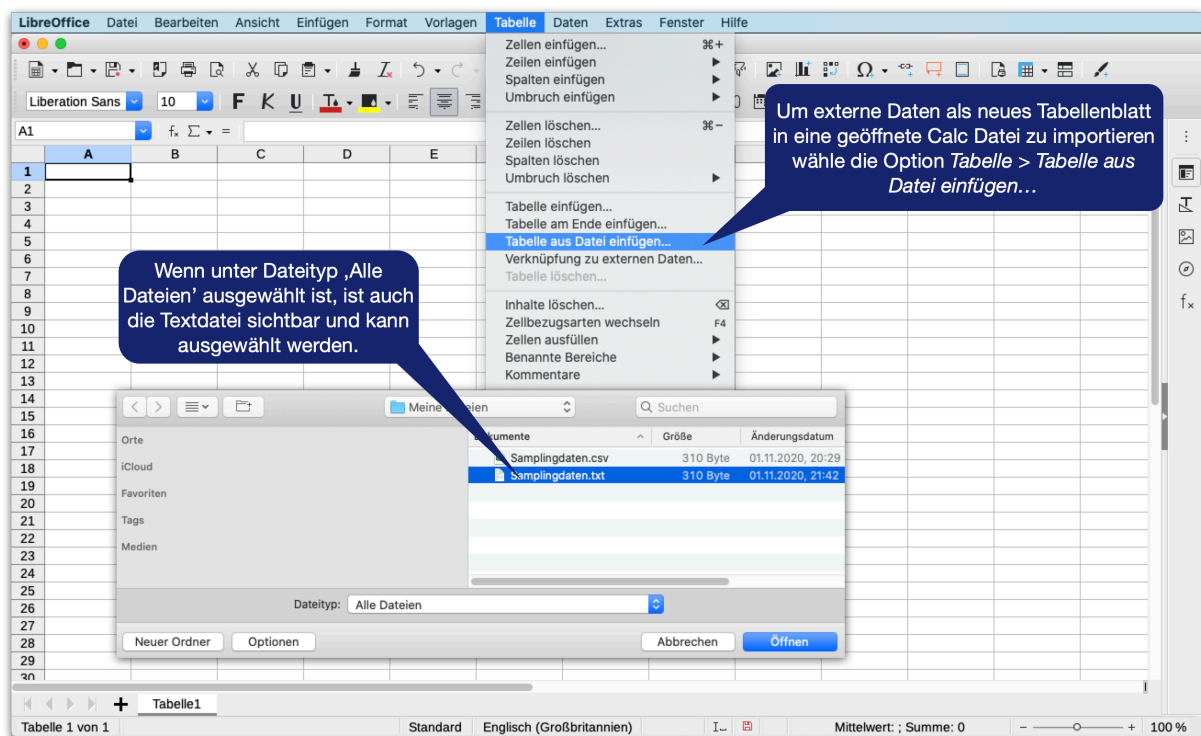


Abbildung 2.6: Hinzufügen einer TXT-Datei in eine geöffnete Calc Datei durch den indirekten Import.

Wenn als Trennzeichen kein Tabulator, Komma, Semikolon oder Leerzeichen verwendet wird, muss das **Trennzeichen explizit angegeben** werden. Im Beispiel in Abb. 2.7 wurde ein Schrägstrich ('forward slash') verwendet. Erst wenn dieses Zeichen explizit angegeben wird, kann Calc die Variablen in ihre jeweiligen Spalten trennen.

Anschließend kann festgelegt werden, ob die Tabelle vor das aktuelle Tabellenblatt eingefügt werden soll oder danach.

Ein **kleiner Tipp**: Das Trennzeichen lässt sich am besten aus der **1. Zeile** ablesen, welche gewöhnlich die Variablennamen enthält. Da die Namen selten Sonderzeichen enthalten, kann hier leichter ein sich regelmäßig wiederholendes Sonderzeichen (wie hier der Schrägstrich) erkannt wer-

2 Tabellen anlegen

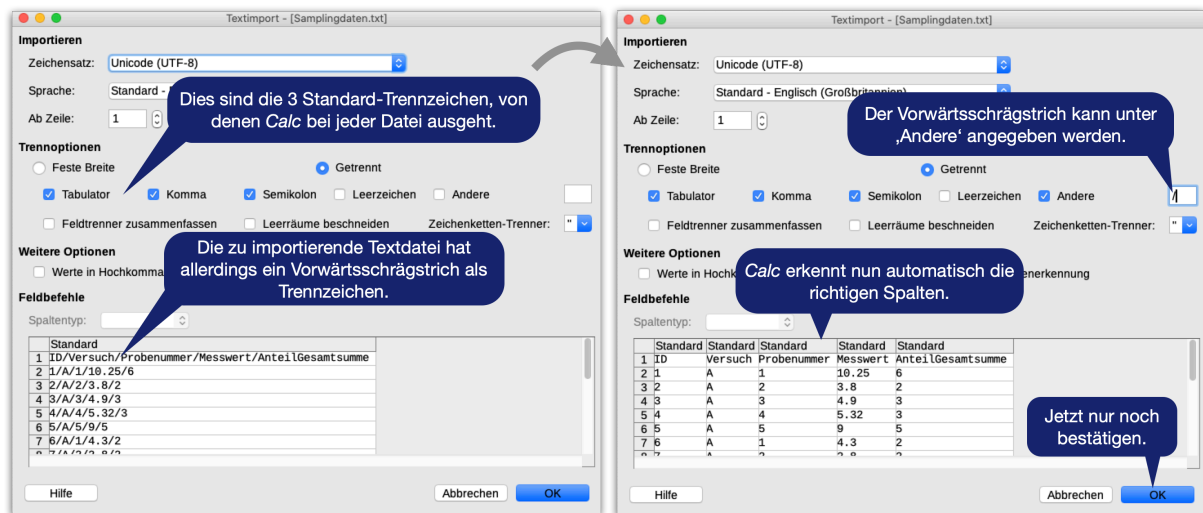


Abbildung 2.7: Angabe des Trennzeichens zum korrekten Import.

den. In den anderen Zeilen befinden sich die Elemente oder Werte der Variablen und die können zum Beispiel Zahlen enthalten, die (nach deutscher Schreibweise) Kommas als Dezimaltrennzeichen verwenden. Hier kann es schnell passieren, dass man diese Dezimaltrennzeichen als Spaltentrennzeichen interpretiert.

2.3 Dateispeicherung nicht vergessen

Dateien sollten sofort nach dem Öffnen oder Importieren abgespeichert werden, um Datenverlust zu vermeiden. Hierzu entweder die Tastenkombination <strg> (bzw. <cmd>) und <shift> und <s> nutzen, im Hauptmenu *Datei* > *Speichern unter...* auswählen oder die Schnellzugriff-Schaltfläche *Speichern* (erkennbar am Disketten Icon) nutzen. Im sich öffnenden Dialogfenster den Speicherort auswählen und einen Dateinamen angeben. Als Dateityp sollte der Calc-spezifische ODS Typ genommen werden (ODS steht für **O**pen**D**ocument **S**preadsheet) (Abb. 2.8).

2.4 Tabellen übersichtlicher gestalten

2.4.1 Spaltenbreite und Zeilenhöhe ändern

Manchmal kommt es vor, dass der Inhalt der Zelle größer ist als die eigentliche Zelle. Dann kannst Du aber problemlos die Breite dieser Spalte oder die Höhe dieser Zeile ändern, eine einzelne Zelle selbst lässt sich nicht ändern. Du kannst zwischen 3 Varianten wählen:

1. Bewegst Du die Maus zwischen zwei Spalten (oder zwei Zeilen) ändert sich der Cursor in einen Doppelpfeil (siehe Abb. 2.9). Nun kannst Du durch Ziehen die Breite bzw. die Höhe ändern.
2. Klickst Du doppelt mit der Maus zwischen zwei Spalten bzw. zwischen zwei Zeilen, so wie in 1. beschrieben, dann wird die Spaltenbreite bzw. die Zeilenhöhe automatisch an die Breite des Inhalts angepasst. → Tipp: Wenn Du das ganze Tabellenblatt markierst, indem Du in das graue Feld links von Spalte A mit der Maus einmal klickst, und dann zwischen Spalte

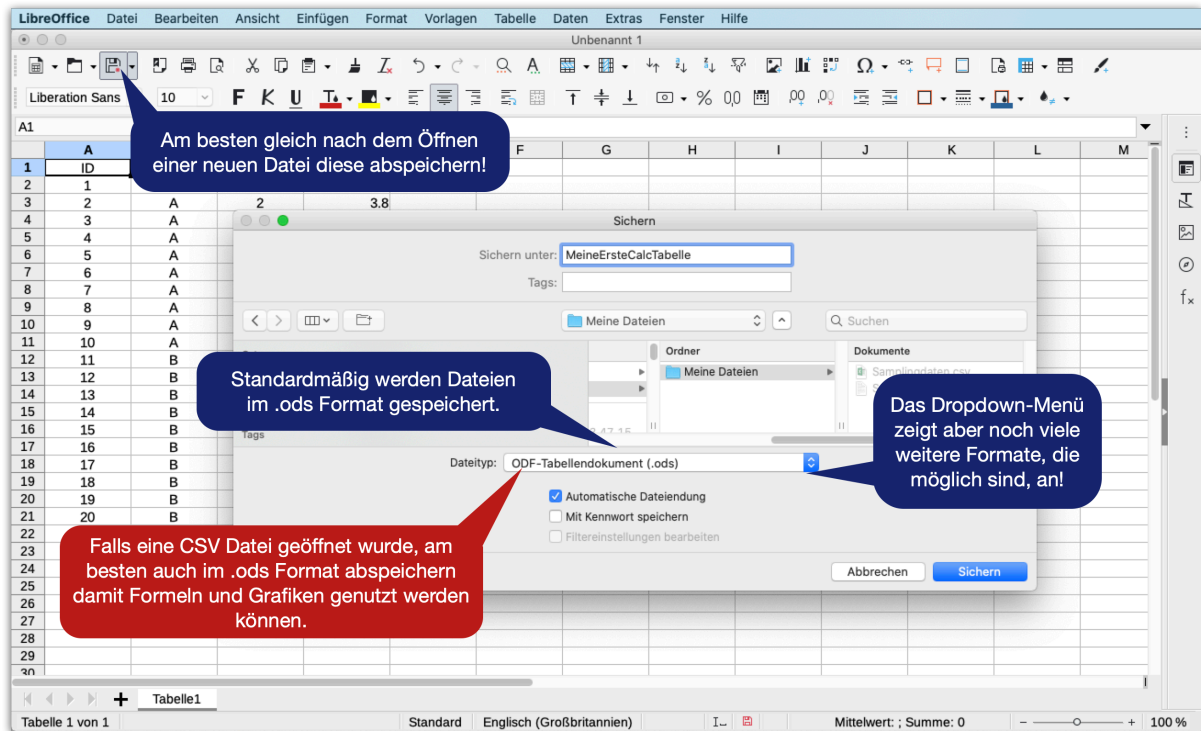


Abbildung 2.8: Speichern der Datei im ODS-Standardformat.

A und B (bzw. Zeile 1 und 2) doppelt klickst, wird die Breite aller Spalten (bzw. Höhe aller Zeilen) gleichzeitig an den Inhalt angepasst!

3. Klickst Du mit der rechten Maustaste auf eine Spalten- oder Zeilenüberschrift (den Buchstaben oder die Zahl), kannst Du im nun erscheinenden Kontextmenü den Eintrag *Spaltenbreite* bzw. *Zeilenhöhe* auswählen und im darauf folgenden Eingabefenster eine spezifischen Wert eintragen. Gib dann in das Eingabefenster den gewünschten Wert ein.

Bei gedrückter <shift> Taste kannst Du auch mehrere Spalten oder Zeilen, die direkt neben- oder untereinander liegen, auswählen. Bei gedrückter <strg> bzw. <cmd> Taste lassen sich Spalten und Zeilen auswählen, die nicht direkt neben- oder untereinander liegen.

2.4.2 Textausrichtung

Calc richtet Text automatisch linksbündig aus, Datum/Zeit und Zahlen rechtsbündig. Manchmal kann diese Ausrichtung aber ungünstig sein, wenn z.B. eine Textspalte rechts von einer Zahlenspalte liegt wie in Abb. 2.10. Hier sieht es aus, als wären Zahlen und Buchstaben zusammengehörig. Zur Änderung der Ausrichtung nutze die 3 Schaltflächen in der Formatierungs-Symbolleiste.

2.4.3 Zeilen und Spalten fixieren

Wenn Zeilen und Spalten über den sichtbaren Bereich des Tabellenblatts hinaus gehen, kann das Blatt vertikal wie horizontal gescrollt werden. Das wiederum führt dazu, dass die Spaltennamen oder Zeilenamen, IDs, etc. nicht mehr sichtbar sind. Um dies zu verhindern, gibt es die Möglichkeit, die ersten Spalten und Zeilen über die Schaltfläche *Zeilen/Spalten fixieren*

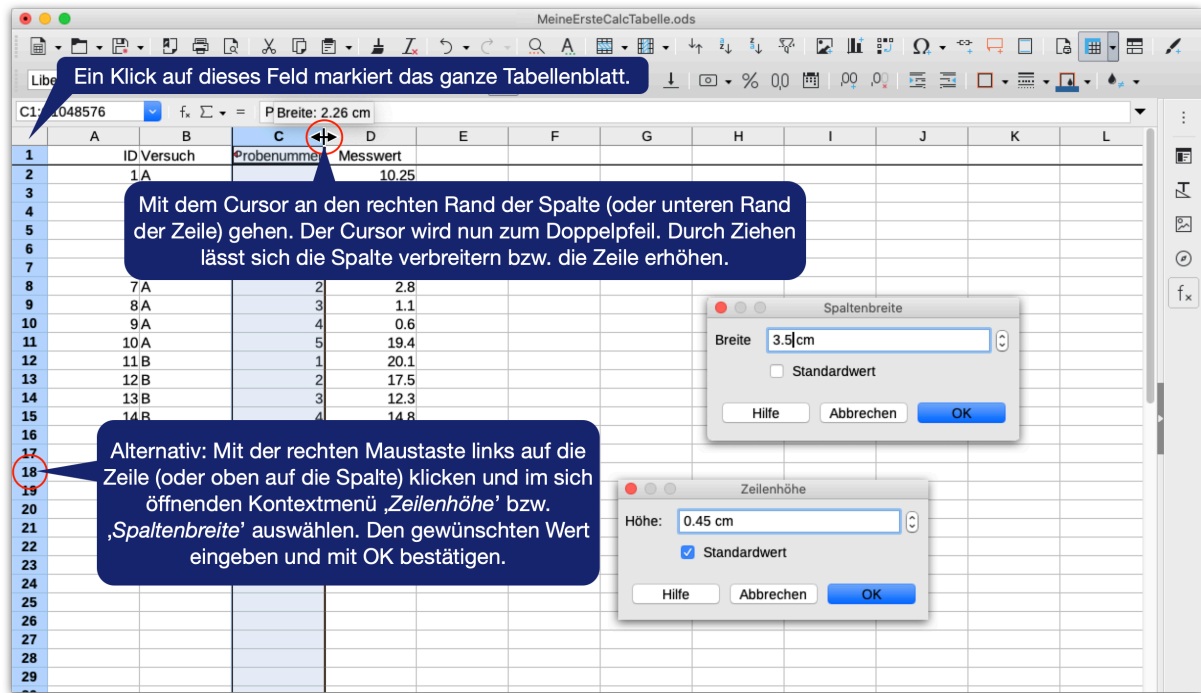


Abbildung 2.9: Veränderung der Spaltenbreite und Zeilenhöhe über das Kontextmenü oder einen Kurzbe-
fehl.

einzufrieren bzw. zu fixieren (Abb. 2.11).

2.4.4 Daten filtern und sortieren

Bei größeren Datensätzen kann es sehr sinnvoll sein, diesen nach verschiedenen Spalten zu sortieren, um sich einen besseren Überblick zu verschaffen. Dazu gibt es drei Schaltflächen in der Standard-Symbolleiste (s. Abb. 2.12), wobei die linke davon mehr Optionen bietet: hier kann genau definiert werden, nach welchen Spalten sortiert werden soll (es können bis zu drei Spalten ausgewählt werden) und was die Sortieroptionen sein sollen. Bevor eine der drei Schaltflächen gedrückt wird, sollte allerdings der zu sortierende Datenbereich markiert werden. Es ist dabei möglich, nur ein paar zusammenhängende Zellen innerhalb des gesamten Datensatzes zu markieren und diese zu sortieren. Allerdings wird es eher vorkommen, dass Du den ganzen Datensatz nach bestimmten Spalten sortieren willst. Dann musst Du alle Spalten oder das gesamte Tabellenblatt über die linke obere Ecke markieren.

Das Filtern der Daten, genauer der Zeilen, funktioniert über die Schaltfläche *Autofilter* direkt daneben (Abb. 2.12). Dazu muss aber erst die oberste Spalte mit den Spaltennamen markiert werden. Nach Auswahl der Schaltfläche erscheint neben jedem Spaltennamen eine Dropdown-Liste. Hierüber können für die entsprechende Spalte

- einzelne Zellen ausgewählt bzw. ausgeschlossen (fehlendes Häkchen) werden,
- die obersten 10, leere oder nicht leere Zellen ausgewählt werden oder
- über den *Standardfilter* Operatoren genutzt werden, um z.B. nur Werte zwischen 3 und 8 anzeigen zu lassen.

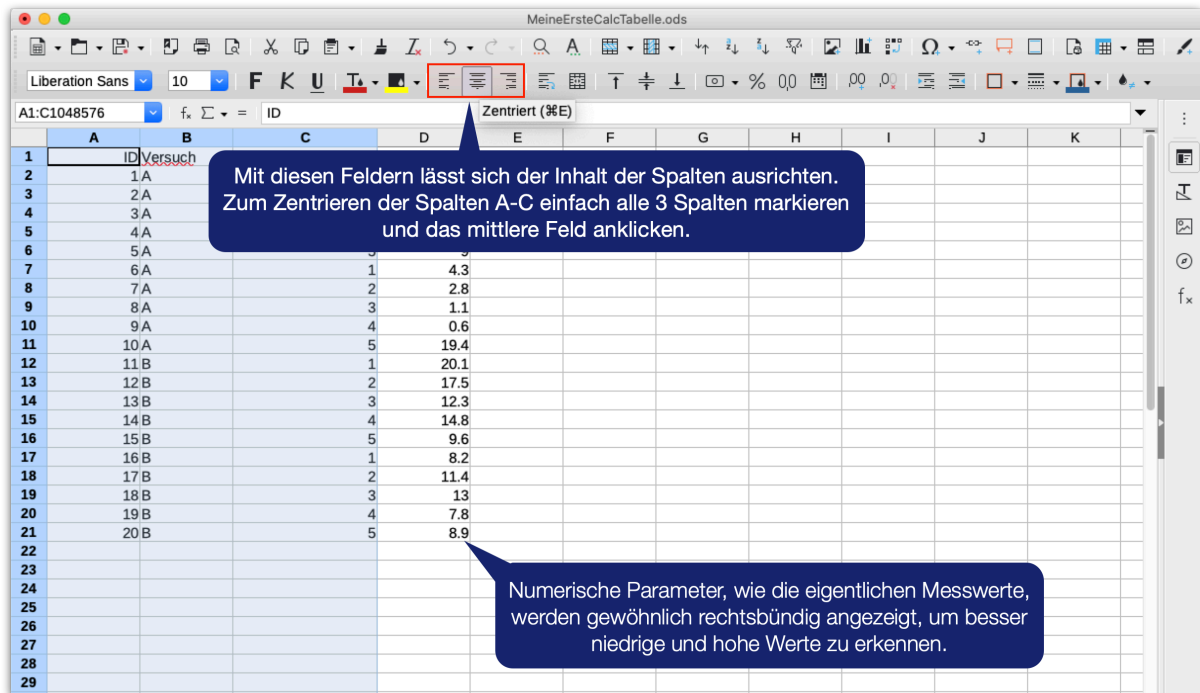


Abbildung 2.10: Inhalte von Spalten lassen sich zentrieren oder links- bzw. rechtsbündig ausrichten.

2.4.5 Zellen formatieren - Komma vs. Punkt

Es gibt eine Reihe von Möglichkeiten, Zellen zu formatieren, z. B. über die Symbolleiste zur Formatierung. Ein besonders einfacher Weg, auf alle Formatierungsmöglichkeiten zugreifen zu können, ist das Zellen-spezifische Kontextmenü. Dazu einfach die entsprechende Zelle auswählen, über die rechte Maustaste das Kontextmenü aufrufen und hier die Option *Zellen formatieren* anklicken (Abb. 2.13).

Wie im Abschnitt **Direkte Dateneingabe und Datentypen** schon erwähnt, unterscheidet Calc zwischen Zahlen mit denen es rechnen kann und Text bzw. strings. Nun werden Nachkommaziffern von den ganzen Zahlen durch den sogenannten Dezimaltrenner getrennt. Im deutschen Sprachraum ist dies häufig ein Komma. im englischen Sprachraum und vor allem in der Programmierung ist es aber stets ein Punkt. Kommata werden im Englischen meist als Tausendertrennzeichen verwendet, während im Deutschen hier jetzt ein Punkt verwendet wird. Calc kann allerdings mit beidem umgehen, genau wie die Betriebssysteme. Programmiersprachen wie R können das aber nicht. Daher ist meine Empfehlung, nicht nur Calc sondern Dein **ganzes Betriebssystem** anzupassen und die Darstellung von **Dezimalzahlen** auf die **englische Schreibweise** einzustellen.

