

Stefan Einbock

STATISTIK für Ingenieure und Naturwissenschaftler (mit Excel)

endlich verständlich!

Ihre Vorteile:

- + 26 gratis Excel-Tools
- + Erklärende Beispiele
- + Klare Sprache

Auf den Punkt:

- Stichprobenumfänge festlegen (Power)
 - Versuche mit statistischen Tests auswerten und interpretieren
 - Messsystemanalysen durchführen
 - Ausreißer finden und bewerten
 - Vertrauensbereiche verstehen
-

Dr.-Ing. Stefan Einbock



studierte allgemeinen Maschinenbau an der Hochschule Esslingen und promovierte an der TU Dresden im Bereich der Betriebsfestigkeit.

Das theoretische Wissen zur Betriebsfestigkeit, Statistik und Zuverlässigkeit vermittelt er in Kooperation mit dem Verein deutscher Ingenieure (VDI) als erfolgreicher Seminarleiter der Seminare „Betriebsfestigkeitsberechnung“ sowie „Bauteile robust auslegen und effizient erproben“. Außerdem hält er regelmäßig Vorträge an Hochschulen.

Zusätzlich ist er Autor mehrerer Bücher zur Auslegung von Bauteilen und deren statistischer und experimenteller Absicherung.

Bei der Robert Bosch GmbH leitet er im Geschäftsbereich Powertrain Systems das Kompetenzzentrum für Metalle.

Stefan Einbock

STATISTIK FÜR INGENIEURE (MIT EXCEL)

Datenauswertung schnell verstehen & anwenden

Mit

330 Seiten,

106 Grafiken,

26 Excel-Tools und

begleitendem Blog <http://einbock-akademie.de/blog>



Bibliografische Informationen der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliographische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© 2018 Stefan Einbock, <http://einbock-akademie.de>

Herstellung und Verlag: **BoD – Books on Demand**, Norderstedt.

ISBN: 9783752837704

Statistik ist eine Zusammenfassung von Methoden, welche uns erlauben vernünftige Entscheidungen im Falle von Ungewissheit zu treffen.

W. Allen Wallis; Harry V. Roberts (1956)

1 EINFÜHRUNG

Um es sofort vorweg zu sagen: ich bin ein großer Freund statistischer Methoden. Ich bin überzeugt, dass mit Hilfe der Statistik Daten deutlich intensiver ausgewertet und interpretiert werden können und dass die Methoden von jedem richtig angewandt werden können. Zusätzlich lassen sich häufig auch Versuche, wenn sie statistisch richtig geplant sind deutlich effizienter durchführen. Außerdem behaupte ich, dass Statistik auch viel Spaß macht. Warum? Einfach weil man damit viele neue Dinge entdecken kann, und das alleine macht ja schon Spaß (schon für kleine Kinder ist das Erforschen Ihrer Umwelt sehr spannend!). Mit diesem Buch möchte ich Ihnen genau diese Möglichkeiten an die Hand geben. Es richtet sich deswegen an

- Interessierte Einsteiger.
- Ingenieure, Naturwissenschaftler und Studenten aus den technischen Berufen.
- alle die wenig Zeit für die Einarbeitung haben.
- alle, die schnell Daten auswerten oder statistische Methoden verstehen wollen.

Natürlich kenne ich auch die Vorurteile gegenüber der Statistik. Vielfach höre ich Sätze wie „traue keiner Statistik die du nicht selbst gefälscht hast!“. Hier zeigt sich die Befürchtung, Daten falsch auszuwerten und damit keine vernünftigen Aussagen treffen zu können.

Mühsam und zeitintensiv habe ich mich in die Statistik eingearbeitet und dieses Wissen in zahlreichen Seminaren der EinbockAKADEMIE an Anwender und Einsteiger weitergegeben. Deswegen ist dieses Buch entstanden, das

- sich auf die wichtigsten Methoden der Statistik für Ingenieure und Naturwissenschaftler fokussiert,
- Hinweise für eine richtige und einfache Anwendung der Methoden enthält,
- eine schnelle Einarbeitung bietet und verständlich geschrieben ist,
- sich sehr stark an der einfach bedienbaren Software Excel orientiert,
- zahlreiche Praxistipps enthält, die Ihnen den „Sparringspartner“ ersetzen sollen
- und für eine einfache Anwendung praktische Excel-Tools bereitstellt.

Kurz: das Ihnen ermöglicht die Statistik schnell zu verstehen und anzuwenden.

