

---

# **Statistik mit Excel 2013**

**Themen-Special**

W-EX2013S

Peter Wies

1. Ausgabe, Februar 2014



**HERDT**

## 3 Statistische Maßzahlen

### In diesem Kapitel erfahren Sie

- ✓ wie Sie Daten klassifizieren
- ✓ wie Sie Häufigkeiten berechnen und darstellen
- ✓ wie Sie Mittelwerte berechnen
- ✓ wie Sie Quantile berechnen
- ✓ wie Sie Streuungsmaße und Spannweiten berechnen

### Voraussetzungen

- ✓ Mit Formeln und Funktionen arbeiten
- ✓ Diagramme erstellen und bearbeiten

### 3.1 Allgemeines zu Häufigkeit und Klassenbildung

#### Klassifizierung von Merkmalen

Ihnen liegen die Daten einer Kundenbefragung, mit deren Hilfe die *Movement GmbH* den Zeitraum bis zum ersten Auftreten von Mängeln bei Stühlen ermitteln will, in Form einer ungeordneten Liste vor.

	A	B	C	D	E	F	G	H	I	J
1	<b>Kundenbefragung Fertigungslinie A: Stühle</b>									
2										
3	<b>Zeitraum bis zum Auftreten der ersten Mängel [in Jahren]</b>									
4	2,0	0,5	1,5	3,0	1,5	0,5	5,0	4,5	4,0	4,5
5	3,5	3,5	3,0	2,5	1,5	1,5	2,0	1,5	0,5	1,0
6	4,5	4,0	2,0	1,5	3,0	2,0	3,0	2,5	0,5	1,5
7	1,0	1,5	0,5	0,5	4,0	3,0	2,0	5,0	3,5	2,5
8	4,0	3,0	2,0	2,0	5,0	0,5	1,5	3,5	3,0	5,0

Da stetige Merkmale für statistische Auswertungen unübersichtlich zu verwalten sind, wird versucht, solche Merkmale mittels einer Klassenbildung in diskrete Merkmale umzuwandeln. Zur Auswertung der Kundenbefragung ist eine Einteilung der Zeiträume in **Klassen** sinnvoll.

Bei der Klassenbildung wird der Wertebereich, in dem alle Messwerte enthalten sind, in mehrere Teilräume (Intervalle) zerlegt. Die so entstandenen Klassen haben definierte Parameter, wie z. B. die Klassenmitte, die Klassenbreite und die untere bzw. obere Klassengrenze. Im abgebildeten Beispiel ist die Klassenbreite aller Klassen gleich. Eine Einteilung in gleich große Klassenbreiten ist sinnvoll, da sich daraus rechnerische Vereinfachungen ergeben.

Klassen		
0,0	bis	1,0
>1,0	bis	2,0
>2,0	bis	3,0
>3,0	bis	4,0
>4,0	bis	5,0

In manchen Häufigkeitsanalysen kann es jedoch der besseren Darstellung dienen, unterschiedlich große Klassen zu wählen. Ein Beispiel dafür wäre eine statistische Auswertung von Alterswerten, die Sie in die Klassen Säugling, Kind und Erwachsener einteilen, obwohl die drei Altersklassen verschieden große Zeitspannen abdecken.

## Absolute und relative Häufigkeiten

Zur Charakterisierung von Häufigkeitsverteilungen können Sie einsetzen:

- ✓ Lokationsmaße, wie zum Beispiel Mittel-, Maximal- und Minimalwerte;
- ✓ Dispersionsmaße (Streuungsmaße), wie zum Beispiel Standardabweichung, Varianz.

Die der statistischen Auswertung zugrunde liegenden Daten sind meist umfangreich und unübersichtlich. Deshalb ist es wesentlich, die entsprechenden Daten sinnvoll zusammenzufassen, beispielsweise durch die Berechnung von Häufigkeiten.

Häufigkeiten werden in **absolute** und **relative** Häufigkeiten unterschieden. Wenn das Merkmal mit  $X$  bezeichnet wird, dessen Ausprägungen mit  $x_1 \dots x_k$  und die Anzahl der Untersuchungseinheiten mit  $n$ , so gelten für die absolute und die relative Häufigkeit folgende Definitionen:

- ✓ **Absolute Häufigkeit** ( $H_n$ ):  $H_n(x_j) =$  Anzahl der auftretenden Untersuchungseinheiten mit Ausprägung  $x_j$

Die absolute Häufigkeit hängt damit stark von der Anzahl der Untersuchungseinheiten ab. Vergleiche von Häufigkeitsverteilungen mit unterschiedlicher Anzahl an Untersuchungseinheiten sind daher nicht mithilfe der absoluten Häufigkeit durchführbar. Zum Vergleich solcher Häufigkeitsverteilungen eignet sich nur eine Größe, die unabhängig von der Anzahl der Untersuchungseinheiten ( $n$ ) ist, nämlich die:

- ✓ **Relative Häufigkeit** ( $h_n$ ):  $h_n(x_j) = \frac{1}{n} H_n(x_j)$

Die relative Häufigkeit bezeichnet also den prozentualen Anteil der Untersuchungseinheiten, welche die Ausprägung  $x_j$  tragen.