Zur Änderung der Ländereinstellungen bei Windows 10, klicke auf die Schaltfläche *Start* und dann auf *Systemsteuerung*. Wähle *Datums-, Uhrzeit- oder Zahlenformate ändern*. Das Dialogfeld *Region* wird angezeigt. Wähle anschließend auf der Registerkarte *Formate* in der Dropdown-Liste das Format aus, das Du ändern möchtest. Klicke nun auf die Schaltfläche *zusätzliche Einstellungen*. Klicke im darauf folgenden Dialogfeld *Format anpassen* auf die Registerkarte *Zahlen* und wähle hier den Punkt als Dezimaltrennzeichen.

Wenn Du nun beim Eintippen der Zahlen ein Komma verwenden, der Dezimaltrenner aber auf Punkt eingestellt ist, erkennt Calc die Zahl nicht als Zahl, sondern als Text. Dies ist daran zu erkennen, dass die Zahl dann links- statt rechtsbündig in der Zelle angezeigt wird (was man aber auch manuell ändern kann und deswegen nicht auffallen muss). Darum der dringende Rat: Nutze zur

2 Tabellen anlegen

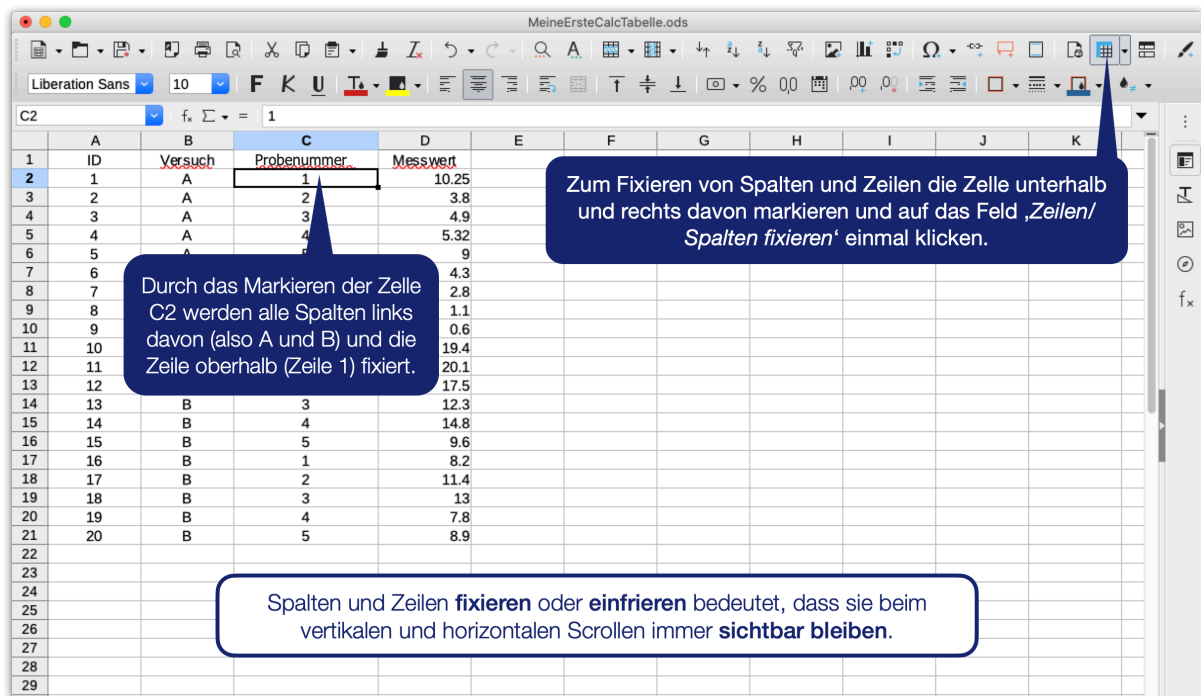


Abbildung 2.11: Zusammenhängende Zeilen und Spalten dauerhaft sichtbar machen.

Eingabe von Zahlen den „Zehnerblock“ auf der rechten Seite der Tastatur. Die Kommataste, die dort neben der 0 sitzt, fügt in Calc IMMER den Dezimaltrenner ein.

2.4.6 Umbenennen von Tabellenblättern

Für mehr Übersicht sorgen auch gut beschriftete Arbeitsblätter. Mit der Auswahl *Tabelle umbenennen* im Kontextmenü (über die rechte Maustaste) oder einem doppelten Maustastenklick auf den Karteireiter Deiner Tabelle oder Deines Diagramm-Arbeitsblattes kannst Du Dir einen eindeutigeren Namen aussuchen (Abb. 2.14).

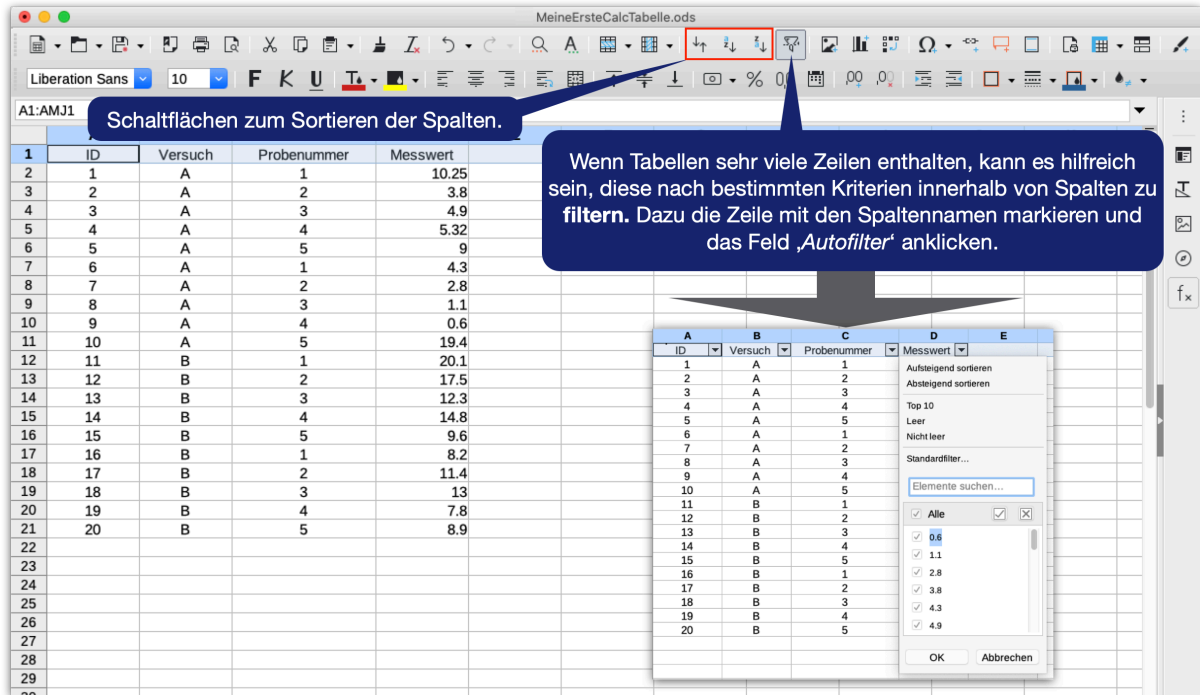


Abbildung 2.12: Die Filterfunktion hilft, Teildatensätze zu untersuchen.

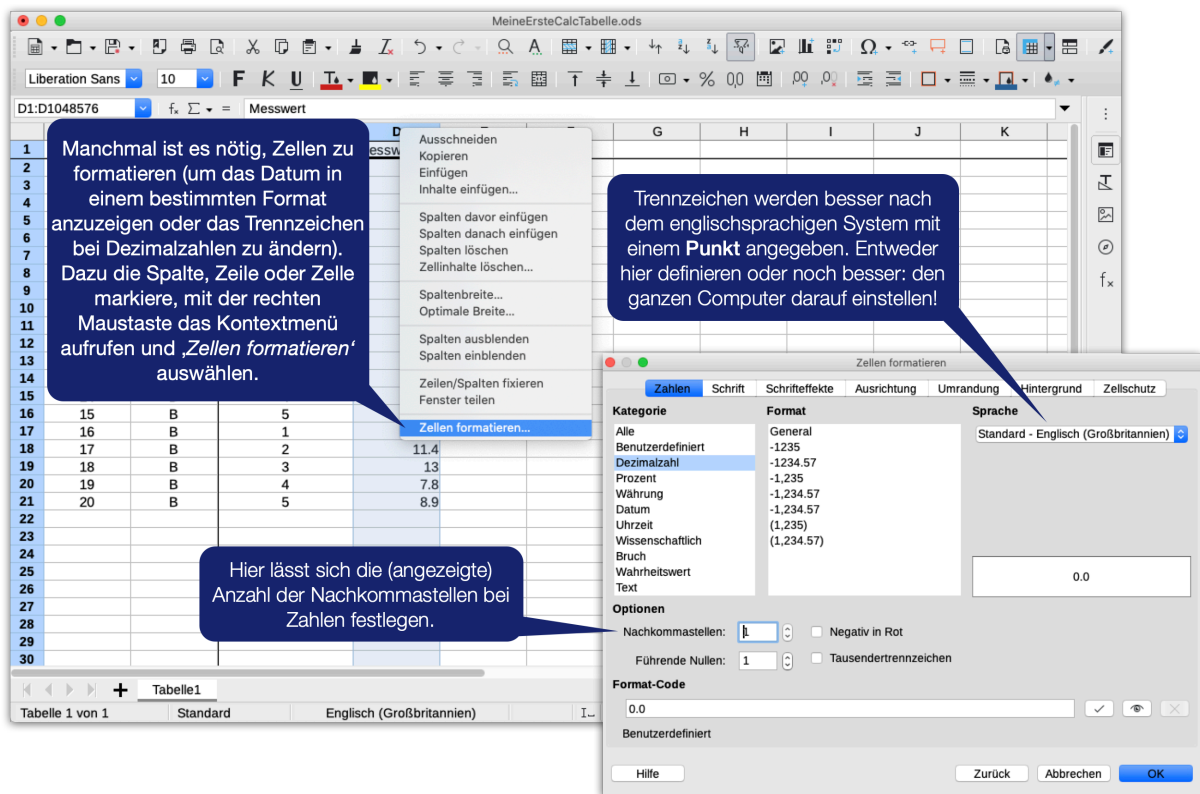


Abbildung 2.13: Umfangreiche Möglichkeiten, Zellen über das Kontextmenü zu formatieren.

2 Tabellen anlegen

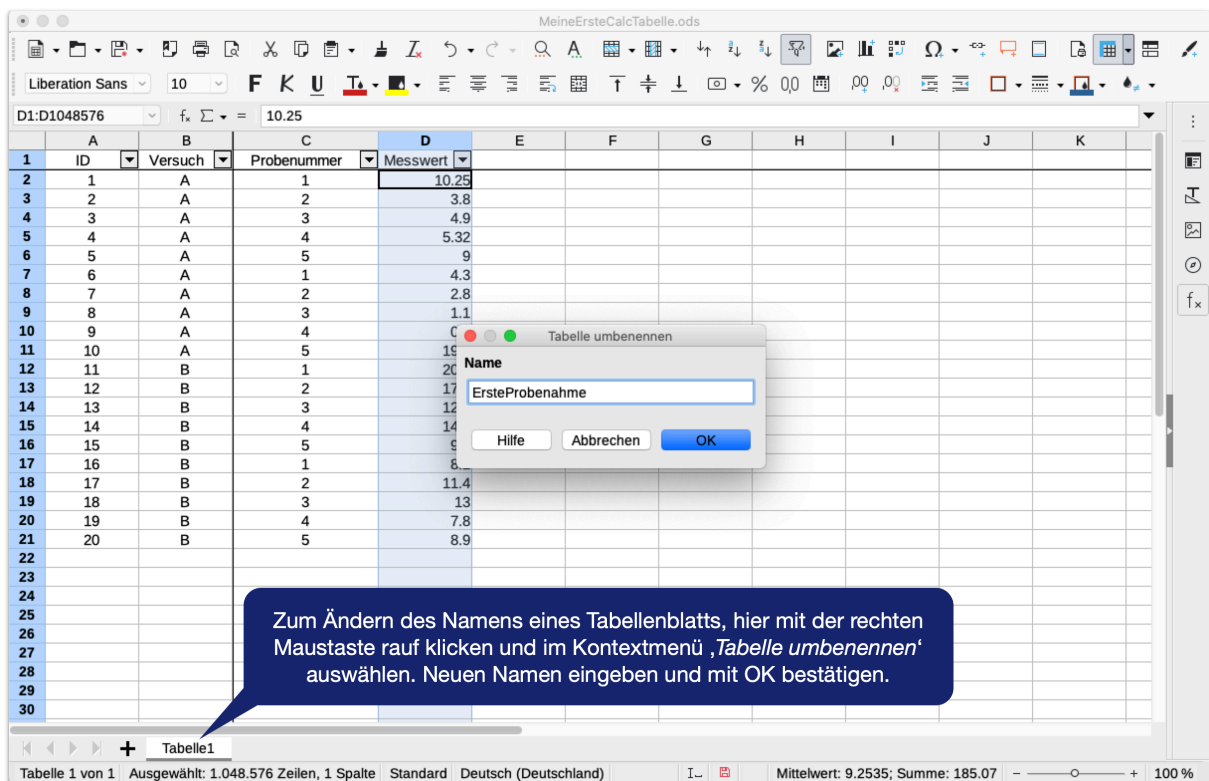


Abbildung 2.14: Umbenennung von Tabellenblättern.

3 Rechnen in Calc

3.1 Wie geht das?

Hauptsächlich wird ein Tabellenkalkulationsprogramm zum Rechnen verwendet. Damit das Programm aber weiß, dass Du jetzt keinen Inhalt in eine Zelle einfügst, sondern rechnen möchtest, musst Du es dem Programm mitteilen.

Das geschieht durch die Eingabe eines '='. Jetzt weiß das Programm, dass Du rechnen möchtest. Nun kannst Du deine Rechnung eingeben. Wie beim Taschenrechner auch, kannst Du alle arithmetischen Operatoren (+, -, *, /, ^ zur Addition, Subtraktion, Multiplikation, Division und Potenzierung) sowie Klammern verwenden (s. Abb. 3.1):

Beispielformel: =1+5*10/(20-4)

In dieser Formel beziehen wir uns auf Werte die in den Zellen A2-E2 in Abb. 3.1 enthalten sind. Nachdem Du die Formel in die Zelle eingegeben hast, drückst Du die Eingabetaste (<enter> Taste). Mit der Eingabetaste verlässt Du die Zelle und lässt sie berechnen. Fehlerhafte Formeln werden von Calc mitgeteilt.

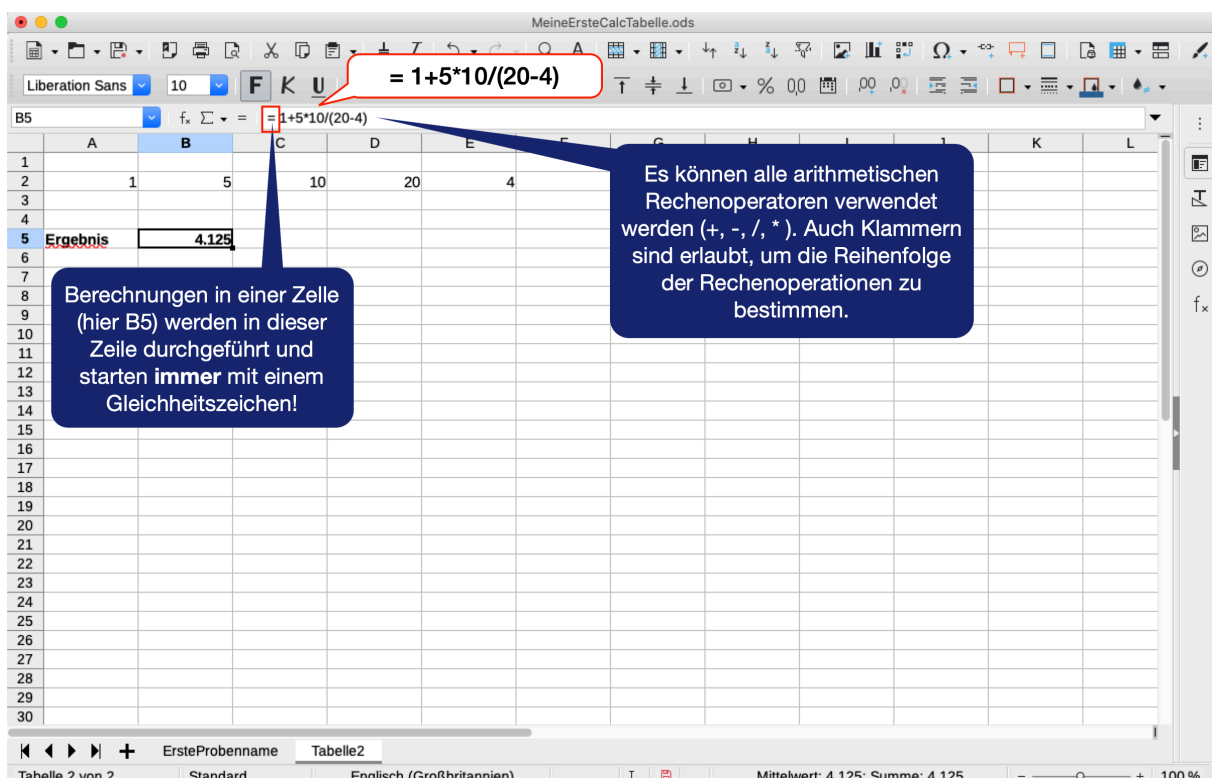


Abbildung 3.1: Einfache Berechnungen mit Zahlen und arithmetischen Operatoren.

Wie Du in der Abbildung Abb. 3.1 erkennen kannst, wird in der ausgewählten Zelle (hier B5) nun das Ergebnis der Rechnung angezeigt. Die Formel ist aber nicht verschwunden, Du siehst die Formel

nach wie vor in der Eingabezeile und kannst sie dort auch ändern.

Würdest Du diese Zelle jetzt kopieren, hättest Du 2 Möglichkeiten: entweder kopierst Du die Formel, welche in der Eingabezeile sichtbar ist (und auch den wahren Inhalt darstellt), oder das Ergebnis was in der Zelle sichtbar ist. Letzteres geht allerdings nur über das Kontextmenü und die Option *Inhalte kopieren*, welche Du bereits im Kapitel **Daten in Calc transponieren** kennengelernt hast. Dazu noch mehr in Kapitel **Export von CSV Dateien**.

3.2 Warum nicht mit Zahlen rechnen?

Natürlich kannst Du einfach mit den Werten aus Zellen direkt rechnen wie in Abb. 3.1. Eine andere Möglichkeit wäre, statt der Werte, also der Zellinhalte, die **Zellnamen** zu verwenden. Calc wird dann automatisch die Inhalte raussuchen und mit diesen die Rechenoperation durchführen. Die Beispielformel würde dann lauten:

$$=A2+B2*C2/(D2-E2)$$

Das Ergebnis wäre in beiden Fällen zunächst natürlich dasselbe. Wenn Du nun aber statt den Zahlen 5 und 10 die Zahlen 500 und 100 in die Zellen B2 und C2 eingibst, wäre das Ergebnis im ersten Fall nach wie vor 4.125 und nicht wie zu erwarten wäre 3126. Dadurch, dass Du stattdessen mit den Namen oder 'Adressen' der Zellen rechnest, wird das **Ergebnis automatisch geändert** (s. Abb. 3.2):

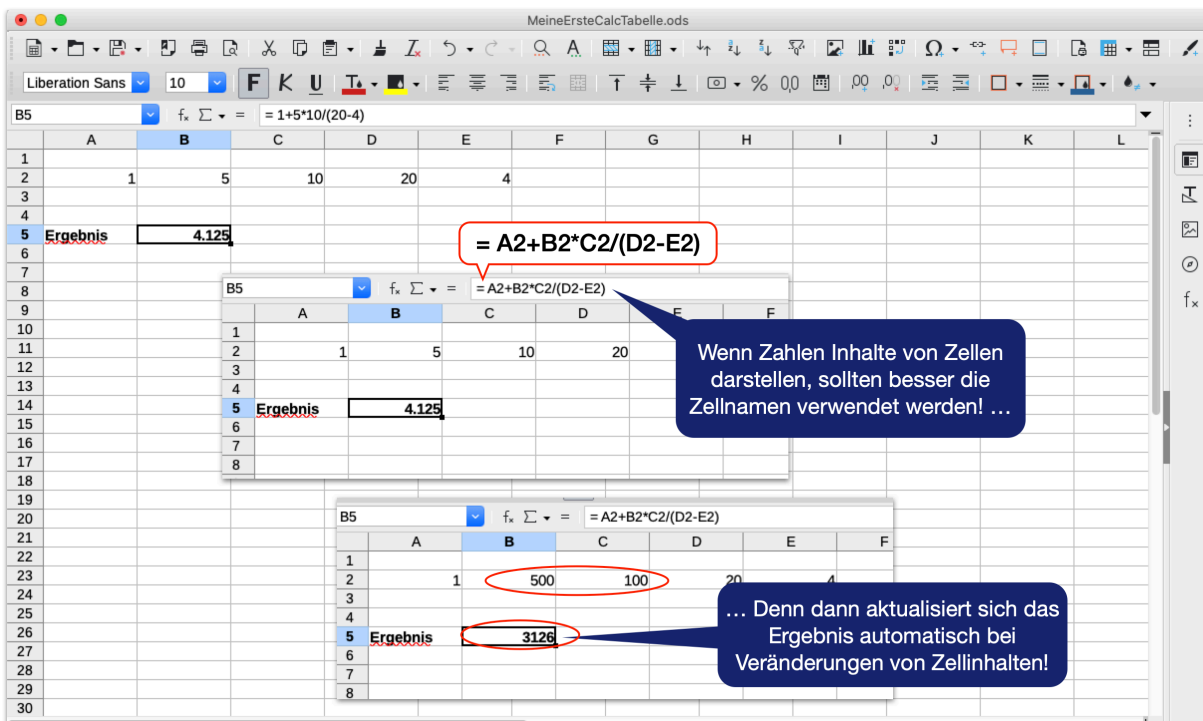


Abbildung 3.2: Verwendung von Zellnamen statt Zahlen in Rechenoperationen.