Um eine **schnelle Einarbeitung** zu gewährleisten, fokussiert dieses Buch auf die wichtigsten Methoden und beschränkt sich auf die absolut notwendige Mathematik. Zusätzlich finden Sie noch hilfreiche Tipps und Erfahrungen zu einer deutlichen Steigerung der Lerneffizienz. Für einen extrem schnellen Einstieg haben wir die wichtigsten Kapitel mit  versehen. Wenn Sie sich auf diese konzentrieren, ist dies der schnellst mögliche Einstieg. Mit Hilfe der Assistenten können Sie außerdem direkt die richtige Methode für Ihre Fragestellung finden.

Für eine **verständliche Vermittlung** des Inhaltes werden viele Abbildungen genutzt, die das Geschriebene untermalen. Zusätzlich werden komplizierte Sachverhalte durch praxisrelevante Beispiele erklärt. Es wird außerdem bewusst eine einfache, klare Sprache verwendet (der berufliche Alltag ist kompliziert genug). Zur Festigung des Verständnisses wird die Theorie zusätzlich durch umfangreiche praxisnahe Beispiele ergänzt. Jedes Kapitel schließt mit einer

kurzen Zusammenfassung. Links sind auch als QR Code eingefügt, den Sie bequem mit dem Handy fotografieren und darauf zugreifen können. Zusätzlich werden wichtige Aussagen oder Formeln häufiger wiederholt. Das hilft dem Verständnis und dem Lesefluss.

Zur **einfachen Anwendung** der Methoden finden Sie nützliche Excel-Tools und am Ende eines jeden Kapitels finden Sie die konkrete Vorgehensweise in Excel. Sie können diese Tools hier

http://einbock-akademie.de/download/buch_statistik

herunterladen. Das Passwort finden Sie in der Fußnote von Seite 177. Für eine sichere Anwendung werden für jede Methode die einzuhaltenden Randbedingungen und mögliche Risiken übersichtlich angegeben. Mit Hilfe der Assistenten am Anfang des Buches und in dem jeweiligen Kapitel gelingt ihnen sehr schnell der Einstieg und sie können zielsicher die richtigen Methoden auswählen. Da die Datenauswertung oftmals mit Excel geschieht, werden für die wichtigsten Gleichungen die Excel-Formeln in folgender Form angegeben:

EXCEL: = MITTELWERT($x_1; x_2; \dots x_n$).

Dieses Buch liefert Ihnen somit

- einen effizienten Einstieg in die Statistik,
- die Möglichkeit Daten selbständig zu planen und fachmännisch auszuwerten,
- einen selbständigen, berufsbegleitenden Einstieg in die Statistik,
- praxisorientierte Übungen zur Vertiefung des Gelernten.

Ich wünsche Ihnen genauso viel Freude beim Lesen und Anwenden der Methoden, wie ich sie beim Schreiben hatte und bin auf Ihre Rückmeldungen gespannt!

Stefan Einbock

Sommer 2018, Stuttgart

2 FEEDBACK WILLKOMMEN!

Da dieses Buch von Ingenieuren für Ingenieure geschrieben ist, möchte ich es gerne in Diskussion mit Ihnen weiterentwickeln.

Dieses Buch gefällt Ihnen? Dann freue ich mich auf eine ehrliche Rückmeldung auf www.amazon.de oder schreiben Sie mir eine Email an kontakt@einbock-akademie.de.

Haben Sie einen Fehler gefunden, der sich trotz größtmöglicher Sorgfalt eingeschlichen hat? Oder möchten Sie Feedback geben? Dann freue ich mich ebenfalls über eine kurze Email.

Fallen Ihnen weitere Themen ein, die Sie außerdem gerne in einer künftigen Auflage behandelt hätten? Bitte senden Sie einfach Ihre Themenwünsche per Email an mich. Ich werde diese sammeln und evtl. werde ich diese dann in meinem Blog veröffentlichen:

www.einbock-akademie.de/blog

Ich freue mich von Ihnen zu hören!

Stefan Einbock



INHALTSVERZEICHNIS

1	Einführung.....	6
2	Feedback willkommen!	7
3	Tipps: effiziente Einarbeitung ✂	14
4	Überblick ✂	17
5	Nützliche Helfer: Die Assistenten ✂	19
6	Einführung in Excel	23
6.1	Umgang mit Daten ✂	24
6.2	Daten einfügen.....	27
6.3	Mit Daten rechnen (Formeln)	28
6.4	Das geheime Excel-Tool ✂	30
6.5	Auf den Punkt	31
Teil 1: Grundlagen		32
7	Wichtige Begriffe und Merkmale ✂	33
8	Beschreibung der mittleren Werte und Streuungen ✂	37
8.1	Der Mittelwert (Lage).....	37
8.2	Der Median (Lage).....	38
8.3	Die Spannweite (Streuung).....	40
8.4	Die Varianz (Streuung).....	40
8.5	Die Standardabweichung (Streuung)	41
8.6	Der Variationskoeffizient (Streuung).....	42
8.7	Die Quartile (Streuung)	42
8.8	Auf den Punkt	44
8.9	Arbeiten mit Excel ✂	45
9	Wichtige Grafiken ✂	47
9.1	Das Histogramm und die Dichtefunktion	47
9.2	Das Wahrscheinlichkeitsnetz	50
9.3	Der Box-Whisker Plot	57
9.4	Auf den Punkt	59