## Häufigkeiten grafisch darstellen

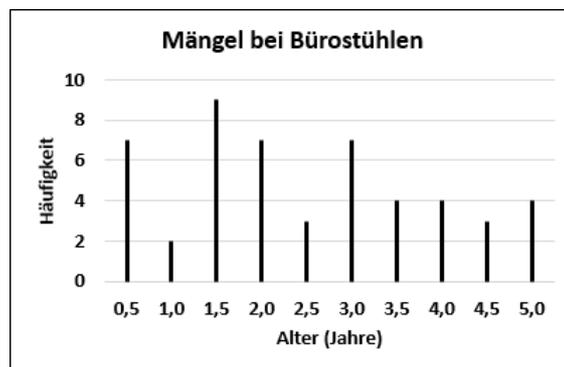
Jede Art von Zahlenmaterial wird übersichtlicher, wenn es grafisch aufbereitet wird. Zur Darstellung von Häufigkeiten werden vorwiegend zwei Diagrammtypen eingesetzt.

- ✓ **Stabdiagramme** zur Darstellung unklassierter Werte
- ✓ **Histogramme** zur Darstellung klassierter Werte

### Stabdiagramme

Bei Stabdiagrammen werden auf der Kategorieachse (x-Achse bzw. Abszissenachse) die Ausprägungen abgebildet, während die Wertachse (y-Achse bzw. Ordinatenachse) die entsprechenden Häufigkeiten zeigt.

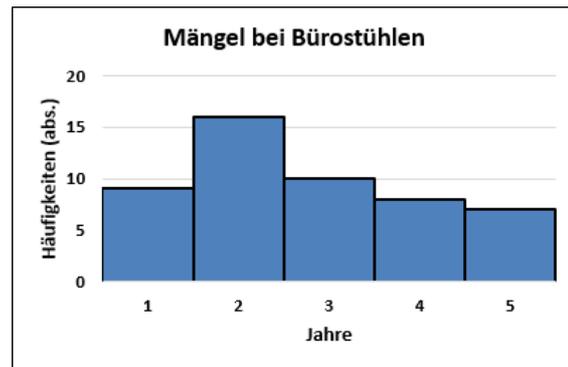
Stabdiagramme werden besonders bei der Darstellung diskreter Merkmale eingesetzt.



Stabdiagramme können z. B. Säulen- oder Balkendiagramme sein. Daneben können Sie ein Stabdiagramm auch (wie in der Abbildung oben) als Liniendiagramm erstellen. Dabei weisen Sie der entsprechenden Datenreihe sogenannte Bezugslinien zu (Register *ENTWURF*, Gruppe *Diagrammlayouts*, *Diagrammelement hinzufügen - Linien - Bezugslinien*) und blenden die Linie der Datenreihe aus (Register *FORMAT*, Gruppe *Formenarten*, *Formkontur - Kein Rahmen*).

## Histogramme

Histogramme werden zur Darstellung stetiger Merkmale eingesetzt (vgl. Abschnitt 3.3). Diese stetigen Werte werden klassifiziert und die Klassen mit der jeweiligen Klassenbreite auf der Kategorieachse dargestellt. Die Wertachse zeigt die absoluten bzw. relativen Häufigkeiten, wenn alle Klassen gleich breit sind. Auf der Diagrammfläche werden Rechteckflächen abgetragen. Deren Fläche ist proportional der Klassenhäufigkeit.



## 3.2 Häufigkeiten in Excel berechnen

<b>Beispieldateien</b>	<i>Kundenbefragung, Kundenbefragung-E</i>
------------------------	---

### Absolute Häufigkeiten mit der Funktion HÄUFIGKEIT berechnen

Excel bietet zur Berechnung der absoluten Häufigkeit die Funktion HÄUFIGKEIT. So können Sie Ereignisse innerhalb verschiedener Bereiche zählen. Die Funktion HÄUFIGKEIT besitzt folgende Syntax:

<b>HÄUFIGKEIT</b> ( <i>Daten; Klassen</i> )
---

- ✓ Das Argument **Daten** bezeichnet die Gesamtheit der zu analysierenden Daten.
- ✓ Das Argument **Klassen** bezeichnet die Obergrenzen der Klassenintervalle.



Die Funktion HÄUFIGKEIT muss als Matrixformel eingegeben werden (vgl. folgende Erläuterungen).

### Was sind Matrixformeln?

Mit Matrixformeln können Sie in Excel mehrere Datenbereiche zur Berechnung kombinieren. Als Matrix (Plural: Matrizen) bezeichnet man einen rechteckigen Zellbereich, für den bestimmte Rechnungen ausgeführt werden sollen. Mit einer Matrixformel können Sie für unterschiedliche Matrizen mehrere Berechnungen auf einmal durchführen. So müssen Sie nicht für jede dieser Berechnungen eine separate Formel eingeben.

- ✓ Matrixformeln werden von geschweiften Klammern {} umschlossen. Darüber hinaus besitzen Matrixformeln keine spezielle Syntax.
- ✓ Alle in Matrixformeln verwendeten Matrizen (Matrixargumente) müssen **dieselbe** Anzahl von Zeilen und Spalten aufweisen.
- ✓ Matrixformeln berechnen jeweils zuerst das erste Element der Matrixargumente, anschließend jeweils das zweite Element usw.
- ✓ Matrixformeln liefern je nach Anwendung entweder ein Ergebnis in einer Zelle oder unterschiedliche Ergebnisse in mehreren Zellen.

Die Eingabe von Matrixformeln ähnelt der Eingabe von Standardformeln. Achten Sie jedoch darauf, dass die einzelnen Argumente dieselbe Spalten- und Zeilenzahl aufweisen. Die Eingabe einer Matrixformel schließen Sie mit der Tastenkombination **Strg** **↵** **↵** ab.



In Excel werden Matrixformeln auch als Arrayformeln bezeichnet.