3.3 Eingeben von Funktionen

Auch eine Funktion fängt in Calc immer mit einem '=' an. Danach folgt der Name der Funktion und **runde Klammern**, innerhalb dessen dann Verweise auf eine oder mehrere Zellen stehen und evtl. **Parameter**.

Häufige Funktionen sind SUMME(), MITTELWERT() und ANZAHL() sowie MIN() und MAX() für den Minimal- und Maximalwert.

Grundsätzlich werden in Calc Funktionen **groß geschrieben**. Allerdings ist Calc sehr benutzerfreundlich und wird Deine Formel automatisch korrigieren, falls Du die Funktion mal klein schreiben solltest. Die Programmiersprache R ist da viel strenger und lässt keinen Fehler durchgehen!

3.3.1 Built-in Funktionen

Calc stellt eine Vielzahl von sog. 'built-in' Funktionen in den Bereichen

- Datenbank
- Datum und Zeit
- Finanzen
- Information
- Logische Operationen
- Mathematik
- Matrix
- Statistik
- Tabelle
- Text
- Plug-in

zur Verfügung. Darüber hinaus kannst Du über 'Add-ins' weitere Funktionen hinzufügen (siehe auch **Aktivieren von Add-Ins**).

Eine Übersicht aller Funktionen erhältst Du über die Seitenleiste. Klicke hier auf die Schaltfläche f_x ganz rechts und es wird die Seitenleiste für Funktionen eingeblendet.

Es gibt nun mehrere Möglichkeiten, diese Funktionen zu nutzen:

1. Direkte Eingabe - wenn der Name der Funktion bekannt ist, kann man diesen direkt in die Eingabezeile schreiben.
2. Verwendung des Funktions-Assistenten in der Befehlszeile. Hier hast Du die Möglichkeit, gezielt nach Funktionen zu suchen.
3. Verwendung des Funktions-Helfer in der Seitenleiste.

In Abbildung Abb. 3.3 siehst Du, wie die Summen-Funktion mittels Funktions-Assistent in der Seitenleiste verwendet wird. Gehe dabei in folgenden Schritten vor:

1. Klicke mit dem Cursor die Zelle an, in die die Formel eingetragen werden soll. In diesem Beispiel soll in Zelle D24 die Summe aus den Zellen D2 bis D21 gebildet werden.
2. Suche und markiere in der Seitenleiste die Funktion, die eine Summe aus Zellen bildet. Wenn eine Funktion markiert ist, erscheint auch ganz unten in der Seitenleiste die Information was die Funktion macht und wie sie genutzt werden soll.
3. Nun klicke die Schaltfläche *Funktion in Tabelle einfügen* in der Seitenleiste an. Es erscheint anschl. in der Zelle wie auch in der Eingabezeile der Befehl '=SUMME(Zahl)'.

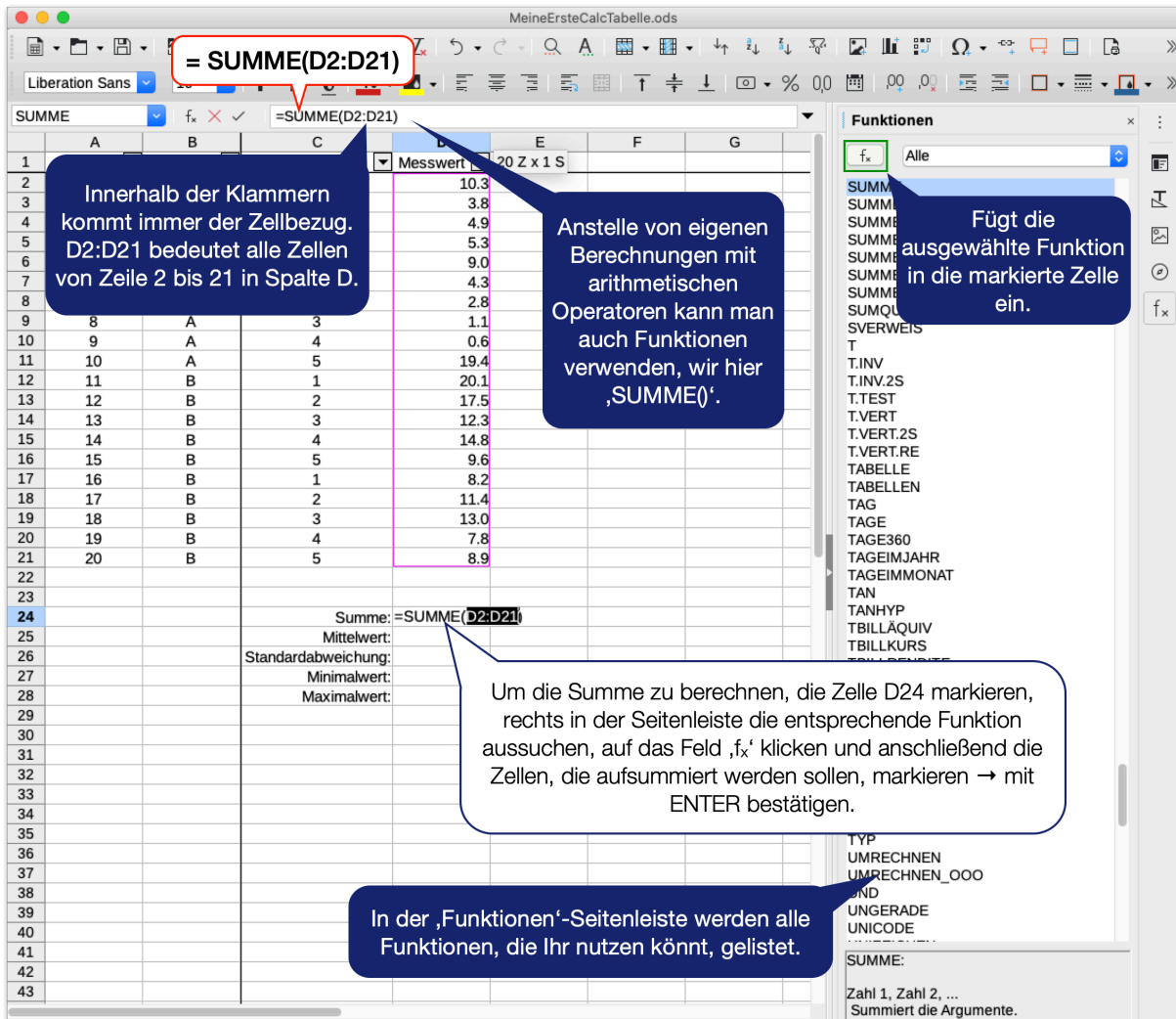


Abbildung 3.3: Verwendung von built-in Funktionen.

4. Ohne etwas einzutippen kannst Du direkt mit der Maus die Zellen markieren, die Du in die Summenbildung mit einschließen willst. Automatisch ändert sich der Befehl und statt 'Zahl' steht jetzt der markierte Zellbereich in den Klammern der Funktion (in diesem Fall '=SUMME(D2:D21)'). Bestätige mit der Eingabetaste.

Jetzt erscheint in der Zelle D24 das Endergebnis und die Zelle darunter ist aktiviert. Wenn Du nun wieder die Zelle mit dem Summenergebnis (hier D24) anklickst und anschließend mit dem Cursor in die Eingabezeile gehst und diese anklickst, kannst Du die Zellbezüge ändern. Calc hilft Dir dabei auch noch und kennzeichnet die Zellen, auf die in der Formel verwiesen wird, mit einer farblichen Umrahmung. Dabei wird die gleiche Farbe verwendet wie für die entsprechenden Zellbezüge in der Formel, die in der Zelle sichtbar sind (nicht in der Eingabezeile). Diese Funktion ist ausgesprochen hilfreich bei Fehlersuchen in Formeln, insbesondere wenn Zellen nicht aneinander liegen wie in Abb. 3.4.

Formeln können genauso kopiert werden wie Werte (also entweder mit kopieren und einfügen, oder durch herunterziehen des 'Anfassers' (kleines schwarzes Kreuz)). Wenn die Zellbezüge nicht fixiert werden, beziehen sich kopierte Formeln allerdings auf andere Zellen (siehe auch **Zellbezüge mit \$ fixieren**).

3.3.2 Zellauswahl

Zellen werden bei Auflistungen in Formeln durch ein Komma getrennt. Falls Zellen innerhalb einer Spalte aneinandergrenzen, wie z.B. Zelle D2, D3 und D4, dann kann man sich Tipparbeit sparen und gibt nur die erste und letzte Zelle dieser 'Zellgruppe' mit einem Semikolon dazwischen an: D2:D4. Calc weiß jetzt, dass alle Zellen von D2 bis D4 eingeschlossen werden sollen. Solche Zellgruppen werden auch durch Kommata getrennt (s. Abb. 3.4).

Statt die Zellen einzutippen kannst Du auch mehrere Einzelzellen und Zellgruppen mit der Maus auswählen, indem Du die <strg> bzw. <cmd> Taste gedrückt hältst während Du auf die jeweiligen Zellen klickst. Um große Zellgruppen zu markieren kannst Du auch auf die erste Zelle der Zellgruppe klicken und dann mit gedrückter Maustaste den Cursor runter ziehen und wieder loslassen (dabei immer die <str> Taste gedrückt halten!).

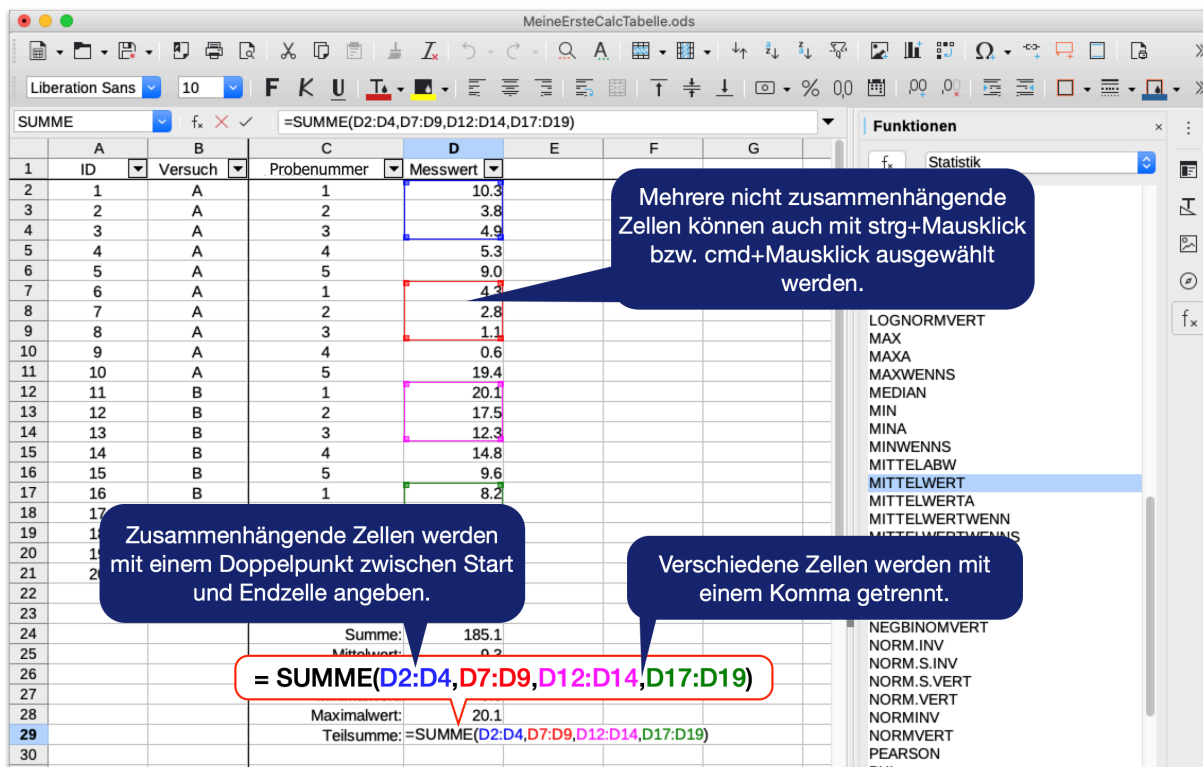


Abbildung 3.4: Format für die Mehrfachnennungen von Zellen.

3.3.3 Zellbezüge mit \$ fixieren

Im Kapitel **Logische Reihen** hast Du gelernt, dass Inhalte (Zahlenreihen, Wochentage oder Monate) automatisch nach unten ausgefüllt werden können. Das gilt auch für Formeln. Allerdings ändert sich mit jeder Zeilenerhöhung auch die Zeilenzahl des Zellbezugs in der Formel, und zwar um genau den gleichen Wert. In Abb. 3.5 wird in Zelle B2 die Summe aus Zelle A1 und A3 gebildet und anschl. die Formel in die folgenden Zellen B2 bis B15 kopiert. Da sich die Adresse von Zelle B2 gegenüber B1 um 1 Zeile erhöht hat, wird der Zellbezug in der Formel um den Wert 1 erhöht. Es wird in Zelle B2 nun die Summe aus A3 und A5 gebildet. Genauso verhält es sich mit Zelle B14. Hier hat sich die Adresse zu B1 um 13 Zeilen erhöht, daher soll hier nun die Summe aus A1+13 also A14 und A3+13 also A16 gebildet werden. Hier ist nun das Problem (wie auch bei Zelle B15), dass es

keinen Wert in A16 (bzw. A17) gibt. Darauf sollte immer geachtet werden.

Gleichermaßen verhält es sich mit Spalten. Würde die Formel aus B1 in die Zelle C1 kopiert werden, würde sich die Zelladresse um eine Spaltenzahl ändern. Entsprechend würde sich auch der Spaltenbezug in der Formel anpassen. In Zelle C2 würde dann die Summe aus B1 und B3 gebildet werden.

Hier soll in Zelle B1 die Summe von Zelle A1 und A3 gebildet werden ($1+3=4$).

Was passiert, wenn die anderen Zellen automatisch gefüllt werden (über das Klicken der unteren rechten Ecke)?

Weil sich die Formelzelle von B1 auf B2 um 1 Zeile erhöht hat, erhöht sich auch der Zellbezug in der Formel um 1 Zeile.

Achtung: den letzten Zellen fehlen Werte für den 2. Summanden, daher ist der Wert der exakt gleiche wie in Spalte A. Immer darauf achten!

Abbildung 3.5: Zelladressierung.

Diese automatische Zeilen- bzw. Spaltenanpassung ist nicht immer gewünscht. Wenn Du z. B. immer den Wert aus einer bestimmten Zelle verwenden möchtest (in Abb. 3.6 ist es beispielsweise der Wert aus Zelle A3), dann musst du die Zellbezüge 'fixieren' (Achtung: nicht zu verwechseln mit dem Zeilen/Spalten fixieren was in Zeilen und Spalten fixieren beschrieben ist). Das geschieht durch das Zeichen \$ (Dollar-Zeichen), welches vor die Zeilen- oder Spaltennummer (oder beides) in die Zelladresse gesetzt wird:

Auch hier soll in Zelle B1 die Summe aus A1 und A3 gebildet werden.

Und hier?

Da der Zeilenbezug des 2. Summanden fixiert wurde (durch '\$' bleibt Zeile 3 erhalten) erhöht sich nur der 1. Summand um 1 Zeile!

Zellbezüge fixieren bedeutet, dass beim Kopieren von Formeln in andere Zellen, weiterhin auf **dieselbe** Zelle hin verwiesen wird.

Abbildung 3.6: Fixierter Zellbezug.

Anwendung des \$ Zeichen:

- =A1 + A\$3 → damit hast Du bei Zelle A3 die **Zeile 3 fixiert** ⇒ bei **vertikalem** Kopieren der Formel bleibt der Zellbezug in Zeile 3, bei horizontalem Kopieren aber erhöht sich die Spalte zu B3, C3, D3, etc.
- =A1 + \$A3 → damit hast Du bei Zelle A3 die **Spalte A fixiert** ⇒ bei **horizontalem** Kopieren der Formel bleibt der Zellbezug in Spalte A, bei vertikalem Kopieren aber erhöht sich die Zeile zu A4, A5, A6, etc.
- =A1 + \$A\$3 → damit hast Du bei Zelle A3 die **Spalte A und die Zeile 3 fixiert** ⇒ die Zelle A3 wird adressiert, egal ob die Formel vertikal oder horizontal kopiert wird!

Kurzbefehl-Tipp: Du musst übrigens die \$ nicht immer tippen. Du kannst auch während der händischen Eingabe der Zellen, oder nachdem Du sie angeklickt hast oder nachträglich in der Eingabezeile, den Cursor auf den Zellbezug setzen und dann **<F4> drücken**. Wenn Du einmal drückst, wird aus A3 zunächst \$A\$3. Ein nochmaliges Drücken der <F4> bewirkt A\$1 und das nächste <F4> bewirkt \$A3, was den Zellbezug nur horizontal fixiert. Schließlich löst ein weiteres Betätigen der Taste ein A3 aus, der Zellbezug ist nicht mehr fixiert.

3.3.4 Beispiel einer typischen Rechenanwendung

Oft greifen Formeln oder Rechenoperationen auf aggregierte Werte oder Statistiken wie die Summe oder den Mittelwert. In Abb. 3.7 wird beispielsweise in Spalte E der prozentuale Anteil des Messwertes berechnet, wozu es die Summe aller Werte braucht. Die allgemeine mathematische Formel für alle Messwerte ist:

$$p = \frac{x_i}{\sum x_i} * 100$$

x_i steht dabei für den *i-ten x-Wert*, also für jeden einzelnen Messwert.

Nun könntest Du versuchen diese mathematische Formel ohne Zwischenschritte in eine einzige Calc Formel umzuwandeln. Auf das Beispiel in Abb. 3.7 bezogen, wäre eine naheliegende Lösung:

- =D2/SUMME(D2:D21)*100 (wobei D2 für x_1 steht)

Würdest Du nun diese Formel in Zelle E3 kopieren, würde sich leider auch der Zellbezug ändern:

- =D3/SUMME(D3:D22)*100 → somit stimmt die Summe aller Messwerte nicht mehr. Es gilt daher, den Zellbezug der Summe zu fixieren. Korrekt wäre daher:
- =D3/SUMME(D\$1:D\$21)*100.

Anstelle von verschachtelten Formeln ist es jedoch ratsam, **Zwischenergebnisse** explizit in einer Zelle **abzuspeichern** und auf diese in anderen Formeln dann zu **verweisen**. In Abb. 3.7 wurde erst in Zelle D24 die Summe gebildet, auf welche dann in der Formel in E2 verwiesen wird. Damit beim Kopieren der Formel in die nächste Zelle nicht der Mittelwert (D25) anstelle der Summe verwendet wird, wird der Zellbezug mit D\$24 fixiert. In dieser Formel wird zusätzlich die RUNDEN() Funktion verwendet, um die Dezimalzahl als ganze Zahl anzeigen zu lassen.

Die Funktion RUNDEN() unterscheidet sich im übrigen von der Zellformatierung, wo Du die Anzeige der Nachkommastellen bestimmen kannst: bei letzterem werden die Rohdaten nicht verändert sondern nur die Anzeige. Du kannst jederzeit die Formatierung auf 2,3 oder mehr Nachkommastellen setzen und diese erscheinen dann wieder. Die RUNDEN() Funktion geht mit **Datenverlust** einher, denn es 'verschwinden' die Nachkommastellen dauerhaft.

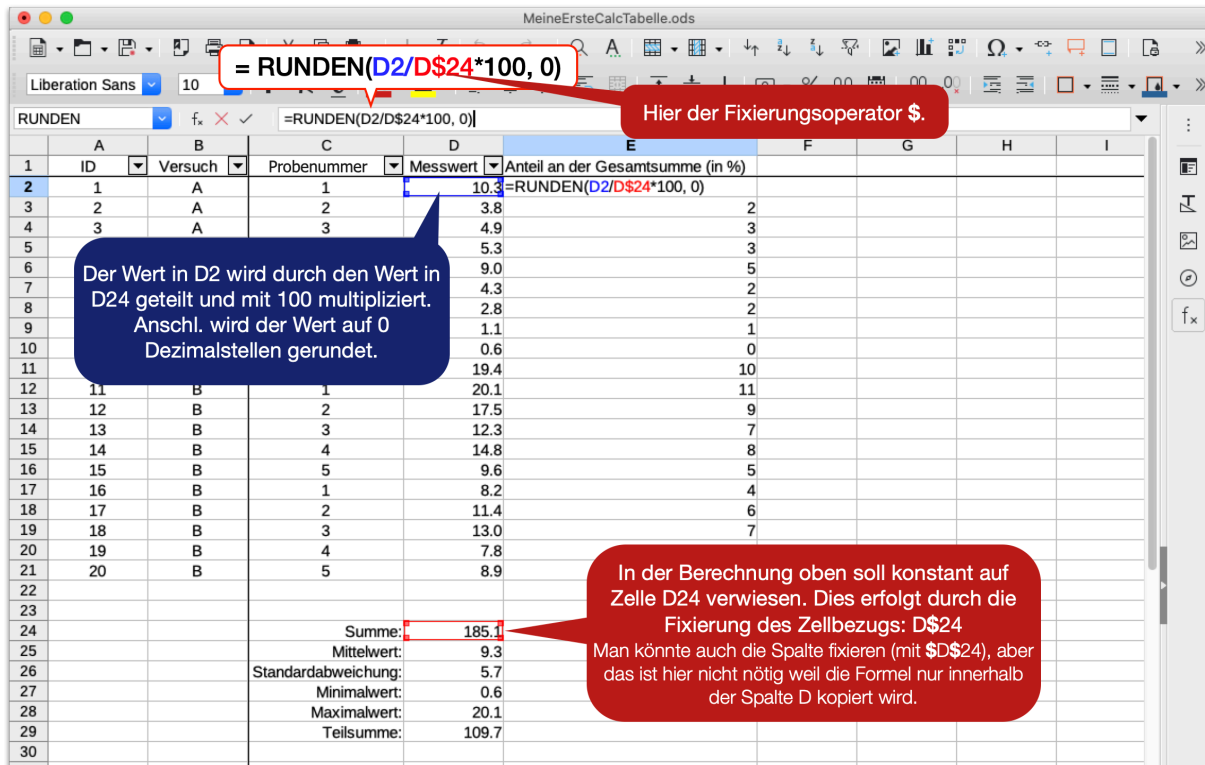


Abbildung 3.7: Beispiel einer typischen Anwendung von fixierten Zellbezügen.