9.5	Arbeiten mit Excel 	60
9.5.1	Das Histogramm-Tool	60
9.5.2	Das Wahrscheinlichkeitsnetz-Tool	61
9.5.3	Das Box-Whisker-Plot-Tool	62
10	Wichtige statistische Verteilungen 	63
10.1	Die Normalverteilung	63
10.2	Die logarithmische Normalverteilung	68
10.3	Die Weibullverteilung	70
10.4	Auf den Punkt	73
11	Der Vertrauensbereich 	74
11.1	Der Vertrauensbereich von Verteilungen	76
11.2	Der Vertrauensbereich des Mittelwertes	79
11.3	Der Vertrauensbereich der Standardabweichung	81
11.4	Der Vertrauensbereich des Weibullexponenten b	82
11.5	Der Vertrauensbereich des Skalenparameters λ	82
11.6	Auf den Punkt	82
11.7	Wichtige Formeln	83
11.8	Arbeiten mit Excel 	85
Teil 2: Daten erheben		86
12	Stichproben erheben	87
12.1	Warum sollten die Daten ermittelt werden?	88
12.2	Welche Daten / Messgrößen sollen erfasst werden?	89
12.3	Wie werden die zu erwartenden Streuungen abgedeckt? 	89
12.4	Mit welchem Messsystem werden diese Daten gemessen?	91
12.5	Wie sollen die Daten dargestellt werden?	91
12.6	Datenerhebungsplan 	92
12.7	Auf den Punkt	92
13	Messsystemanalyse (MSA)	93
13.1	Assistent einer Messsystemanalyse 	94
13.2	Grundlagen	94
13.3	Schritt 1: Bewertung der Auflösung des Messgerätes	97

13.4	Schritt 2: Funktionsprobe des Messgerätes.....	97
13.5	Schritt 3: Verfahren 1 (Wiederholpräzision)	97
13.6	Schritt 4: Verfahren 2 (Vergleichspräzision)	101
13.7	Schritt 5: Verfahren 3 (Vergleichspräzision)	108
13.8	Interpretation von Messsystemanalysen	108
13.9	Auf den Punkt	110
13.10	Arbeiten mit Excel ✂	111
13.11	Wichtige Formeln	115
Teil 3: Daten auswerten und interpretieren		116
14	Zusammenhänge von Daten finden (Regression und Korrelation)	117
14.1	Assistent Analyse von Zusammenhängen ✂	117
14.2	Grafische Auswertungen und ihr Nutzen	118
14.3	Modellrechnungen und ihre Ziele	120
14.3.1	Korrelationsanalyse oder: gibt es Zusammenhänge?	120
14.3.2	Regressionsanalyse, oder: wie ist der Zusammenhang?	129
14.3.2.1	Die Methode der kleinsten Quadrate	134
14.3.2.2	Maximum Likelihood Methode für unzensierte Daten	139
14.3.2.3	Maximum Likelihood Methode für zensierte Daten	142
14.3.3	Vertrauensbereiche der Regressionsmodelle.....	145
14.4	Auf den Punkt	152
14.5	Arbeiten mit Excel ✂	153
14.6	Wichtige Formeln	155
15	Unterschiede untersuchen (statistische Tests)	157
15.1	Assistent für statistische Tests ✂	160
15.2	Grundlagen statistischer Tests ✂	162
15.2.1	Nullhypothese und Alternativhypothese	163
15.2.2	Fehler erster Art / Signifikanz	165
15.2.3	Fehler zweiter Art / Power	166
15.3	Test auf Verteilungen.....	168
15.3.1	Kolmogorow-Smirnow-Test	169
15.3.2	Anderson-Darling-Test	175

15.3.2.1	Test auf Normal- und logarithmische Normalverteilung	176
15.3.2.2	Test auf Weibullverteilung	182
15.3.3	Arbeiten mit Excel ✂	187
15.4	Test auf Streuungen für 2 Stichproben	189
15.4.1	F-Test für normalverteilte Daten	190
15.4.2	Siegel-Tukey Test für nicht normalverteilte Daten	194
15.4.3	Arbeiten mit Excel ✂	199
15.5	Test auf Streuungen für > 2 Stichproben	202
15.5.1	Bartlett-Test für normalverteilte Daten	202
15.5.2	Levene-Test für nicht normalverteilte Daten	205
15.5.3	Arbeiten mit Excel ✂	210
15.6	Mittelwerttests für ≤ 2 Stichproben	212
15.6.1	Der t-Test (normalverteilte Daten)	213
15.6.1.1	Der Einstichproben t-Test	218
15.6.1.2	Zweistichproben t-Test für gleiche Varianzen	222
15.6.1.3	Zweistichproben t-Test für ungleiche Varianzen nach Welch	224
15.6.1.4	Zweistichproben t-Test für gepaarte Daten	225
15.6.1.5	Test des Korrelationskoeffizienten	229
15.6.2	Wilcoxon-Test (nicht normalverteilte Daten)	230
15.6.3	Arbeiten mit Excel ✂	234
15.7	Mittelwerttests für mehr als 2 Stichproben	239
15.7.1	Varianz-Analyse (ANOVA) für normalverteilte Daten	239
15.7.2	Kruskal-Wallis-Test für nicht normalverteilte Daten	245
15.7.3	Arbeiten mit Excel ✂	251
15.8	Erfahrungen aus der Praxis	254
15.9	Auf den Punkt	255
16	Ausreißer bewerten	257
16.1	Der Assistent für Ausreißertests ✂	258
16.2	Robuste Kennwerte gegen Ausreißer	259
16.3	Grafischer Test: der Box-Whisker-Plot	259
16.4	Tests für normalverteilte Daten	263

16.4.1	Dean-Dixon-Test für kleine Stichproben	263
16.4.2	Pearson-Test für größere Stichproben.....	266
16.5	Walsh-Test für nicht normalverteilte Daten.....	269
16.6	Umgang mit Ausreißern.....	271
16.7	Auf den Punkt	274
16.8	Arbeiten mit Excel ✎	275
17	Poweranalyse: Oder wie groß muss die Stichprobe sein?	279
17.1	Grundlagen der Poweranalyse ✎	279
17.2	Poweranalyse für den t-Test.....	280
17.2.1	Berechnung der Power	280
17.2.2	Berechnung der Stichprobengröße	285
17.2.3	Arbeiten mit Excel ✎	288
17.3	Poweranalyse für den F-Test	290
17.3.1	Berechnung der Power	290
17.3.2	Berechnung der Stichprobengröße	290
17.3.3	Arbeiten mit Excel ✎	291
17.4	Ausblick auf weitere Poweranalysen	293
Teil 4:	Daten präsentieren	294
18	Daten präsentieren.....	295
18.1	Die richtige Einstellung ✎	295
18.2	Daten präsentieren	297
18.3	Mit Statistik tricksen.....	299
18.3.1	Absolute anstelle von relativen Zahlen verwenden.....	299
18.3.2	Nullpunkt bei Achsen unterdrücken.....	300
18.3.3	Daten weglassen	302
18.3.4	Stichprobenanzahl sehr klein wählen.....	303
18.3.5	Nur einen Teil der Wahrheit sagen	303
18.3.6	Falsche Signifikanzniveaus wählen	304
18.4	Auf den Punkt	308
19	Englische Begriffe.....	309
19.1	Deutsch-Englisch	309

19.2	Englisch-Deutsch	314
20	Anhang	320
20.1	Quantile der t-Verteilung	320
20.2	Quantile der F-Verteilung	321
20.3	Quantile der χ^2 -Verteilung	322
20.4	Schranken der Normalverteilung	323
20.5	Vertrauensbereiche	325
20.6	Wahrscheinlichkeitsnetze	327
20.6.1	Normalverteilung	327
20.6.2	Logarithmische Normalverteilung	328
21	Literaturverzeichnis	329
22	Stichwortverzeichnis	332

3 TIPPS: EFFIZIENTE EINARBEITUNG

Für viele wird die Einarbeitung in die Betriebsfestigkeit parallel zur Arbeit erfolgen. Deshalb ist es wichtig, diese so effizient wie möglich zu gestalten. Die Kapazität, welche zur Bearbeitung der Aufgaben zur Verfügung steht, kann man sich als ein leeres Glas vorstellen. Die Aufgaben kann man sich als Kugeln denken, wobei die Größe der Kugeln den Aufwand für die Aufgaben darstellt (siehe Abbildung 2-1).