Nicht jede mathematische Formel lässt sich jedoch 1:1 mit einer einzigen Calc Formel bzw. einem einzigen Rechenschritt abbilden. Ein Beispiel ist die **Varianz**:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Hier soll im Zähler die quadrierte Differenz zwischen jedem x-Wert (also jedem Messwert) und dem Mittelwert von x zum Schluss aufsummiert werden. Und darin liegt das Problem. Folgende Formel könnte man ausprobieren, allerdings wird keine zum Erfolg führen:

- =SUMME((D2:D21-MITTELWERT(D\$2-D\$21))^2)
- =SUMME((D2-MITTELWERT(D\$2-D\$21))^2 : (D21-MITTELWERT(D\$2-D\$21))^2)

Lösung: Es muss erst (in einer weiteren Spalte) für jeden Messwert die quadrierte Differenz berechnet werden, das wäre für Zeile 2:

- =(D2-MITTELWERT(D\$2-D\$21))^2.

Und diese Werte der neuen Spalte könnten dann aufsummiert und durch die Freiheitsgrade (dafür steht $n - 1$) geteilt werden.

3.4 Zellen verbinden mit der VERKETTEN() Funktion

Manchmal ist es notwendig, **IDs** zu erstellen, die eine **Kombination aus verschiedenen Merkmalen** darstellen. In unserem Beispiel-Datensatz gibt es jeweils 2 Proben die zu der gleichen Kombination aus Probenummer und Versuch gehören. Welche genau muss immer aus 2 Spalten (B und

C) abgelesen werden. Man kann aber beide in einer einzigen zusammenbringen. Dies kann dann beim Aggregieren und v.a. bei den Diagrammen hilfreich sein. Man nennt dies auch 'verketteten' von Zellen (im Englischen 'concatenate'). Bei der gleichnamigen Calc Funktion VERKETTEN() werden die zu verkettenden Zellen einfach mit Kommata gelistet. Möchtest Du ein bestimmtes Sonderzeichen oder Leerzeichen zwischen den Zellen anzeigen, wie in Abb. 3.8 der Bindestrich, dann muss dieses in Anführungszeichen an entsprechender Stelle mit gelistet werden.

Die neue ID-Spalte kann entweder hinten angehängt werden, oft ist es aber wünschenswert, diese Spalte am Anfang stehen zu haben. Dafür die Spalte neben der sich die neue befinden soll anklicken (oben den Buchstaben anklicken) und über die rechte Maustaste das Kontextmenü aufrufen und *Spalten davor einfügen* bzw. *Spalten danach einfügen*. Eine Spaltenüberschrift vergeben, in der 2. Zeile die VERKETTEN() Funktion anwenden und anschl. die Formel in die unteren Zellen kopieren. Fertig ist die neue ID-Spalte!

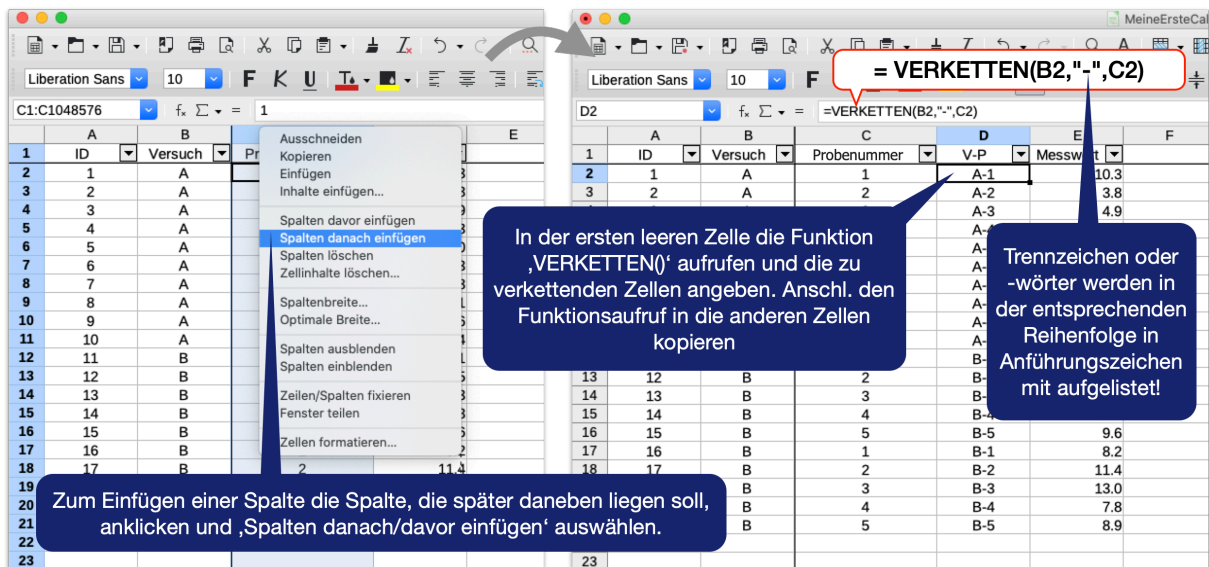


Abbildung 3.8: Verbinden von Zellen. Hier werden die Zellen in Spalte B und C mit einem Bindestrich dazwischen verbunden.

3.5 Datenaggregation mit Pivot-Tabellen

Pivot-Tabellen sind ein wichtiges **Werkzeug zur Datenauswertung** in einem Tabellenkalkulationsprogramm. In ihr werden die Rohdaten so zusammengefasst, dass wichtige Eigenschaften der Daten auf den Punkt dargestellt sind. Das führt zwar zu einem Verlust von Details, dem gegenüber steht aber ein großer Zugewinn an Übersichtlichkeit. So kann eine Pivot-Tabelle eine Ursprungstabelle mit 10000 Zeilen und 100 Spalten auf 5 Zeilen und 2 oder 3 Spalten zusammen fassen. Meistens handelt es sich dabei um typische **Kreuztabellen**, die ein **weites Format** haben. Damit aus dem Verlust aus Details allerdings eine sinnvolle Zusammenfassung entsteht, muss Du genau wissen, welchen Zweck Deine Auswertung hat und welche Fragen Du beantworten möchtest.

Im vorliegenden Beispiel möchten wir wissen, ob sich die Messwerte im Mittel zwischen

1. den Versuchen,
2. den Probennummern und
3. den Kombinationen an Probennummer und Versuch unterscheiden.

Dazu werden alle Zellen, die in der Pivot-Tabelle zusammengefasst werden sollen, inkl. der Spaltennamen markiert und der Pivot-Assistent über die Schaltfläche *Pivot-Tabelle einfügen oder bearbeiten* aufgerufen (Abb. 3.9). Im sich öffnenden Dialogfenster kann die *Aktuelle Auswahl* bestätigt werden:

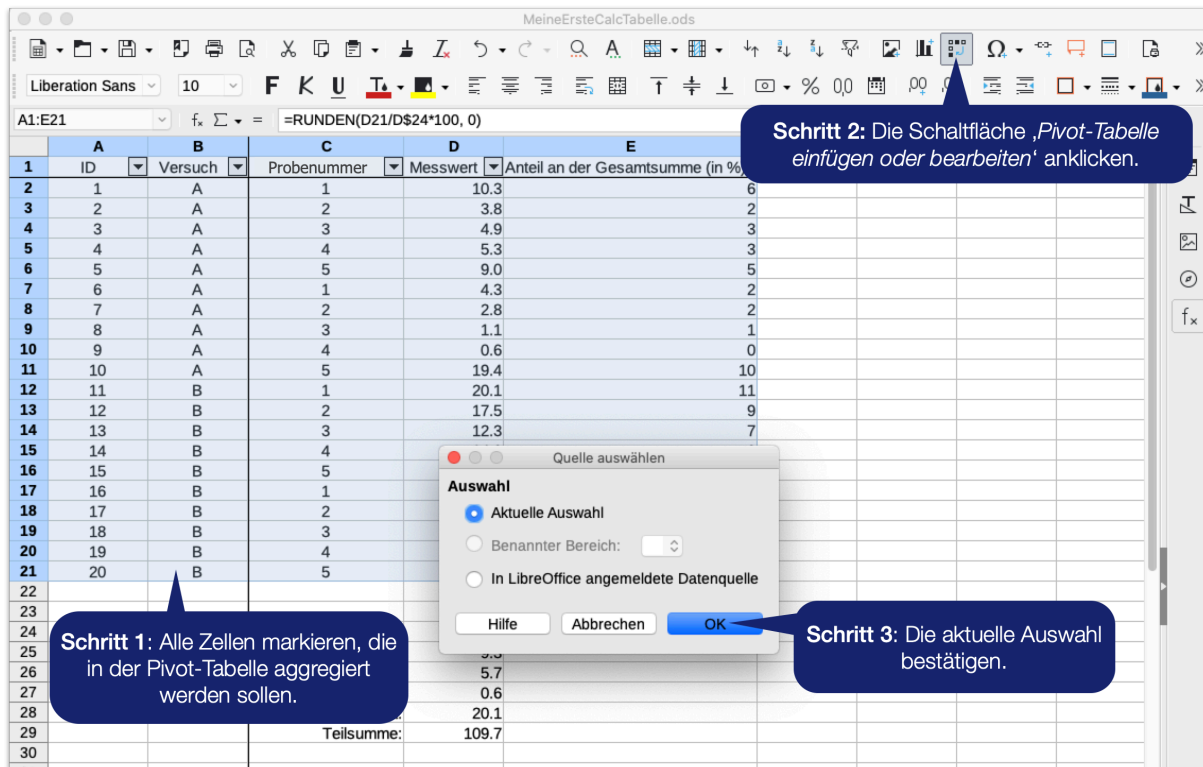


Abbildung 3.9: Erstellung einer Pivot-Tabelle für ausgewählte Zellen.

Nun erscheint der eigentliche Assistent, über dem man das Layout der Pivot-Tabelle definieren kann (Abb. 3.10). Damit ist auch gemeint, welche **deskriptiven Statistiken** für welche Variable berechnet werden soll:

- Im Zeilenfeld kannst Du die Spalten definieren, die im langen Format erscheinen sollen. D.h. diese Spalte wird nicht auf mehrere Spalten in der Pivot-Tabelle aufgeteilt.
- Im Spaltenfeld kannst Du Spalten bestimmen, die ihren Werten entsprechend auf verschiedene Spalten aufgeteilt wird.
- Nicht immer muss eine Spalte im Zeilen- oder Spaltenfeld angegeben sein, aber **IMMER** muss eine Variable im Datenfeld definiert werden, sowie die deskriptive Statistik, nach der aggregiert werden soll. Die Standardeinstellung ist die Summe. Um dies zu ändern, einfach auf 'Summe - Spaltenname' (in Abb. 3.10 ist es Summe - Messwert) doppelklicken und eine andere Funktion im *Datenfeld* Dialogfenster auswählen.

Im vorliegenden Beispiel sollen die Probenummern auf 5 Spalten (weil es die Werte 1-5 gibt) aufgeteilt werden und der Mittelwert für jede Kombination aus Versuch und Probenummer berechnet werden.

Nach Bestätigung der *OK* Schaltfläche, wird die Pivot-Tabelle in einem neuen Tabellenblatt erstellt (Abb. 3.11). Neben den eigentlichen Mittelwerten für die Kombination der verschiedenen Spalten (aka Probenummern) und Zeilen (aka Versuchen), werden in einer Pivot-Tabelle automatisch sog. *Randstatistiken* angezeigt. Das sind Zeilen- und Spaltenspezifische Statistiken, in unserem Beispiel die Mittelwerte pro Zeile und Spalte sowie für alle Werte zusammen. In dieser zusam-

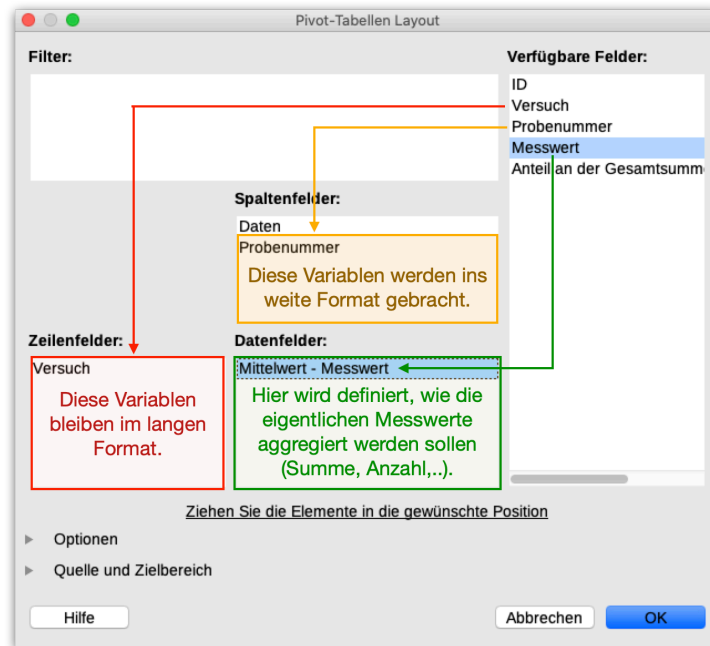


Abbildung 3.10: Mit dem Pivot-Manager das Layout bestimmen.

mengefassten Tabelle ist schnell ersichtlich, dass der mittlere Messwerte in Versuch B doppelt so hoch ist wie in Versuch A. Dies wird nicht so deutlich in der Ursprungstabelle. Weniger groß dagegen sind die Unterschiede zwischen den Probenummern (wenn beide Versuche zusammen gefasst werden):

3.6 Statistische Analysen automatisch generieren

In Calc gibt es neben Funktionen für deskriptive Statistiken (wie Mittelwert, Median, Varianz) auch die Möglichkeit, eine Zusammenfassung der deskriptiven Statistik zu erstellen und Verfahren der **inferenziellen Statistik**, wie z.B. eine Varianzanalyse oder Regression, durchzuführen. Über das Hauptmenü *Daten > Statistiken* lassen sich diese Optionen auswählen (Abb. 3.12).

Nach Auswahl der Analyse müssen im sich öffnenden Dialogfenster die entsprechenden Daten bzw. Variablen ausgewählt und einige verfahrensspezifische Einstellungen vorgenommen werden (Abb. 3.13).

Bei Auswahl der Regression werden einem neben den beiden Koeffizienten der Regressionsgleichung (Achsenabschnitt und Steigung) viele weitere Informationen ausgegeben (siehe Abb. 3.14).

:

- Bestimmtheitsmaß R^2
- Standardfehler der Regression
- Stichprobengröße und Anzahl der Variablen
- ANOVA Tabelle
- Standardfehler und t-Test Ergebnisse der Koeffizienten
- vorhergesagte Y Werte und die Abweichungen zu den beobachteten Y Werten (=Residuen)

The screenshot shows a Calc spreadsheet titled 'MeineErsteCalcTabelle.ods'. The pivot table is set to 'Mittelwert - Messwert'. The data is organized as follows:

	A	B	C	D	E	F	G	
1	Mittelwert - Messwert	Daten						
2	Versuch	1	2	3	4	5	Summe Ergebnis	
3	A	7.3	3.3	3.0	3.0	14.2	6.1	
4	B	14.2	14.5	12.7	11.3	9.3	12.4	
5	Summe Ergebnis	10.7	8.9	7.8	7.1	11.7	9.3	

Callouts in the image provide the following information:

- Die Probenummern sind jetzt in Spalten aufgeteilt.** (The sample numbers are now split into columns.)
- Nicht irritieren lassen: hier stehen keine Summen als Teilergebnisse sondern Mittelwerte!** (Don't be misled: here there are no sums as sub-results but averages!) - This callout points to the 'Summe Ergebnis' column.
- Der Gesamtmittelwert sollte unserem berechneten Mittelwert entsprechen.** (The overall average should correspond to our calculated average.) - This callout points to the value 9.3 in the bottom-right cell.
- Durch die Spaltenaufteilung in ein weites Format gibt es nur noch EINE Zeile pro Versuch A und B.** (Due to the column division in a wide format, there is now only ONE row per experiment A and B.)
- Hier sind die Mittelwerte für Probenummer 3 in Versuch A und B jeweils zu sehen, sowie der Mittelwert für A und B zusammen.** (Here you can see the averages for sample number 3 in experiments A and B, as well as the average for A and B together.) - This callout points to the values 3.0 and 7.8 in columns C and D of row 5.

Abbildung 3.11: Teilergebnisse über Pivot-Tabellen anzeigen lassen.

3.7 Aktivieren von Add-Ins

Manchmal reichen die Standardfunktionen von Calc nicht aus und man bedient sich sogenannter Add-Ins. Dazu musst Du im Menü über *Extras* > *Extension-Manager...* den Erweiterungsmanager aufrufen. Nun wird eine Liste aller aktivierbaren Add-Ins angezeigt. Ist das gesuchte Add-In nicht dabei muss es erst installiert werden, was in der Regel Administrator-Rechte auf der Maschine voraussetzt. Ein in der Biologie sehr oft genutztes und auch bereits installiertes Add-in ist z.B. **Solver für nichtlineare Probleme**.

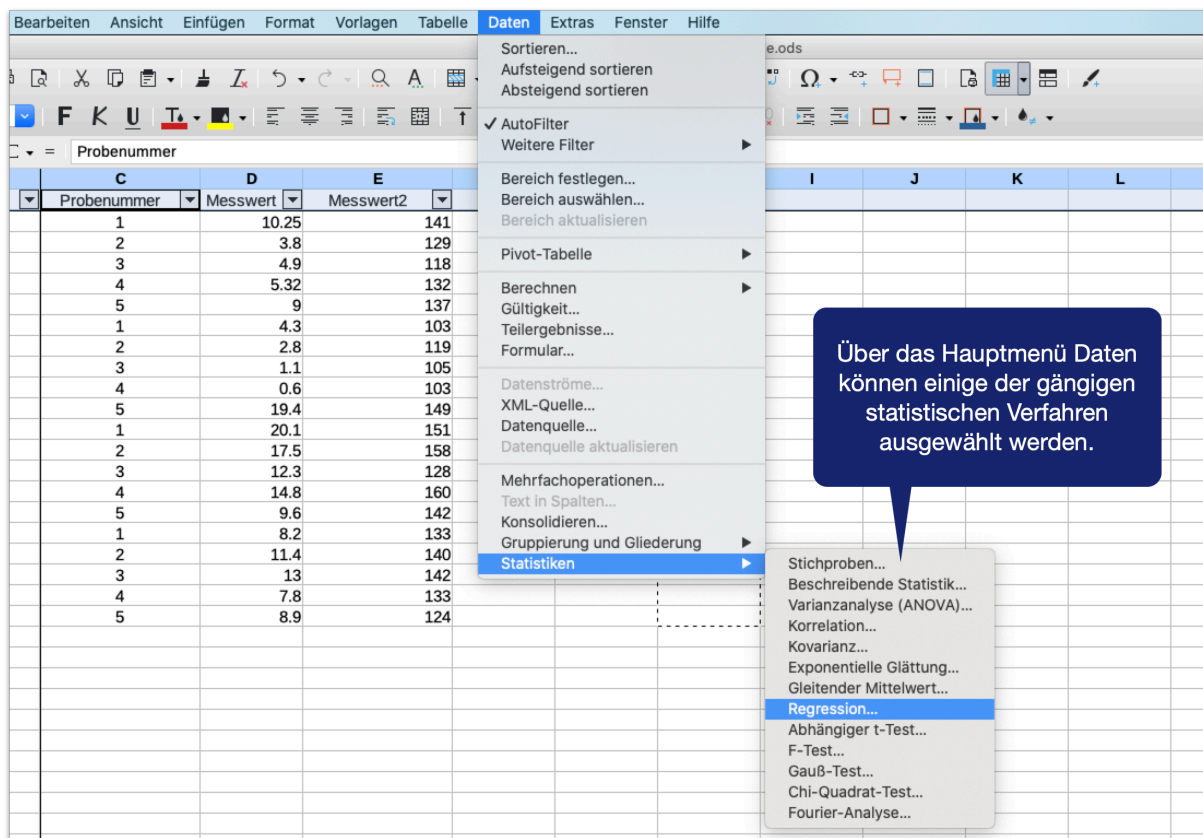


Abbildung 3.12: Auswahl von umfangreicheren Analyseverfahren in Calc.

The screenshot shows the 'Regression' dialog box in LibreOffice Calc. The background spreadsheet has columns C, D, and E selected. Column C is 'Probenummer', D is 'Messwert', and E is 'Messwert2'. The data in these columns is as follows:

Probenummer	Messwert	Messwert2
1	10.25	141
2	3.8	129
3	4.9	118
4	5.32	132
5	9	137
1	4.3	103
2	2.8	119
3	1.1	105
4	0.6	103
5	19.4	149
1	20.1	151
2	17.5	158
3	12.3	128
4	14.8	160
5	9.6	142
1	8.2	133
2	11.4	140
3	13	142
4	7.8	133
5	8.9	124

The 'Regression' dialog box is open, showing the following settings:

- Daten:**
 - Bereich unabhängiger Variablen (X): \$ErsteProbenahme.\$E\$1:\$E\$21
 - Bereich abhängiger Variablen (Y): \$ErsteProbenahme.\$D\$1:\$D\$21
 - X- und Y-Bereiche mit Beschriftung
 - Ergebnisse nach: \$G\$1
- Gruppirt nach:**
 - Spalten
 - Zeilen
- Regressionstypen festlegen:**
 - Lineare Regression
 - Logarithmische Regression
 - Potenzielle Regression
- Optionen:**
 - Konfidenzintervall: 0.95
 - Reste berechnen
 - Schnitt bei Null erzwingen

Two callout boxes provide instructions:

- Left callout: "Bei Auswahl der Regression öffnet sich dieses Fenster: wähle hier die X- und Y-Variable, die Art der Regression (z.B. linear), etc."
- Right callout: "Bestätigen, damit das Ergebnis in Zelle G1 anschl. erscheint."

Abbildung 3.13: Einstellungen bei Durchführung einer Regression.

D	E	F	G	H	I	J	K	L	M
wert	Messwert2		Regression						
10.25	141		Regressionsmodell	Linear					
3.8	129								
4.9	118		RGP Rohausgabe						
5.32	132		0.295286881028337	-29.827718704101					
9	137		0.040464047200644	5.3961124433803					
4.3	103		0.747381487589931	2.9581820044832					
2.8	119		53.253685362461	18					
1.1	105		466.014521110337	157.51513388966					
0.6	103								
19.4	149		Regressionsstatistik						
20.1	151		R ²	0.7473814875899					
17.5	158		Standardfehler	2.9581820044832					
12.3	128		Anzahl der X-Variablen	1					
14.8	160		Beobachtungen	20					
9.6	142		Adjustiertes R ²	0.7333471257894					
8.2	133								
11.4	140		Varianzanalyse (ANOVA)						
13	142			fg	BB	MS	F	Schrittweite F	
7.8	133		Regression		1.466.0145211103	466.0145211	53.25368536	8.843207E-07	
8.9	124		Residuum		18.157.5151338897	8.750840772			
			Insgesamt		19	623.529655			
			Konfidenzintervall		0.95				
				Koeffizienten	Standardfehler	t-Statistik	P-Wert	Untere 95%	Obere 95%
			Achsenabschnitt	-29.827718704101	5.39611244338	-5.527631052	3.007344E-05	-41.16453027	-18.49090714
			Messwert2	0.295286881028337	0.040464047201	7.297512272	8.843207E-07	0.210275072	0.38029869
			Messwert2	Prognostiziertes Y	Messwert	Residuum			
			141	11.807731520895	10.25	-1.557731521			
			129	8.2642889485551	3.8	-4.464288949			
			118	5.0161332572434	4.9	-0.116133257			
			132	9.1501495916401	5.32	-3.830149592			
			137	10.626583996782	9	-1.626583997			
			103	0.5868300418183	4.3	3.713169958			
			119	5.3114201382717	2.8	-2.511420138			
			105	1.177403803875	1.1	-0.077403804			
			103	0.5868300418183	0.6	0.013169958			
			149	14.170026569122	19.4	5.229973431			
			151	14.760600331179	20.1	5.339399669			
			158	16.827608498377	17.5	0.672391502			
			128	7.9690020675267	12.3	4.330997932			
			160	17.418182260434	14.8	-2.61818226			
			142	12.103018401924	9.6	-2.503018402			
			133	9.4454364726684	8.2	-1.245436473			
			140	11.512444639867	11.4	-0.11244464			
			142	12.103018401924	13	0.896981598			
			133	9.4454364726684	7.8	-1.645436473			
			124	6.7878545434134	8.9	2.112145457			

Automatisch generierter Output der Regression

Der Y-Achsenabschnitt a.

Der Steigungsparameter b hinsichtlich der X-Variable 'Messwert2'.

Abbildung 3.14: Automatisch erzeugte Ergebnisse der linearen Regression.

4 Diagramme

Als Grafik bezeichnet man jede nicht-fotografische zweidimensionale Darstellung. Das Diagramm stellt eine Untergruppe der Grafiken dar und ist die **visuelle Darstellung von Daten und Prozessen in abstrahierter Form**. Es ist ein wichtiges Werkzeug um Zahlen besser zu verstehen und Zusammenhänge und Muster besser zu erkennen.

Calc bietet eine Reihe von möglichen Diagrammen, die alle über die Schaltfläche *Diagramm einfügen* in der Standard-Symbolleiste erreichbar sind (s. Abb. 4.1). Vorher einfach die Zellen markieren die visualisiert werden sollen, inkl. der Spaltennamen. Die Schaltfläche startet dann den **Diagramm-Assistenten** in dem man nacheinander den Typ der Grafik, die Zellen, die als Werte, Kategorien oder sonstiges zu verwenden sind, die Achsenbeschriftung, die Diagrammbeschriftung und vieles mehr einstellen kann (Abb. 4.2).

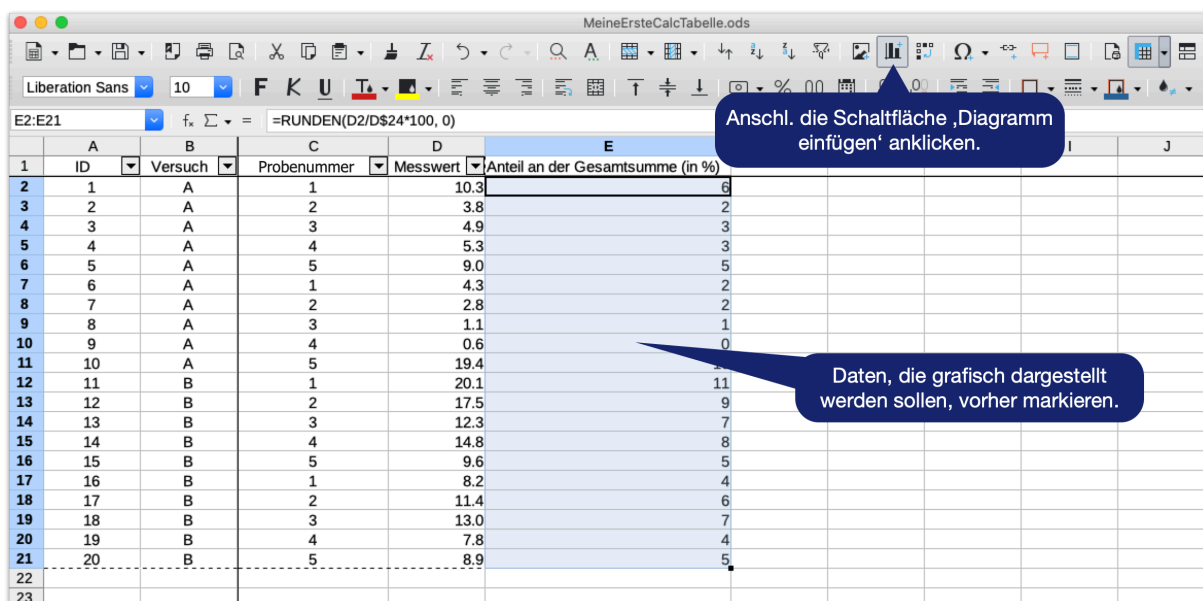


Abbildung 4.1: Starten des Diagramm-Assistenten für ausgewählten Datenbereich.

Schließlich kann man das Diagramm im Tabellenblatt erstellen lassen oder als eigenes Diagramm-Arbeitsblatt außerhalb der Tabelle. Auch nachträglich kann man die Skalierung der Achsen, die Farbe und Form des Diagramms, die Beschriftung, etc. verändern. Um zurück in den Diagramm-Modus zu kommen (erkennbar an der grauen Diagramm-Umrandung und daran, dass oben in der GUI nur eine Diagramm-Symbolleiste zu sehen ist), musst Du auf das Diagramm doppelklicken (s. Abb. 4.3). Probiere selbst ein bisschen rum. Auch hier ist die rechte Maustaste wieder Dein Freund. Wenn Du ein Diagrammelement markierst und dann die rechte Maustaste drückst wird Dir angezeigt, was Du machen kannst. Es erscheinen auch weitere Optionen in der *Eigenschaften*-Seitenleiste, mit der Du die Diagrammelemente anpassen kannst.

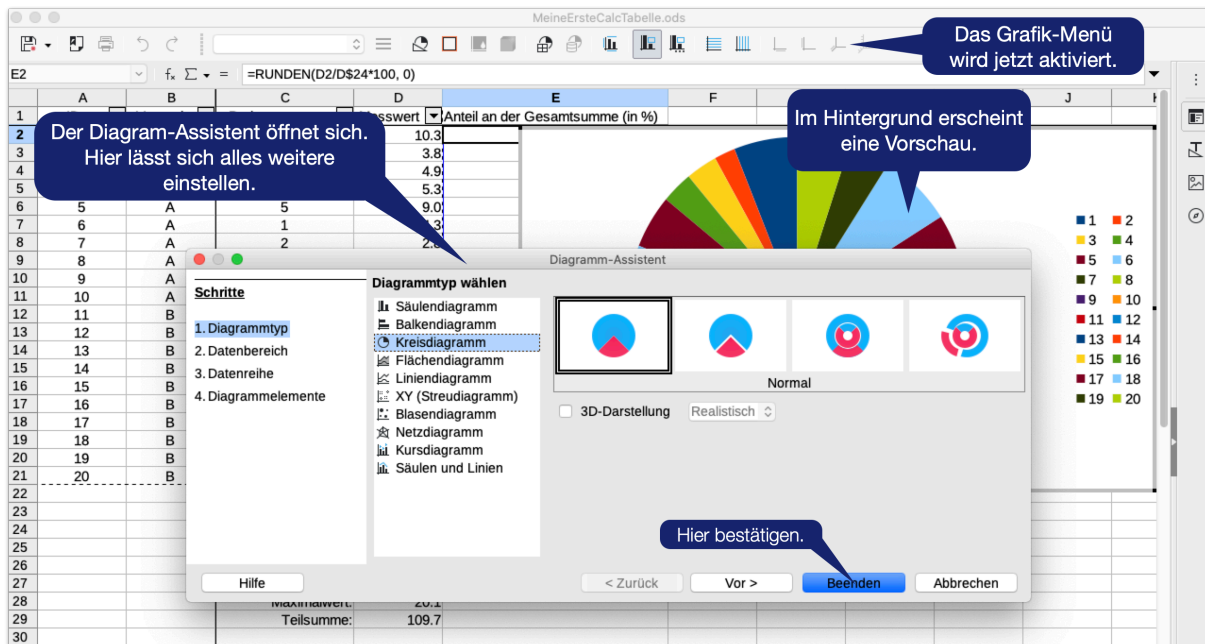


Abbildung 4.2: Auswahl des Diagramm-Typen.

4.1 Einfaches Streudiagramm mit Trendlinie

Beim Streudiagramm müssen meist keine Daten vorab aggregiert werden. Hier reicht es einfach die Spalten mit den Rohdaten auszuwählen und dann über die Diagrammschaltfläche und dem sich darauf öffnenden Diagramm-Assistenten in Schritt 1 (Diagrammtyp) 'XY (Streudiagramm) > Nur Punkte' auszuwählen (s. Abb. 4.4).

Calc stellt automatisch immer die erste Variable (mit dem niedrigeren Buchstaben) auf der X-Achse dar, die zweite Variable auf der Y-Achse. Diese Zuordnung kann aber unter Schritt 3 (Datenreihe) geändert werden. Klicke dazu oben links in dem Feld 'Datenbereiche' auf die 'X'-bzw. 'Y-Werte' und wähle anschl. unter 'Bereich für X- bzw. Y-Werte' die Zellen der Variable, die auf dieser Achse angezeigt werden sollen.

Unter Schritt 4 (Diagrammelemente) sollte immer die Achsenbeschriftung erfolgen (s. Abb. 4.5). Wenn nicht in Schritt 2 (Datenbereich) ein Häkchen gesetzt wurde unter 'Erste Spalte als Beschriftung', wird nicht erkenntlich im Diagramm, welche Variable auf welcher Achse angezeigt wird. Daher ist es wichtig, in Schritt 4 diese Beschriftung nachzutragen. Hier können auch weitere Informationen als nur der Variablenname angezeigt werden, z.B. die Einheiten.

Nachdem das Diagramm erstellt wurde, können verschiedene Modifikationen vorgenommen werden, wie z.B. die Entfernung der Legende, die in diesem Fall keinen Informationsgewinn bringt (s. Abb. 4.6).

Wenn davon ausgegangen werden kann, dass die Variable auf der X-Achse die Werte der Variable auf der Y-Achse beeinflusst, kann auch eine Kurvenanpassung durchgeführt werden. Eine Kurvenanpassung ist eine Technik mit der man versucht, eine gegebene mathematische Modellfunktion bestmöglich an Datenpunkte anzupassen. Meist wird hierbei eine lineare Funktion in der Form $Y = a + b * X$ verwendet und eine lineare/gerade Trendlinie zum Diagramm hinzugefügt (Abb. 4.7). Über das Menü der Trendlinie kann ausgewählt werden, welche Form die Kurve haben soll, ob die Kurve durch den Koordinatenursprung gehen soll (ob z.B. der Achsenabschnitt a in der linearen Gleichung gleich 0 gesetzt werden soll) oder ob die Regressionsgleichung und das

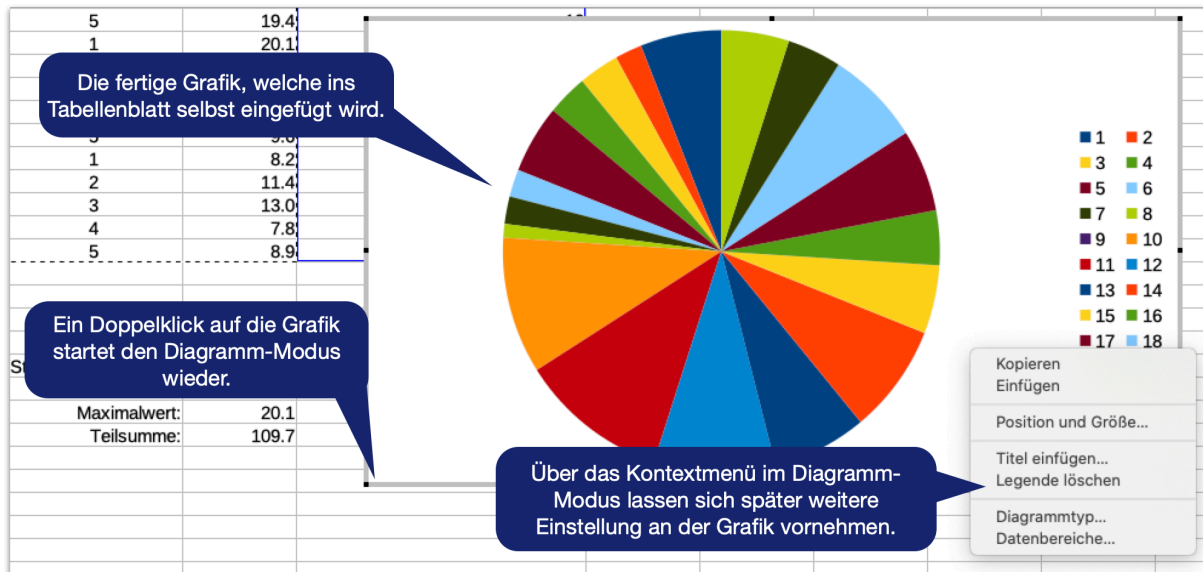


Abbildung 4.3: Nachträgliche Bearbeitung von Diagrammen.

Bestimmtheitsmaß R^2 angezeigt werden sollen (s. Abb. 4.8).

Textfelder wie zum Beispiel die numerischen Ergebnisse der Regression können auch angepasst werden. Abb. 4.9 zeigt, wie die Nachkommastellen der Koeffizienten der automatisch generierten Regressionsgleichung modifiziert werden können, um das finale Diagramm zu erhalten (Abb. 4.10).

4.2 Datenaggregation - Beispiel Säulendiagramm

Für viele der Diagrammtypen müssen die Daten vorab aggregiert und teils auch umstrukturiert werden. Zum Beispiel ein Säulendiagramm zum Vergleich von 2 oder mehr Gruppen hinsichtlich einer numerischen Variable basiert meist auf dem Mittelwert. Dazu wird oft noch als Fehlerbalken die Standardabweichung oder der Standardfehler angezeigt. Dies wird nicht von Calc automatisch berechnet, wenn auf die Rohdaten nur verwiesen wird. Beides musst Du selbst berechnen (s. Abb. 4.11). Du kannst dabei die aggregierten Daten im weiten Format anordnen - wie in Abb. 4.11 - oder auch im langen. In beiden Fällen 'versteht' Calc meist wie die Daten visualisiert werden sollen.

Für das Säulendiagramm werden erstmal nur die Mittelwerte und die Überschriften (für die X-Achse) ausgewählt (s. Abb. 4.12):

In Schritt 2 und 3 des Diagramm-Assistenten können die Standardeinstellung so bleiben, in Schritt 4 wird ein Titel und eine Y-Achsenbeschriftung definiert. (Abb. 4.13):

Die Fehlerbalken zur Darstellung der Streuung der Daten können erst im Anschluss hinzugefügt werden. Dazu auf das fertige Diagramm doppelklicken bis die schwarze Umrahmung erscheint. Da die Fehlerbalken ein Teil der Diagrammelemente 'Säule' sind, müssen diese nun explizit durch einen Mausklick auf eine der beiden Säulen ausgewählt werden. Ein kleines Viereck sollte mittig in jeder Säule erscheinen (Abb. 4.14). Mit der rechten Maustaste das Kontextmenü aufrufen und die Option *Y-Fehlerbalken einfügen...* auswählen. Wäre es ein Balkendiagramm mit horizontalen Balken, wäre X-Fehlerbalken die richtige Wahl. Im sich öffnenden Dialogfenster werden die beiden

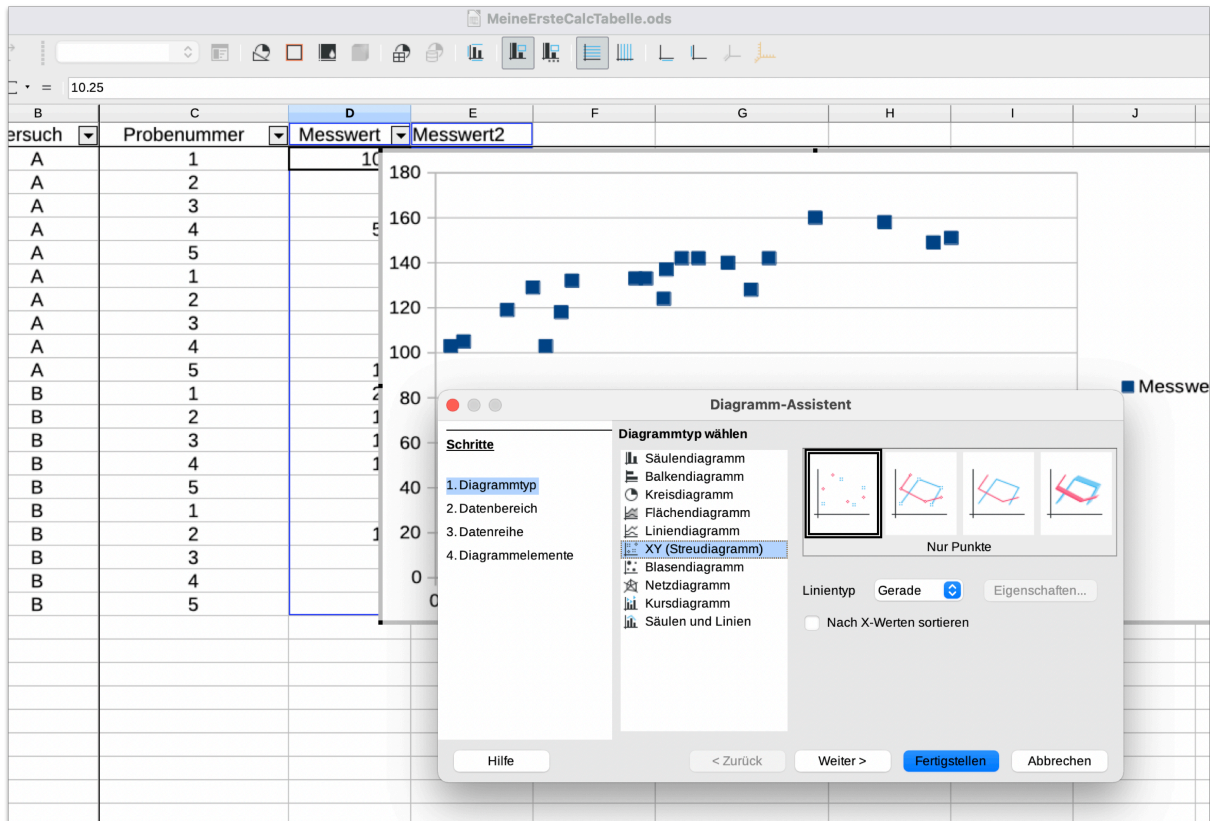


Abbildung 4.4: Erstellung eines einfachen bivariaten Streudiagramms.

Zellen, welche die berechnete Standardabweichung enthalten, unter *Parameter* angegeben. Hier kannst Du auch entscheiden, ob Du die Fehlerbalken in beide Richtungen anzeigen möchtest oder nur nach oben (also in den positiven Bereich). In Abb. 4.14 wird ersteres gewählt. Fertig ist das Diagramm (Abb. 4.15)!

Wichtig: ein Diagramm sollte verständlich sein und selbst erklärend. Es sollte eine Überschrift, Achsenbeschriftungen (gegebenenfalls mit Maßeinheiten) eine Legende für die einzelnen Datenreihen, evtl. eine zweite Y-Achse und eine vernünftige Skalierung der Achsen haben. Daneben sollte jedes Diagramm auch eine Abbildungslegende haben. Hier wird kurz und präzise zusammengefasst, was das Diagramm darstellt und wenn nötig werden einige Diagrammelemente kurz erläutert (wie in Abb. 4.15).