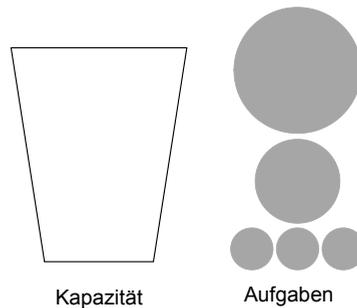


Abbildung 3-1: Kapazität vs. Aufgaben

Größtmögliche Effizienz wird erreicht, wenn möglichst viele Aufgaben innerhalb eines Tages erledigt werden können. Es ist aus Effizienzgründen sinnvoll eine einmal begonnene Aufgabe auch abzuschließen. Es macht beispielsweise keinen Sinn bei einer Einarbeitung immer nur eine halbe Seite des Buches zu lesen und dann am nächsten Tag weiterzumachen. Dadurch können Aufgaben nicht beliebig klein werden.

Die kleineren Aufgaben repräsentieren beispielsweise die Bearbeitung von Emails, den Austausch mit Kollegen oder die schnelle, kurzfristige Beantwortung von Fragen. Die großen Aufgaben sind beispielsweise das Einarbeiten in die Betriebsfestigkeit, oder die Auslegung eines Bauteiles. Aus Effizienzgründen ist es wichtig die Aufgaben in der richtigen Reihenfolge zu bearbeiten (siehe Abbildung 2-2).

Werden zuerst die vielen kleinen Aufgaben erledigt, (Variante 1 aus Abbildung 2-2), dann bleibt am Ende des Tages nicht genügend Zeit, um die großen Aufgaben zu erledigen. Beginnt man dagegen mit den großen Aufgaben zuerst und lässt sich durch die kleinen nicht ablenken, dann ist ausreichend Zeit vorhanden. Im schlimmsten Fall muss dann eine der kleineren Aufgaben auf den nächsten Tag geschoben werden.

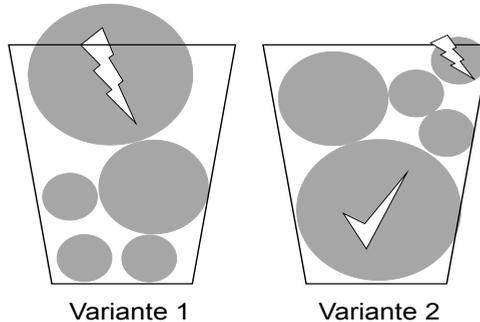


Abbildung 3-2: Einfluss der Reihenfolge bei der Aufgabenbearbeitung

Aus dieser Idee heraus ergeben sich sieben überraschend einfachen Tipps, wie Sie Ihre Lerneffizienz deutlich steigern können!

1. Legen Sie Ziele fest

Setzen Sie sich konkrete Ziele, die Ihnen einen direkten Nutzen bringen. Ideal ist es, wenn Sie z. B. formulieren: „Für das Bauteil meiner Konstruktion werde ich zum xx.xx.xxxx die Sicherheitsfaktoren berechnen.“ Halten Sie diese Ziele schriftlich mit einem Zieltermin fest. Dies gibt Ihnen einen Fokus und motiviert, da erreichte Ziele abgehakt werden können.

2. Setzen Sie sich feste Zeiten

Nehmen Sie sich konkrete Zeiten zum Lernen/Einarbeiten vor. Im beruflichen Alltag bieten sich hier oftmals die Wochentage Donnerstag oder Freitag an. Blocken Sie sich an einem dieser Tage min. zwei Stunden und nutzen diese für die Einarbeitung (das sind „nur“ 5 % Ihrer zur Verfügung stehenden Zeit). Beginnen Sie mit den schwierigsten und größten Aufgaben zuerst.

3. Verstehen Sie den Gesamtzusammenhang

Wenn Sie den Gesamtzusammenhang verstehen, hilft es Ihnen, das Gelernte in eine Struktur einzusortieren. Sie können sich dadurch besser fokussieren. Orientieren Sie sich beim Gesamtzusammenhang an der Gliederung dieses Buches. Die Lernzeit verkürzt sich und das Verständnis steigt.

4. Fertigen Sie Skizzen an

Versuchen Sie das Gelernte so einfach wie möglich in Skizzen festzuhalten. Je einfacher die Skizzen werden, umso besser haben Sie den Zusammenhang verstanden. Skizzen können auch Mind Maps sein oder kurze Skizzen, die den Zusammenhang zwischen Ursache und Wirkung über Blockschaltbilder darstellen.

5. Lassen Sie sich nicht ablenken

Dies bedeutet, dass Outlook geschlossen und das Telefon stumm geschaltet ist. Ideal ist es, wenn Sie im Homeoffice oder in einem abgeschlossenen Raum arbeiten können. Der Fokus auf die eine Aufgabe steigt.