4.3 Datenaggregation - Beispiel Histogramm

Histogramme werden zur Betrachtung der **Verteilung** einer **einzelnen kontinuierlichen Variable** (X-Variable) verwendet. Dafür wird die Variable (X) in gleich große Wertebereiche (Klassen), im Englischen spricht man von '**bins**', aufgeteilt und anschl. wird jede Beobachtung der entsprechenden Klasse zugeordnet. Diese Klassen werden dann als Säulen auf der X-Achse angezeigt, und die Höhe der Säule wird durch die **Häufigkeit** der Beobachtungen definiert (die Y-Achse). Es ähnelt somit einem Säulendiagramm, nur dass die Säulen aufgrund der lückenlosen Klassenaufteilung direkt aneinander liegen.

Die Anzahl der Klassen oder bins bzw. die Größe des Wertebereichs oder die Weite der 'bins', sollte



Abbildung 4.5: Hinzufügen eines Titels und der Achsenbeschriftung in Schritt 4 des Diagramm-Assistenten.

so gewählt werden, dass sich die Verteilungskurve gut erkennen lässt. Es macht beispielsweise keinen Sinn, nur 2 Klassen zu bilden. Hier wirst Du nicht erkennen können, ob die Verteilung eher uni oder bimodal ist, wo in etwa der Median liegt und wie stark die Streuung ist. Genauso wenig empfiehlt es sich so viele Klassen zu wählen, dass in jeder Klasse nur 1 Beobachtung ist. Die Wahl der Anzahl hängt auch von der Stichprobengröße (N) ab. Ist $N = 1000$ kannst Du gut 50 oder mehr Klassen definieren. Ist $N = 50$, sollte die Anzahl nicht größer als 15 sein. Im vorliegenden Beispiel ist $N = 20$, da könnten 5-6 Klassen noch in Ordnung sein.

Die Aggregation der Daten in Klassen und Anzahl der Beobachtungen pro Klasse wird nicht von Calc bei der Erstellung des Diagramms vorgenommen (anders als bei der Programmiersprache R). Sprich, Du musst dies vorab selbst machen. Schau Dir dazu an, wie groß die Stichprobengröße und die Spannweite Deiner kontinuierlichen Variable ist und teile diesen Gesamtwertebereich in eine passende Anzahl gleich großer Intervalle. In Abb. 4.16 ist zu sehen, dass die Spannweite 0.6 - 20.1 beträgt. Hier bietet es sich an, die Daten in 6 Klassen mit jeweils einem Wertebereich von 4 Einheiten zu erstellen (siehe Spalte G).

Nun könntest Du für jede Klasse die Anzahl der Messwerte zählen, die innerhalb des Intervallbereichs dieser Klasse liegen. Bei 20 Werten ist dies noch schnell getan. Ist der Datensatz jedoch größer, bietet es sich an, eine Calc Funktion zu verwenden. Diese nennt sich in diesem Fall **HÄUFIGKEIT()**. Allerdings wird sie etwas anders angewendet als alle bisher besprochenen Funktionen! Wird die Funktion einmalig in einer Zelle aufgerufen, erstellt sie eine komplette **Häufigkeitsmatrix** für diese wie auch die angrenzenden Zellen. Die Formel muss nicht in die anderen Zellen kopiert werden.

Damit das auch wirklich funktioniert, darf bei der Formel die schließende runde Klammer NICHT gesetzt werden. Stattdessen muss die Tastenkombination <shift> und <strg> (bzw. <cmd>) und <enter> gedrückt werden:

In Abb. 4.17 kannst Du sehen, wie in Zelle H1 die Formel für die erste Messwert-Klasse angewendet wird:

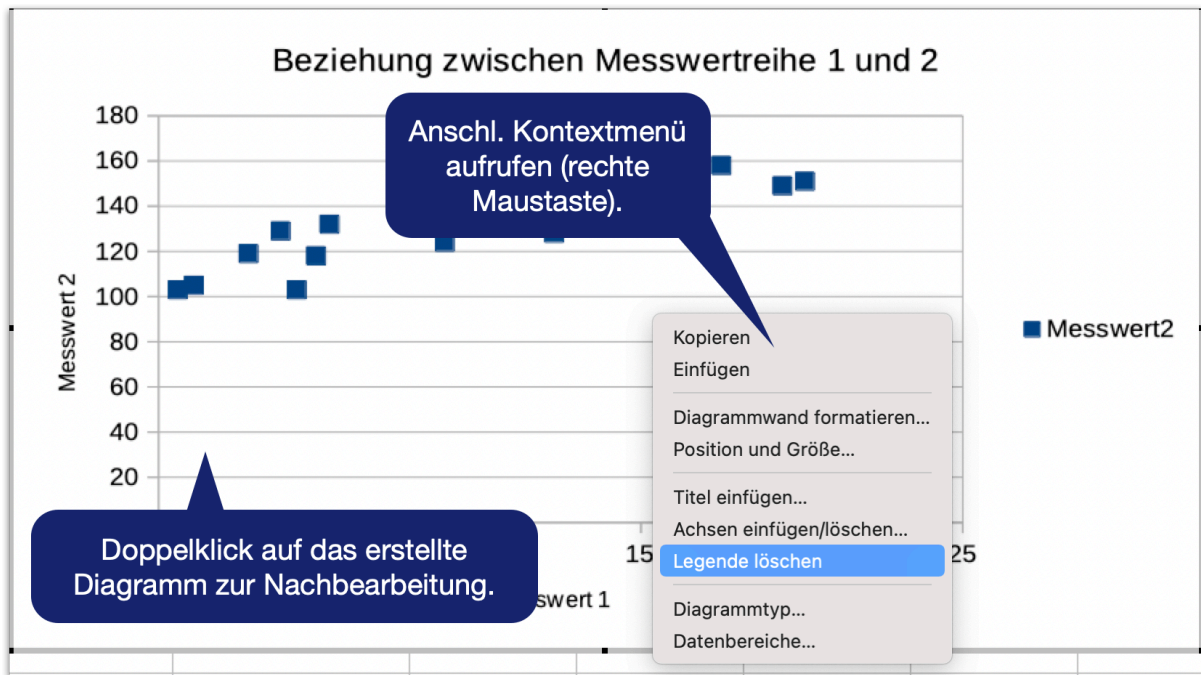


Abbildung 4.6: Entfernung der wenig informativen Legende nach Erstellung des Diagramms über das Kontextmenü.

■ = HÄUFIGKEIT(D2:D21; G2:G7

D2:D21 bezieht sich auf die Zellen mit den einzelnen Messwerten und G2:G7 sind die selbst definierten Klassen, zu denen die Messwerte zugeordnet werden sollen. Wichtig ist hierbei, dass beide Zellbezüge mit einem **Semikolon** getrennt werden.

Nun wird **statt der schließenden Klammer die Tastenkombination gedrückt** und automatisch wird nicht nur für diese Zelle, sondern für alle anderen Zellen in H, die in G eine korrespondierende Klasse haben, die Häufigkeit bestimmt. Calc gibt darüber hinaus immer für 1 weitere Zelle, die keine korrespondierende Klasse hat, die Häufigkeit an. Wenn die Klassen richtig gewählt wurden, sollte dieser Wert immer Null sein.

Obwohl Histogramme eigentlich ein eigener Diagrammtyp sind, gibt es keine Vorlage dazu in Calc. Es muss stattdessen der Umweg über das einfache Säulendiagramm genommen werden (s. Abb. 4.18) und später einige Anpassungen vorgenommen werden:

In Schritt 2 des Diagramm-Assistenten wird definiert, dass die erste Zeile wie auch Spalte nur der Beschriftung dienen (Abb. 4.19):

Das erstellte Säulendiagramm zeigt die typischen Lücken zwischen den Säulen, die es nicht bei einem Histogramm geben kann, da ja die Intervalle aneinander grenzen. Um dies zu ändern werden wieder die Säulen mit einem Klick markiert und im Kontextmenü *Datenreihe formatieren* gewählt (s. Abb. 4.20). Hier nun den Abstand unter *Einstellungen* von 100% auf 0% setzen und bestätigen. Wenn Du willst, kannst Du auch noch die Füllfarbe, die Transparenz und die Umrandungslinie anpassen (Abb. 4.21).

Voila, fertig ist das Histogramm:

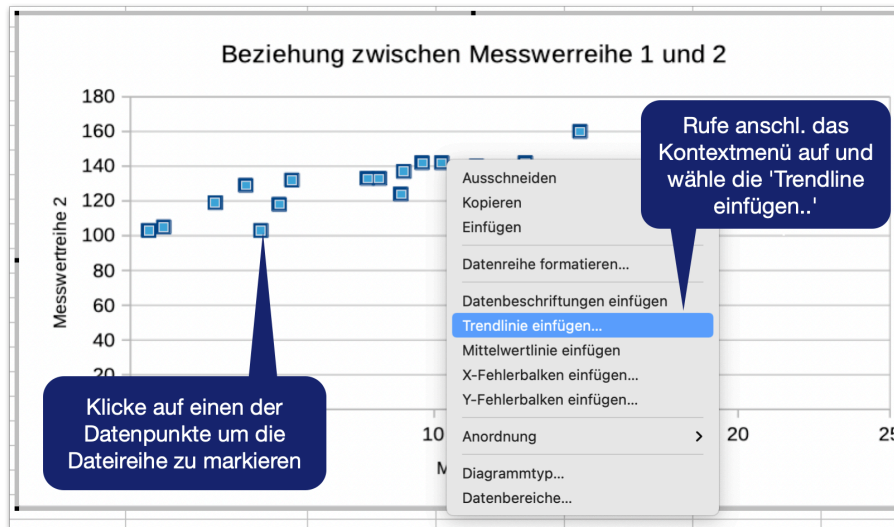


Abbildung 4.7: Hinzufügen einer Trendlinie, indem die Datenreihe markiert und anschl. das Kontextmenü über die rechte Maustaste aufgerufen wird.

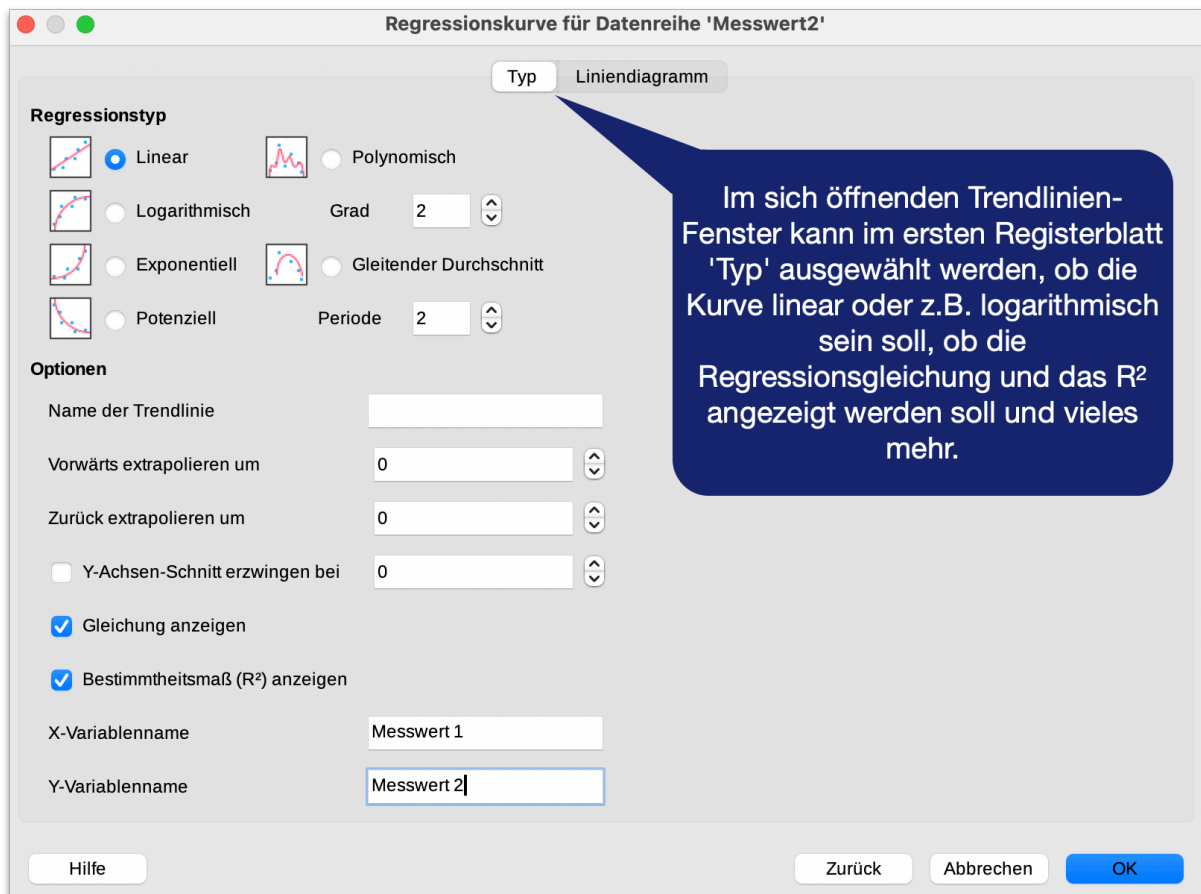


Abbildung 4.8: Im Trendlinien-Menü können verschiedene Einstellungen zur Kurvenanpassung und zur Anzeige der Regressionsergebnisse vorgenommen werden.

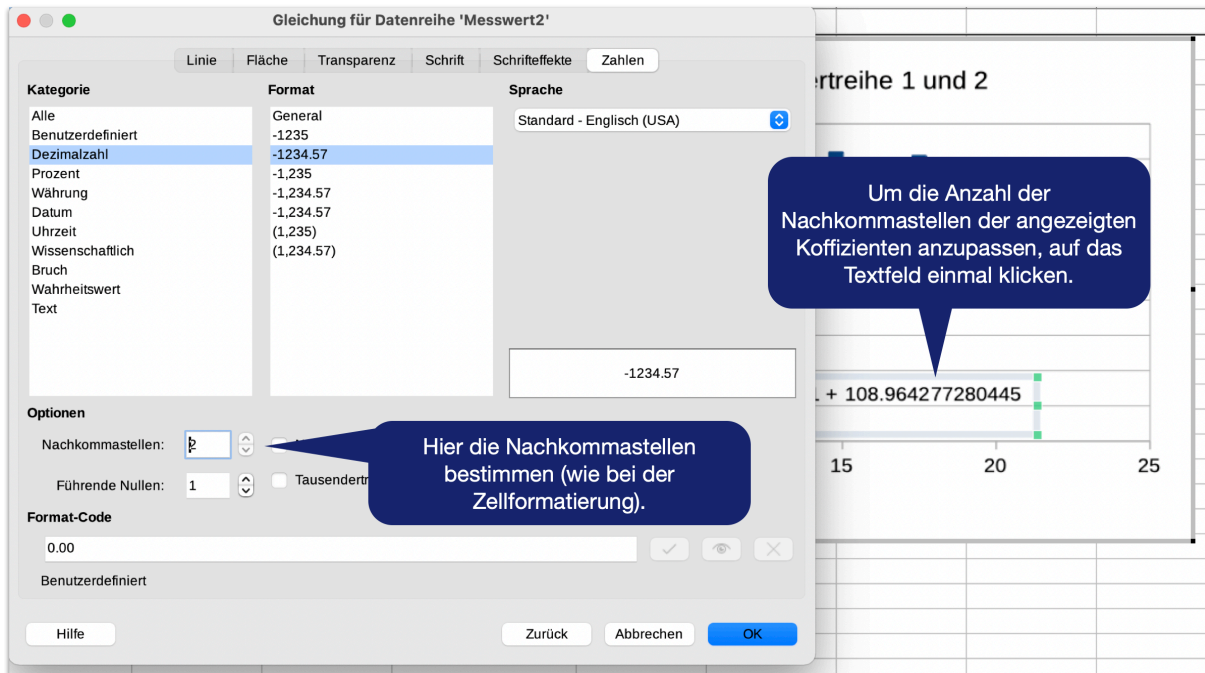


Abbildung 4.9: Anpassung der Regressionsergebnisse über das Kontextmenü des Textfelds.

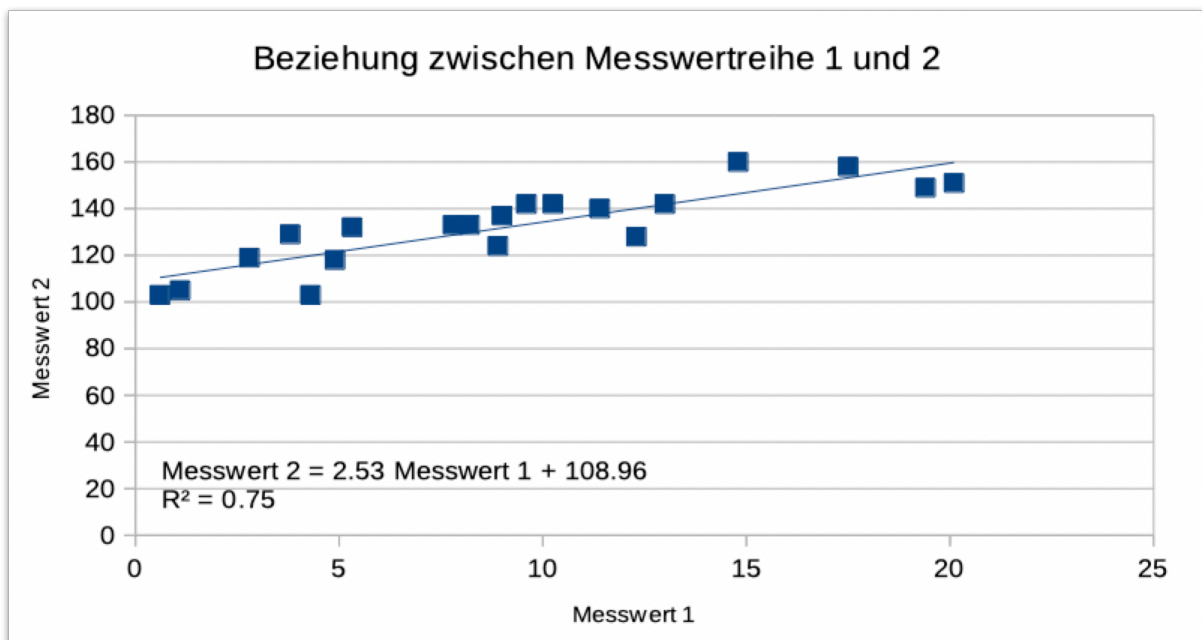


Abbildung 4.10: Das finale Streudiagramm mit Trendlinie und linearer Regressionsgleichung.

4 Diagramme

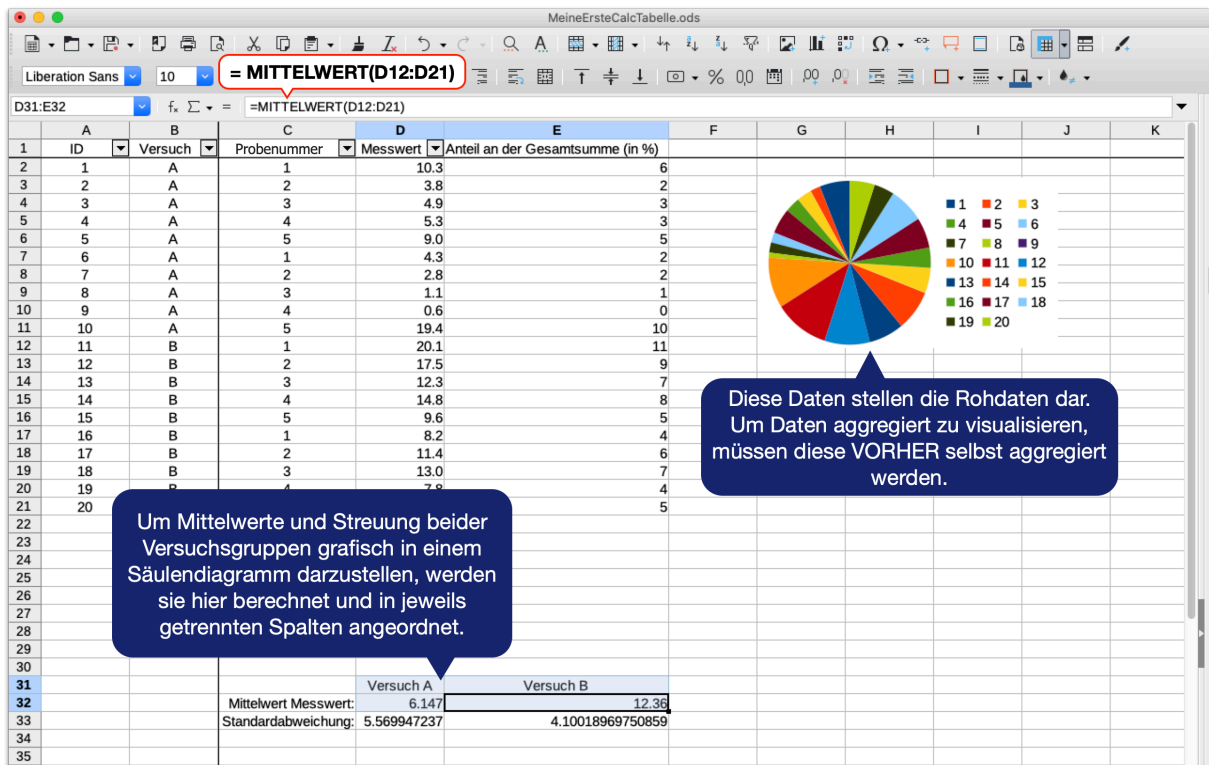


Abbildung 4.11: Datenaggregation als Vorbereitung.