6. Lehren Sie

Erklären Sie Ihren Kollegen und Vorgesetzten Ihr Vorgehen und Ihre Erfahrungen. Je einfacher (und kürzer) Sie erklären und die Rückfragen Ihrer Kollegen beantworten können, umso größer ist ihr Verständnis. Das müssen Sie üben. Können Sie eine Frage nicht beantworten, zeigt dies eine Lücke auf, die Sie durch zusätzliches Studium schließen können. Sie werden merken, dass Ihr Ansehen bei Ihren Kollegen steigt. Sie erreichen schrittweise einen Expertenstatus.

7. Belohnen Sie sich

Belohnen Sie sich nach erreichten Zielen. Dies können auch Kleinigkeiten sein, z. B. ein früherer Feierabend, ein Kaffee mit den Kollegen oder etwas Zeit mit der Familie. Wichtig ist, dass Sie das Gefühl haben, sich etwas Gutes zu tun. Das motiviert!

4 ÜBERBLICK

Die Statistik lässt sich grob in zwei Bereiche unterteilen. Die deskriptive Statistik oder die beschreibende Statistik beschäftigt sich mit der Darstellung empirischer Daten in Kennzahlen, Tabellen oder Grafiken. Sie schafft die Basis für die Verwendung und Interpretation von Daten. Da wir Informationen sehr gut visuell aufnehmen, sind Grafiken von zentraler Bedeutung.

Im ersten Teil (Grundlagen) werden hierfür die wichtigsten Verfahren und Kennwerte vorgestellt. Beispiele hierfür sind durchschnittliche Einkommen, Alterspyramiden oder Verteilungen von Messwerten in Histogrammen. Diesen Teil der Statistik nennt man oft auch beschreibende Statistik. Dafür werden sowohl grafische, also auch rechnerische Verfahren genutzt.

Mit Hilfe des zweiten Teils (Daten erheben) werden wir erfahren, worauf es bei der Erhebung von Daten ankommt. Das schließt die Versuchsplanung und die Messsystemanalyse mit ein.

Der dritte Teil (Daten auswerten) ist die induktive oder schließende Statistik. Mit diesen Methoden können wir „Licht ins Dunkel“ bringen. Sie erlauben Zusammenhänge zu erkennen, Unterscheide festzustellen oder auch von kleinen Stichproben auf größere Mengen zu schließen. Dies geschieht überwiegend in drei Schritten, nach denen alle Tests beschrieben sind.

Beim vierten Teil (Daten richtig präsentieren) erfahren Sie, worauf es bei der Darstellung Ihrer Ergebnisse ankommt.

Teil 1: Grundlagen (Kapitel 7 – 11)

Praxis ohne Theorie leistet immer noch mehr als Theorie ohne Praxis. – Quintilian

Es werden die wichtigsten Verfahren vorgestellt, um Daten zu beschreiben. Dazu werden sowohl grafische als auch rechnerische Verfahren gezeigt. Wichtig hierbei ist, immer mit möglichst vielen verschiedenen Methoden die Daten zu betrachten. Jede Methode ist dabei eine Sicht auf Ihre Daten. Je mehr Sichtweisen Sie bekommen, umso sicherer sind Ihre Aussagen.

Im [Kapitel 7](#) lernen Sie die wichtigsten Grundbegriffe kennen. Dies schafft Klarheit in der Kommunikation und vermeidet Missverständnisse.

In [Kapitel 8](#) werden die zentralen rechnerischen Merkmale der Statistik eingeführt, um Streuungen und mittlere Werte zu beschreiben.

In [Kapitel 9](#) werden die wichtigsten grafischen Methoden der Statistik vorgestellt. Diese sind sehr wertvoll, da wir Menschen mit Grafiken besonders gut umgehen können.

In [Kapitel 10](#) erfahren Sie alles Notwendige zu den wichtigsten statistischen Verteilungen für Ingenieure. Die statistischen Verteilungen kann man auch zu den Grafiken zählen. Da diesen aber eine zentrale Rolle zukommen, werden sie in einem separaten Kapitel behandelt.

In [Kapitel 11](#) behandeln wir den Umgang mit streuenden Daten. Dazu wird der Vertrauensbereich eingeführt, mit dem trotz Unsicherheiten genaue Aussagen möglich sind.

Teil 2: Daten erheben ([Kapitel 12](#) und [13](#))

oder: Wer misst, misst Mist! – Grundgesetz der Messtechnik

In [Kapitel 12](#) lernen Sie, wie sie Daten erheben und worauf Sie bei einer Stichprobenauswahl achten müssen. Dies ist insbesondere deswegen relevant, da praktisch alle Aussagen auf der Basis von Stichproben erfolgen.

In [Kapitel 13](#) erfahren Sie, wie Sie in einfachen Schritten, mit Hilfe der Messsystemanalyse sicherstellen, dass Sie aus den gemessenen Daten auch die richtigen Schlüsse ziehen können und wie Sie Messunsicherheiten minimieren.

Teil 3: Datenauswertung ([Kapitel 14](#), [15](#) und [16](#))

oder: Trends und Exemplarisches erkennen, Zufälliges und Flüchtliges verdrängen - das kann und sollte die Statistik leisten. – Tyll Necker

Das [Kapitel 14](#) versetzt Sie in die Lage, Daten so auszuwerten und grafisch aufzubereiten, dass diese leicht verständlich sind und interpretiert werden können. Zusätzlich lernen Sie, wie Sie zufällige Zusammenhänge (Korrelationen) von signifikanten unterscheiden und wie Sie diese nutzen können.

In [Kapitel 15](#) zeigen wir Ihnen, wie Sie mit Hilfe statistischer Tests einfach überprüfen können, ob Unterschiede zufällig oder handlungsrelevant (statistisch signifikant) sind.

In [Kapitel 16](#) behandeln wir das umfangreiche Thema Ausreißer. Das schließt das Auffinden von Ausreißern und den richtigen Umgang mit Ausreißern ein.

Mit Hilfe von [Kapitel 17](#) können Sie notwendige Stichprobenumfänge berechnen, oder Aussagen über die Power eines statistischen Tests geben.

Teil 4: Daten präsentieren ([Kapitel 18](#) und [18.3](#))

oder: Traue keiner Statistik, die du nicht selbst gefälscht hast – Autor unbekannt.

Meiner Erfahrung nach wird die Statistik manchmal aus Unwissenheit, oder – schlimmer – aus Absicht falsch angewendet, um Daten im eigenen Sinne zu präsentieren. In beiden Fällen sind die Aussagen dann nutzlos. Ziel muss es sein, mit Hilfe der Statistik aus den Daten so viel wie möglich zu lernen. Ihnen also ein Maximum an Informationen zu entlocken. Dazu ist neben der richtigen Einstellung auch das richtige Wissen nötig.

[Kapitel 18](#) gibt Ihnen Hinweise, wie Sie Daten fachmännisch interpretieren sowie präsentieren und wie Sie den Eindruck von Manipulationen vermeiden. Daneben werden auch Hinweise zur richtigen Einstellung bei der Präsentation der Ergebnisse gegeben.

In [Kapitel 18.3](#) erfahren Sie, wie Daten schnell und einfach manipuliert werden. Dieses ist mein Lieblingskapitel. Wir sind davon überzeugt, dass man mit dieser Art der Darstellung am meisten lernt, welche Gefahren bei statistischen Methoden lauern und wie man diese vermeidet.

Vergessen Sie allerdings niemals, dass die Statistik nur ein Hilfsmittel ist! Am wichtigsten ist und bleibt die Ingenieursmäßige Interpretation und Analyse der Aussagen!

5 NÜTZLICHE HELFER: DIE ASSISTENTEN

Da die Methoden vielfältig sind und außerdem die Namen verwirrend, haben wir einen einfachen Assistenten entwickelt, mit dem Sie abhängig von Ihrer Fragestellung schnell und bequem das richtige Kapitel finden, in dem Ihre Fragen beantwortet werden siehe [Abbildung 5-1](#).