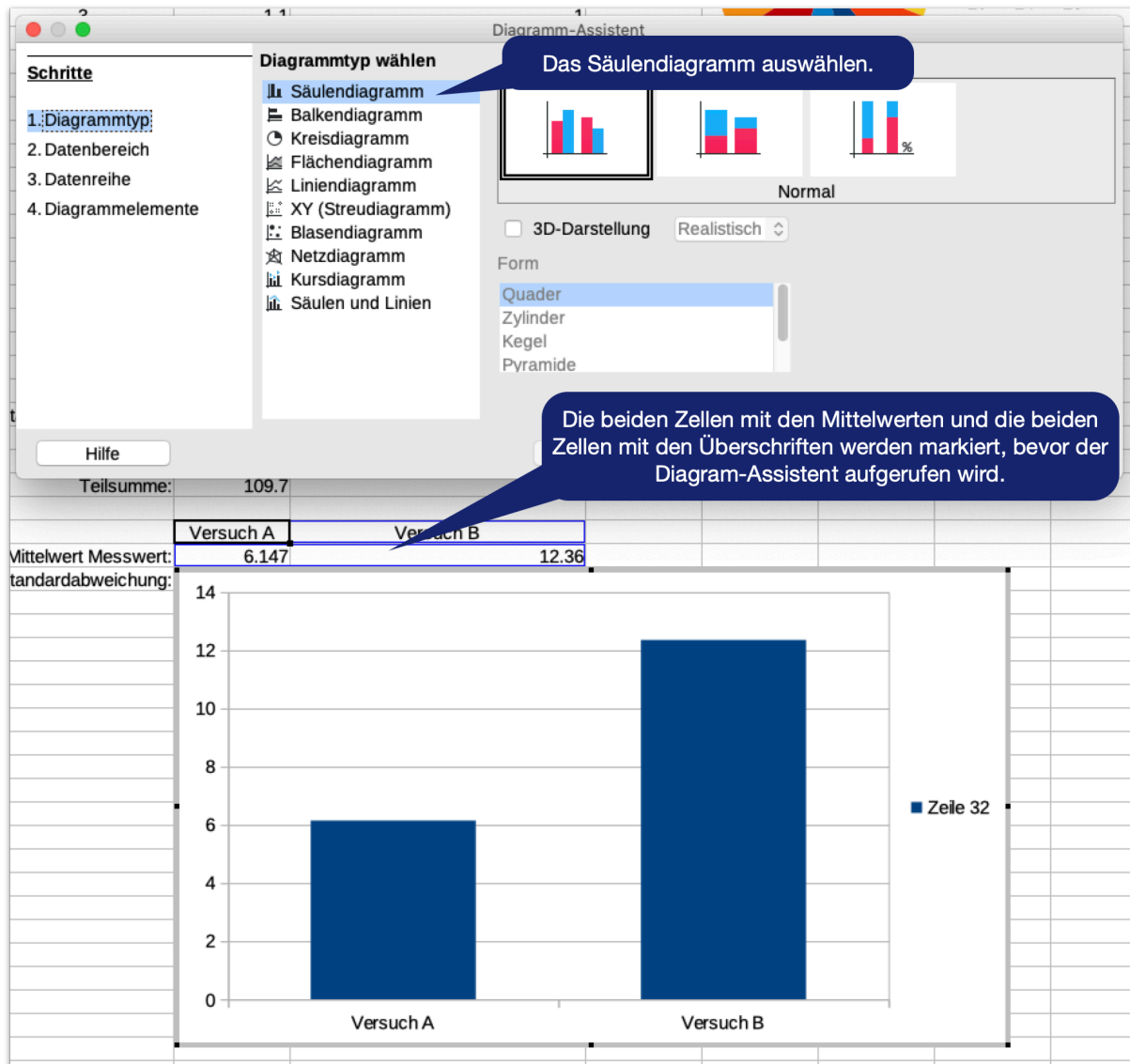


Abbildung 4.12: Säulendiagramm zur Darstellung von Gruppenvergleichen.

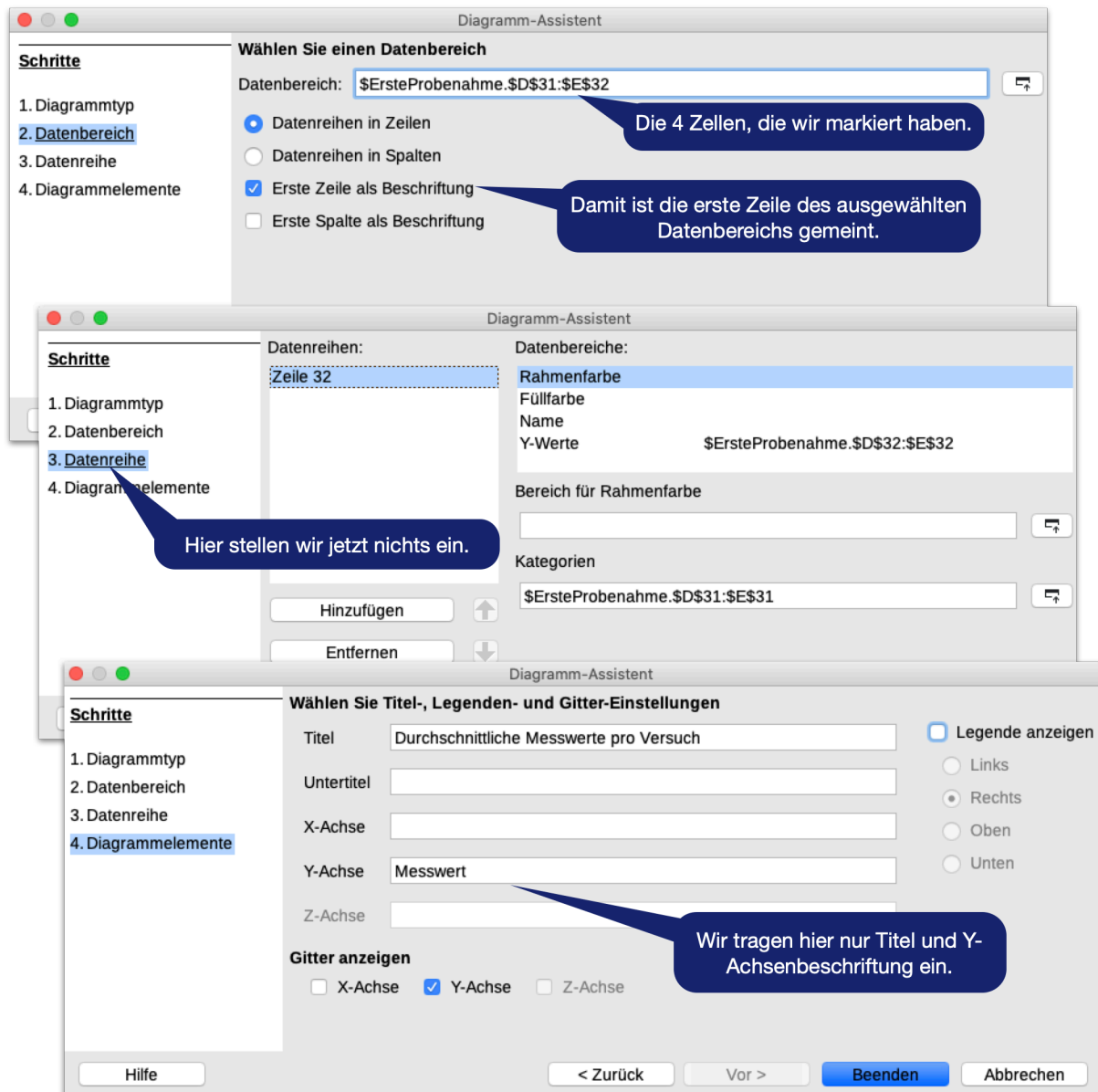


Abbildung 4.13: Die Hauptschritte im Diagramm-Assistenten.

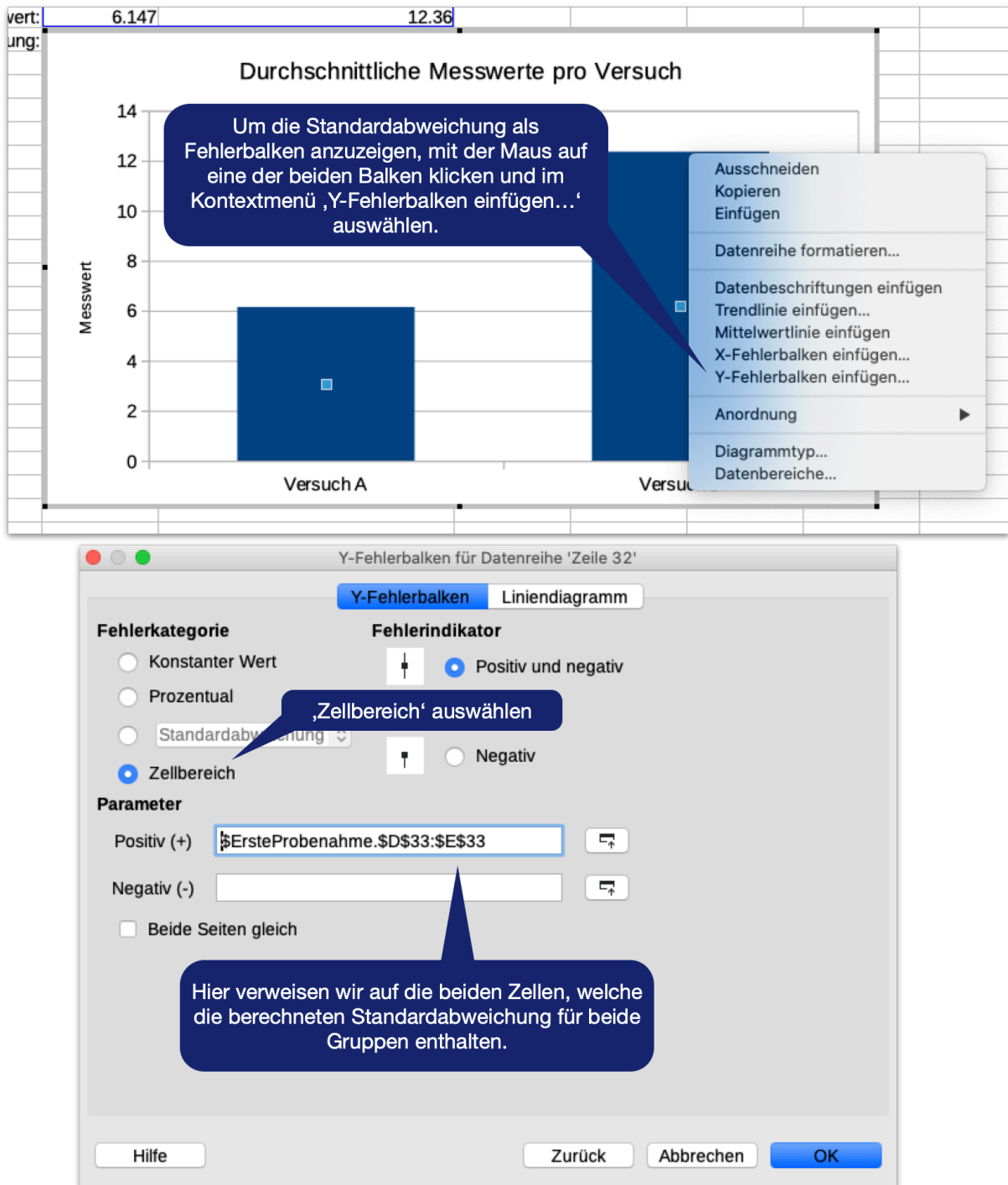


Abbildung 4.14: Hinzufügen von Fehlerbalken.

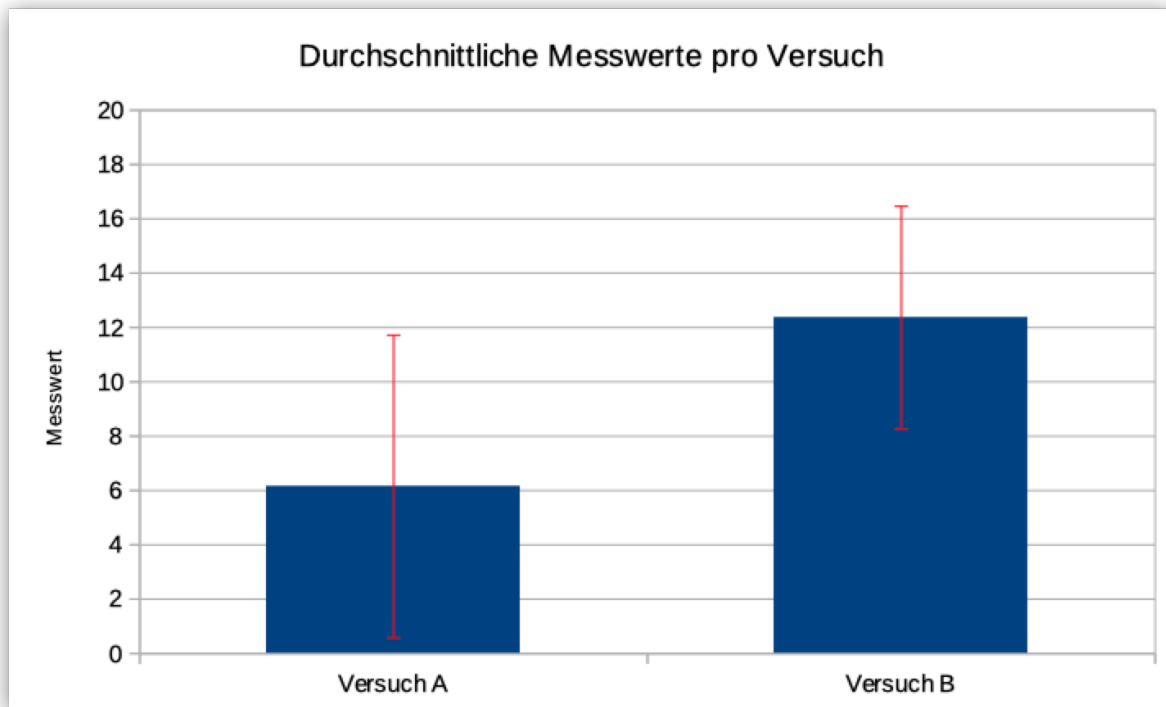


Abbildung 4.15: Vergleich der Messwert-Mittelwerte zwischen Versuch A und B. Die Standardabweichungen werden durch die Fehlerbalken repräsentiert.

ID	Versuch	Probenummer	Messwert	Anteil an der Gesamtsumme (in %)	Messwert-Klassen	Häufigkeit
1	A	1	10.3	10.3		4
2	A	2	3.8	3.8		8
3	A	3	4.9	4.9		12
4	A	4	5.3	5.3		16
5	A	5	9.0	9.0		20
6	A	1	4.3	4.3		24
7	A	2	2.8	2.8		
8	A	3	1.1	1.1		
9	A	4	0.6	0.6		
10	A	5	19.4	19.4		
11	B	1	20.1	20.1		
12	B	2	17.5	17.5		
13	B	3	12.3	12.3		
14	B	4	14.8	14.8		
15	B	5	9.6	9.6		
16	B	1	8.2	8.2		
17	B	2	11.4	11.4		
18	B	3	13.0	13.0		
19	B	4	7.8	7.8		
20	B	5	8.9	8.9		
Summe:			185.1			
Mittelwert:			9.3			
Standardabweichung:			5.7			

Callouts in the image:

- Klasse < 4
- Klasse >12 und ≤16
- Alle Werte liegen zwischen diesem Minimal- und Maximalwert. Entsprechend diesen Bereich in ≥5 gleich große Intervalle aufteilen (z.B. ≤4, ≤8, ≤12, ≤16, ≤20, >20).

Abbildung 4.16: Bildung von gleich großen Klassen basierend auf der Spannweite der Rohdaten, also der Messwerte.

Die (Roh-)Daten.

Die selbst definierten Klassen.

Keine schließende Klammer setzen!!

In Zelle H2 jetzt die Häufigkeitsfunktion verwenden.

shift + strg (bzw. cmd) + enter

Zellen H3-8 werden nach dieser Tastaturkombination automatisch gefüllt (auch für die Klasse >24, die wir gar nicht selbst definiert haben)

ID	Versuch	Probenummer	Messwert	Anteil an der Gesamtsumme (in %)	Messwert-Klassen	Häufigkeit
1	A	1	10.3	6	4	4
2	A	2	3.8	2	8	4
3	A	3	4.9	3	12	4
4	A	4	5.3	3	16	6
5	A	5	9.0	5	20	3
6	A	1	4.3	2	24	2
7	A	2	2.8	2		1
8	A	3	1.1	1		0
9	A	4	0.6	0		
10	A	5	19.4	10		
11	B	1	20.1	11		
12	B	2	17.5	9		
13	B	3	12.3			
14	B	4	14.8			
15	B	5	9.6			
16	B	1	8.2			
17	B	2	11.4			
18	B	3	13.0			
19	B	4	7.8			
20	B	5	8.9			
Summe:			185.1			
Mittelwert:			9.3			
Standardabweichung:			5.7			

Abbildung 4.17: Erstellung einer Häufigkeitsmatrix mithilfe der HÄUFIGKEIT() Funktion.

E	F	G	H	I	J	K	L	M
samtsumme (in %)		Messwert-Klassen	Häufigkeit					
6		4	4					
2		8	4					
3		12	6					
3		16	3					
5		20	2					
2		24	1					
2			0					
1								

Alle zu visualisierenden Zellen markieren INKLUSIVE der Spaltennamen und auf **Diagramm einfügen** klicken.

Das Standard-Säulendiagramm für das Histogramm verwenden.

Abbildung 4.18: Datenauswahl und Erstellung eines einfachen Säulendiagramms.

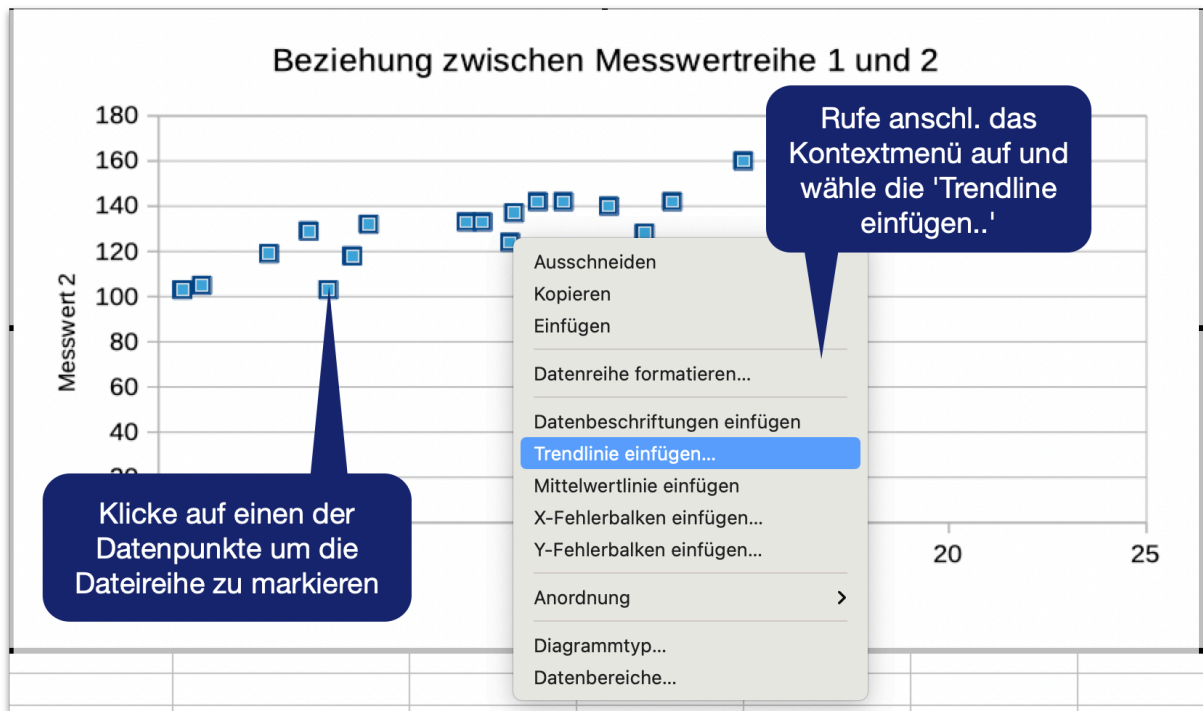


Abbildung 4.19: Auswahl des Datenbereichs und Angabe der Diagrammelemente.

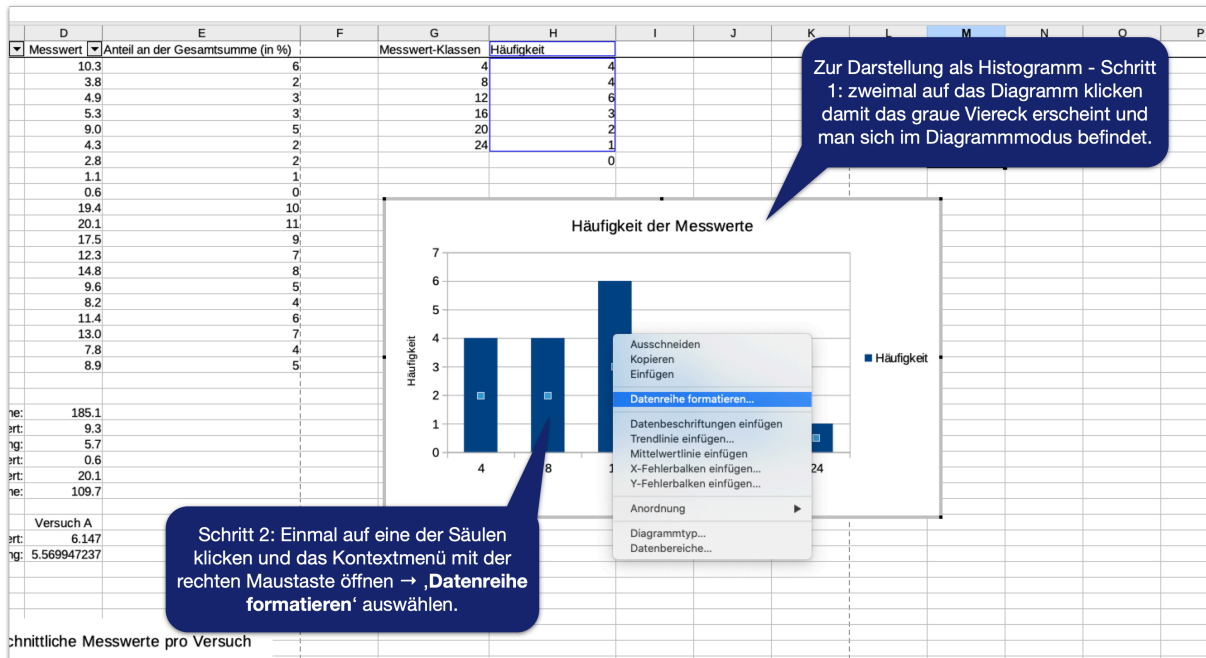


Abbildung 4.20: Formatierung der Datenreihe damit das Säulendiagramm einem Histogramm grafisch entspricht.

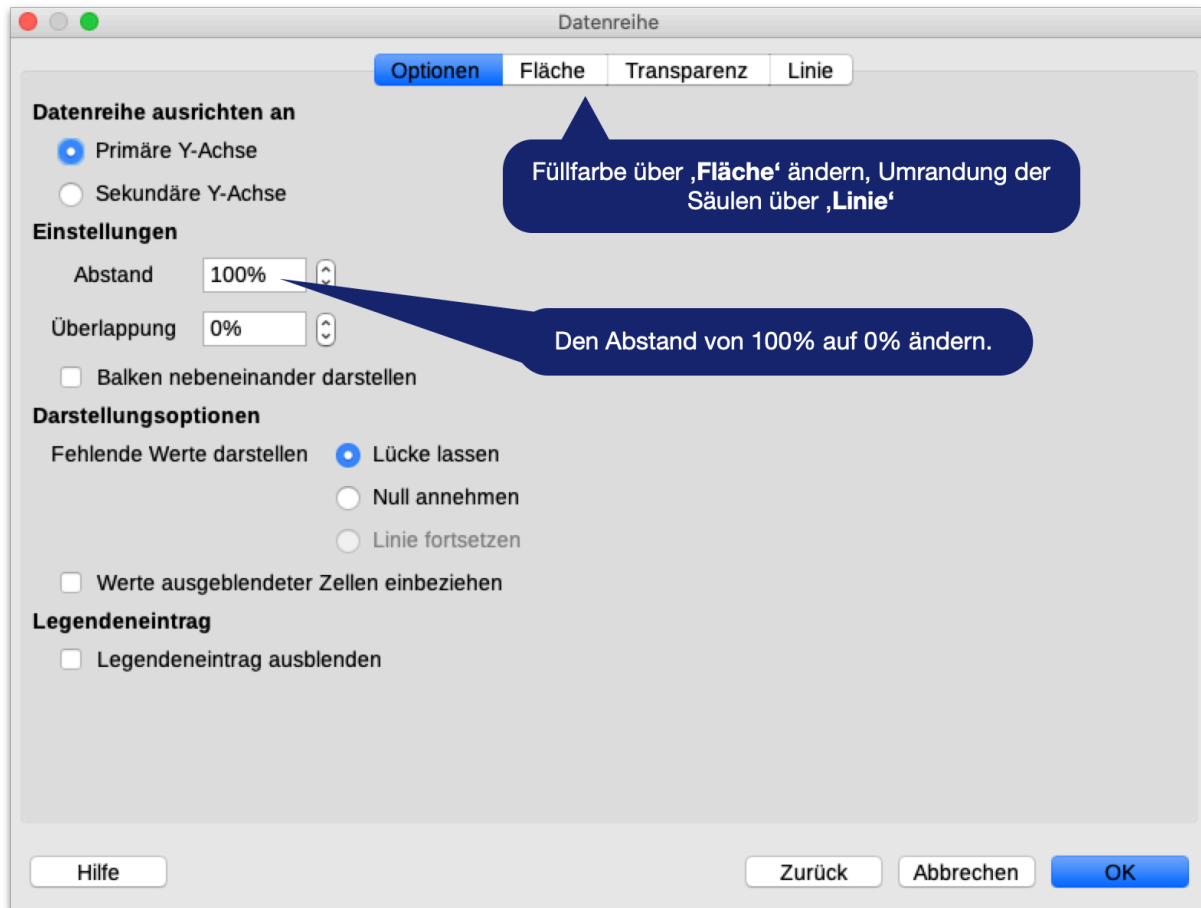


Abbildung 4.21: Reduzierung des Säulenabstands auf 0 Prozent und Änderung der Füllfarbe sowie Hinzufügen einer Umrandungslinie.

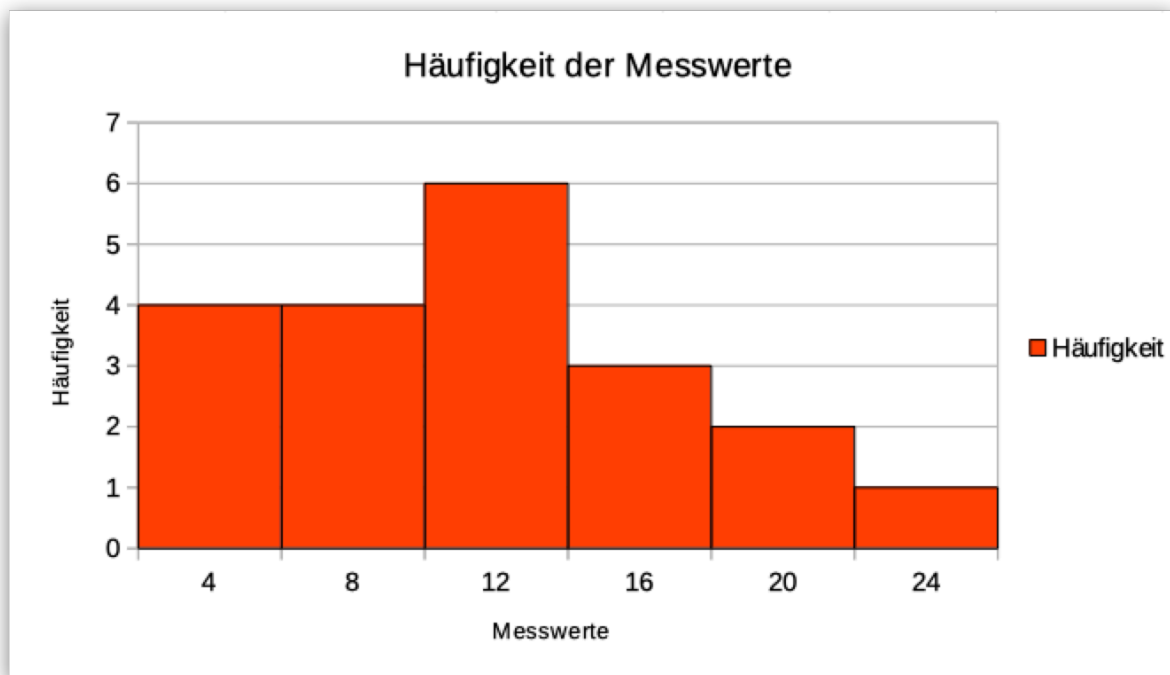


Abbildung 4.22: Histogramm zur Darstellung der Häufigkeitsverteilung der Messwerte (gepoolt über beide Versuche)

5 Tabellenexport zur weiteren Verwendung

5.1 Vorbereitungen zum Export

Auf folgende Punkte sollten Ihr besonders dann achten, wenn die Daten Eurer *Calc*-Tabelle in anderen Programmen weiter verwendet werden sollen:

- **Metainformationen** (z.B. Einheiten, Erklärungen zu Abkürzungen etc.) gehören grundsätzlich in ein **separates Tabellenblatt**.
- **Zwischenberechnungen** (Formeln), Grafiken etc. werden **nicht** mit den Daten zusammen in einer Datei exportiert.
- Keine leeren Spalten.
- Keine leeren Zeilen.
- Auch keine leeren Zellen (diese mit NA füllen = Not Available).
- Die erste Zeile enthält die Spaltennamen, ab Zeile 2 kommen die eigentlichen Werte der jeweiligen Spalte bzw. Variable.
- Spaltennamen sollten **keine Sonderzeichen** enthalten.
- Auch **Leerzeichen** sollten **nicht** in Spaltennamen vorkommen. Um dennoch Wörter voneinander zu trennen zu können, bieten sich folgende Möglichkeiten:
 - Bindestrich
 - Unterstrich
 - die sog. 'camelCase'-Schreibweise

5.2 Export von CSV Dateien

Um Daten zwischen Calc und R auszutauschen empfiehlt sich das CSV-Format. Es kann allerdings nur jeweils **ein Tabellenblatt** in einer CSV-Datei abgespeichert werden. Hast Du mehrere Tabellen in Deiner ODS-Datei, die alle in R importiert werden sollen, dann musst Du jede Tabelle separat exportieren. Das Speichern einer Tabelle als CSV-Datei ist ähnlich wie bei einer ODS-Datei, nur dass als Dateityp nicht *ODF-Tabellendokument (.ods)* gewählt werden darf, sondern *CSV (.csv)* (siehe auch **Dateispeicherung nicht vergessen**).

Nach der Speicherung passiert nichts weiter als das sich der Dateiname oben entsprechend dem Namen, der vergeben wurde, ändert. Das Problem ist jedoch, dass trotz Calc Warnung, dass nur die aktuelle Tabelle gespeichert wird, die anderen Tabellen weiterhin sichtbar sind. Man mag daher geneigt sein, an diesen später weitere Änderungen vorzunehmen. Diese Änderungen würden aber nicht mit gespeichert werden. Nach einem Schließen dieser CSV-Datei und erneutem Öffnen, wird nämlich nur noch die eigentlich gespeicherte Tabelle zu sehen sein.

Um einem solchen möglichen 'Missverständnis' aus dem Weg zu gehen, empfiehlt sich folgende Vorgehensweise, welche in den Abb. 5.1 bis Abb. 5.4 dargestellt ist.

Hier eine kurze Zusammenfassung:

- Die zu exportierende Tabelle in eine neue **Arbeitsmappe kopieren** und dann als CSV-Datei speichern. So hast Du weiterhin die Original ODS-Datei (oder Arbeitsmappe) geöffnet und die neue CSV-Datei.
- Wenn die Tabelle nicht ganz den Konventionen wie sie in **Vorbereitungen zum Export** beschrieben sind entspricht, d.h. es gibt noch Formeln, leere Zeilen und Spalten, evtl. Kommentare, ungeeignete Spaltennamen, etc., mache folgendes: kopiere nur die Zellen, die exportiert werden sollen (ohne evtl. deskriptive Statistiken, etc.) und füge den Inhalt (!) in ein weiteres Tabellenblatt (s. Abb. 5.1). Mache hier alle weitere nötigen Anpassungen und dann **verschiebe** die Tabelle in eine neue Arbeitsmappe (s. Abb. 5.2) und speichere diese als CSV (s. Abb. 5.3 und Abb. 5.4).

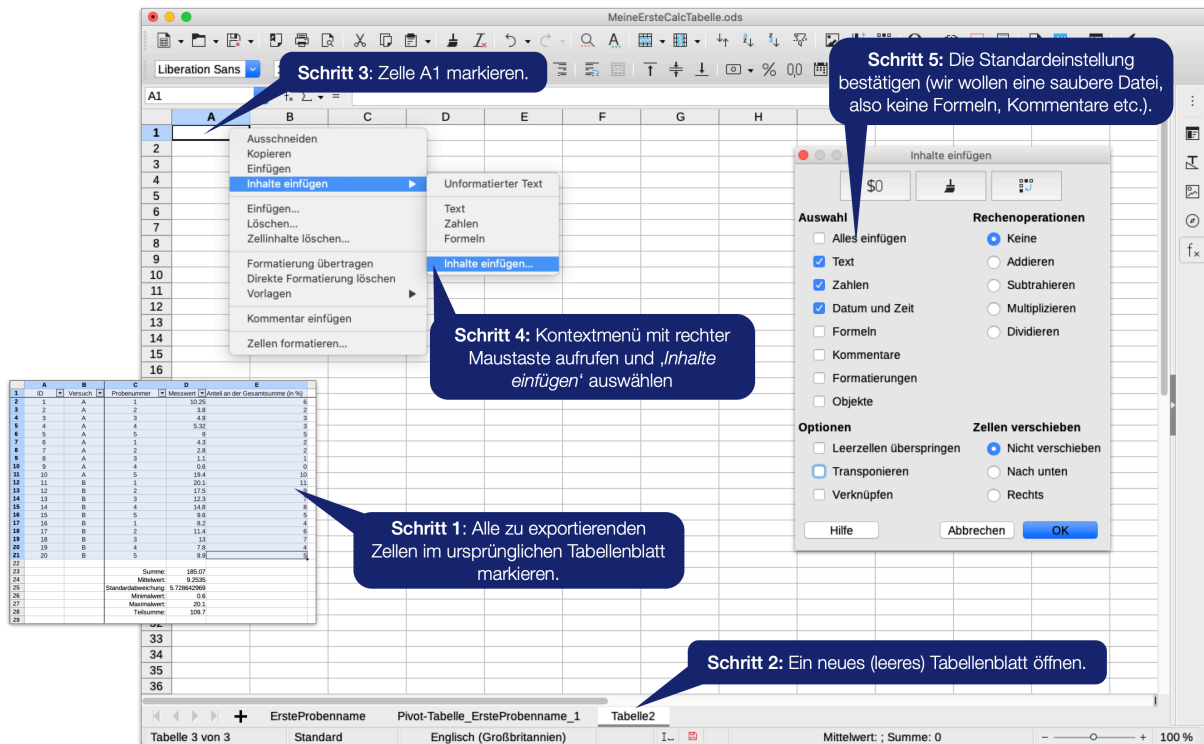


Abbildung 5.1: Datenvorbereitung für den Export.

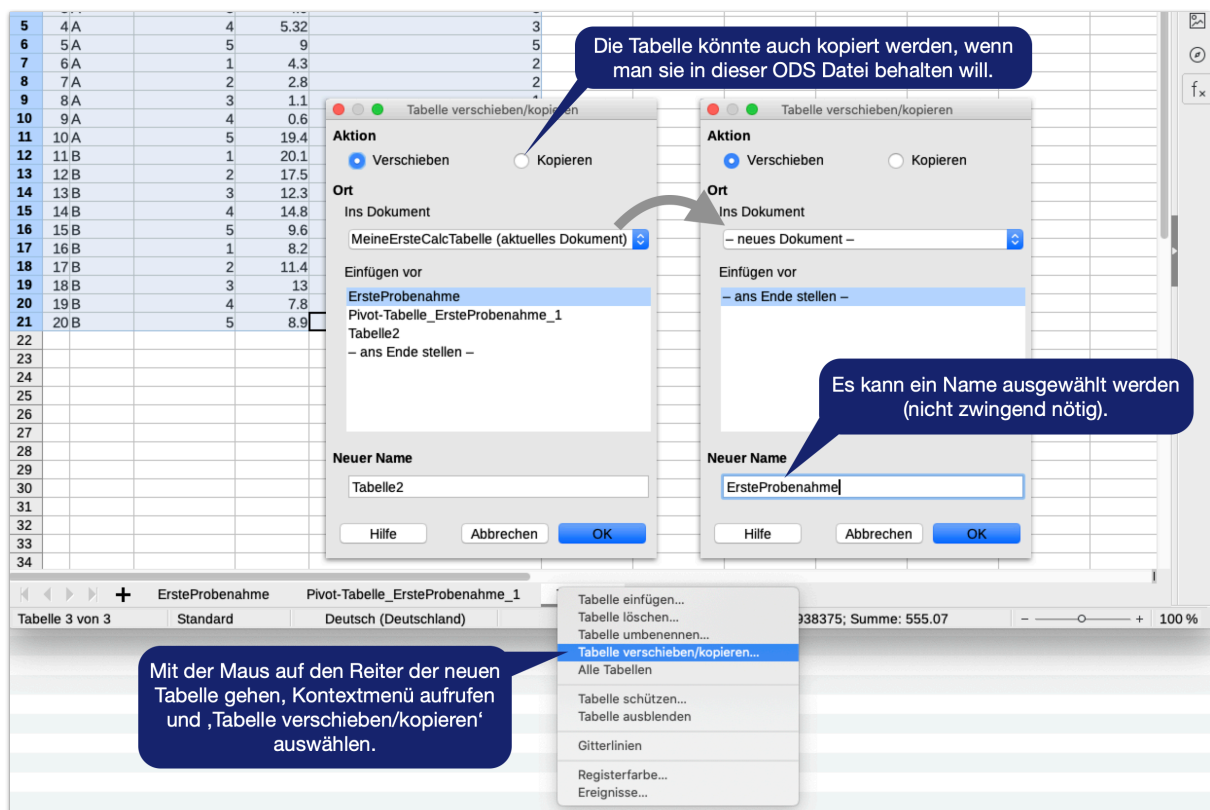


Abbildung 5.2: Tabelle in ein neues Dokument verschieben.

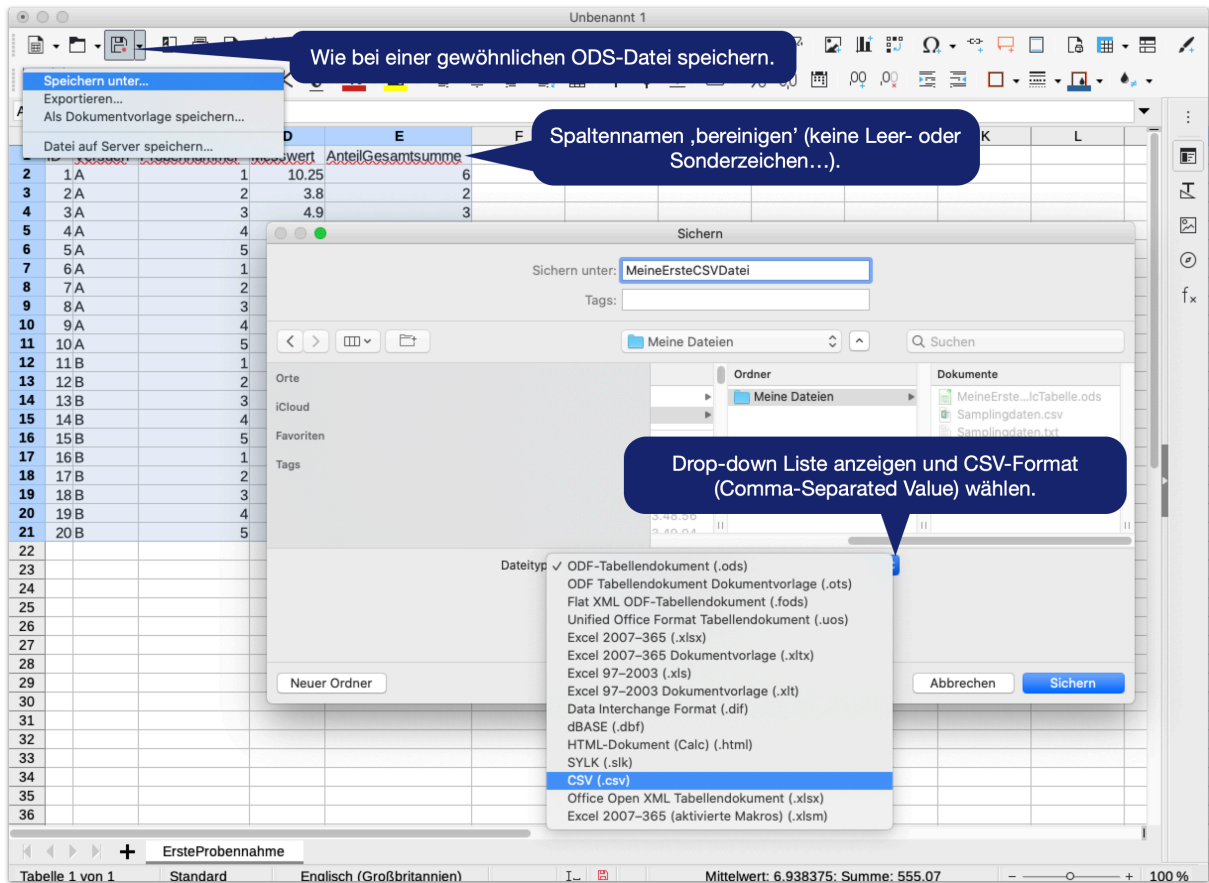


Abbildung 5.3: Tabelle im CSV-Format speichern.



Abbildung 5.4: Feldoptionen festlegen.

6 Problembehandlung

6.1 Wenn Calc Zahlen als Text erfasst

Manchmal kann es vorkommen, dass in Calc Zahlen zu sehen sind, die das Programm nicht als solche erkennt. Es können dann z.B. keine Rechenoperationen vorgenommen werden.

Die Lösung: Lege eine zusätzliche Spalte an, in der Calc die Werte mit Hilfe einer Funktion in echte Zahlen umwandelt. Um als Text gespeicherte Zahlen in echte Zahlen umzuwandeln, kannst Du die Funktion ZAHLOWERT() benutzen. Handelt es sich um ganze Zahlen, so brauchst Du nur den Zellbezug zu den als Text gespeicherten Werten einfügen. Enthalten die Werte Dezimalstellen oder Tausender-Trennzeichen, so müssen diese explizit definiert werden. Der Dezimaltrenner wird als zweiter Parameter in die Funktion eingefügt, der Tausender-Trenner als dritter. Ohne diese Parameter zeigt die Funktion einen Fehler an.

Möchtest Du absichtlich Zahlen als Text erfassen, dann stelle der Zahl bei der Eingabe ein Hochkomma ' voraus. Auf diese Weise kannst Du auch Telefonnummern oder Postleitzahlen mit führender Null korrekt darstellen.