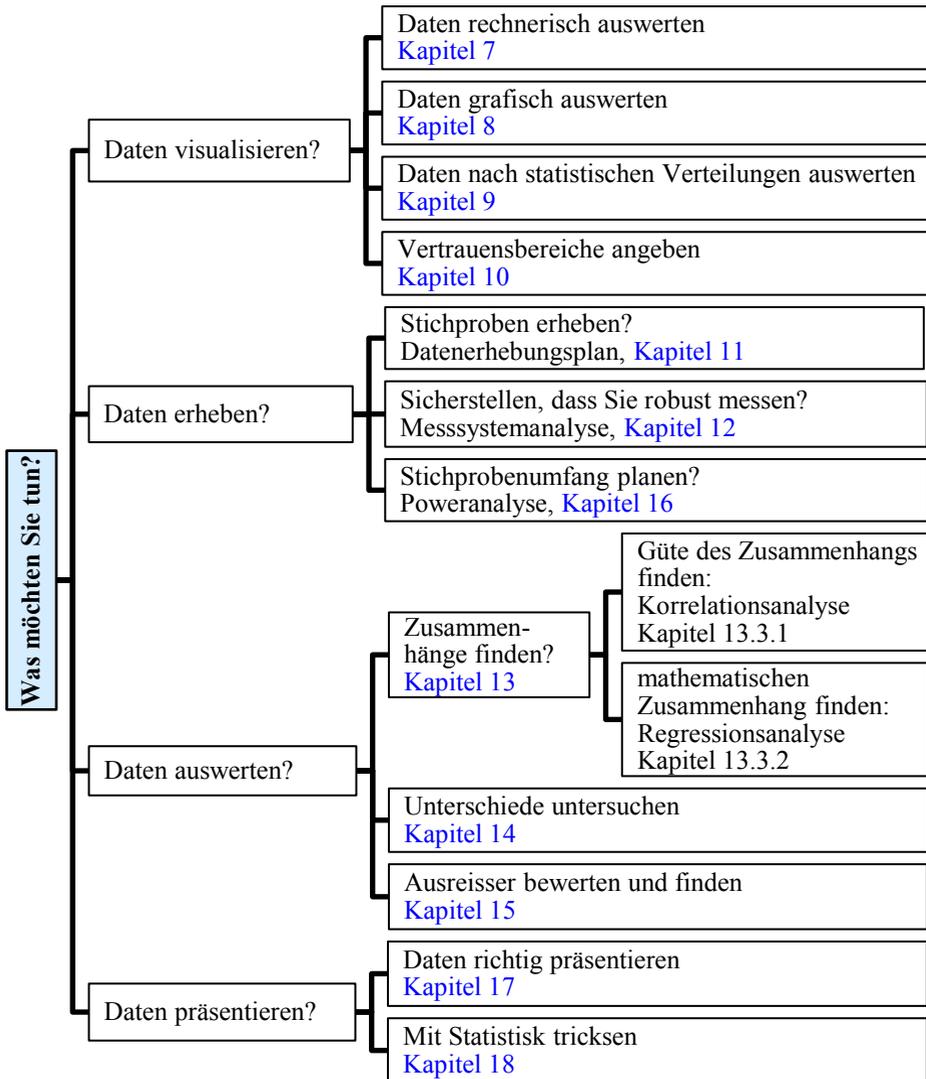


Abbildung 5-1 Assistent zur Datenanalyse (Überblick)

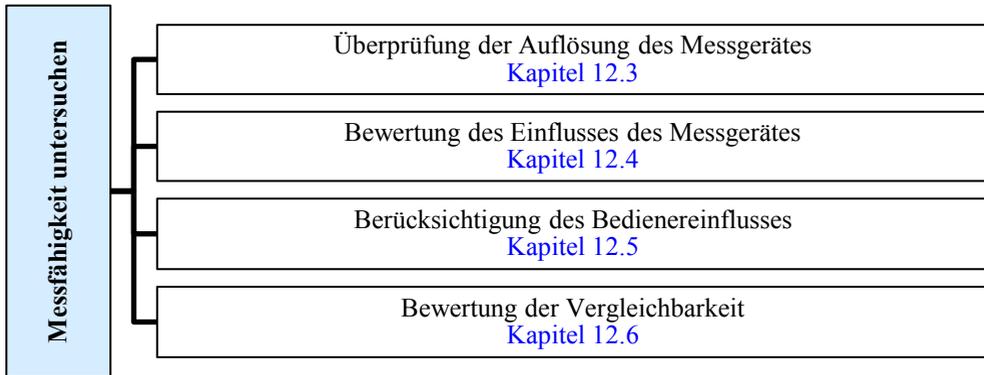


Abbildung 5-2: Assistent der Messsystemanalyse

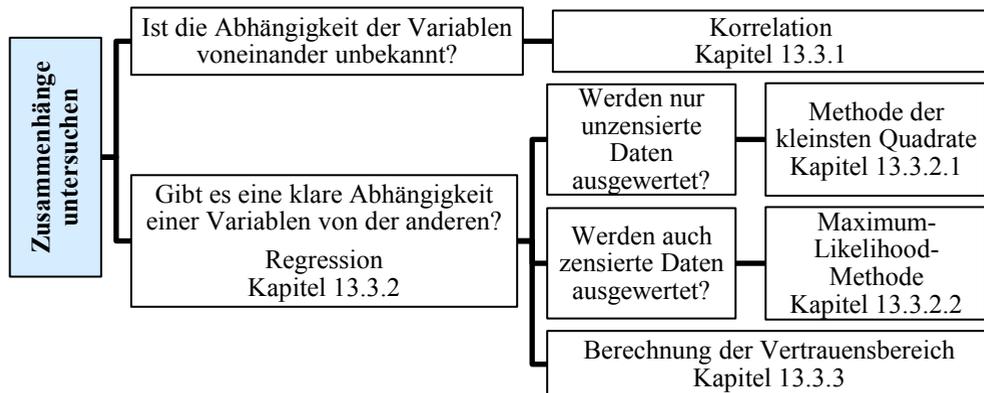


Abbildung 5-3: Assistent zur Analyse von Zusammenhängen

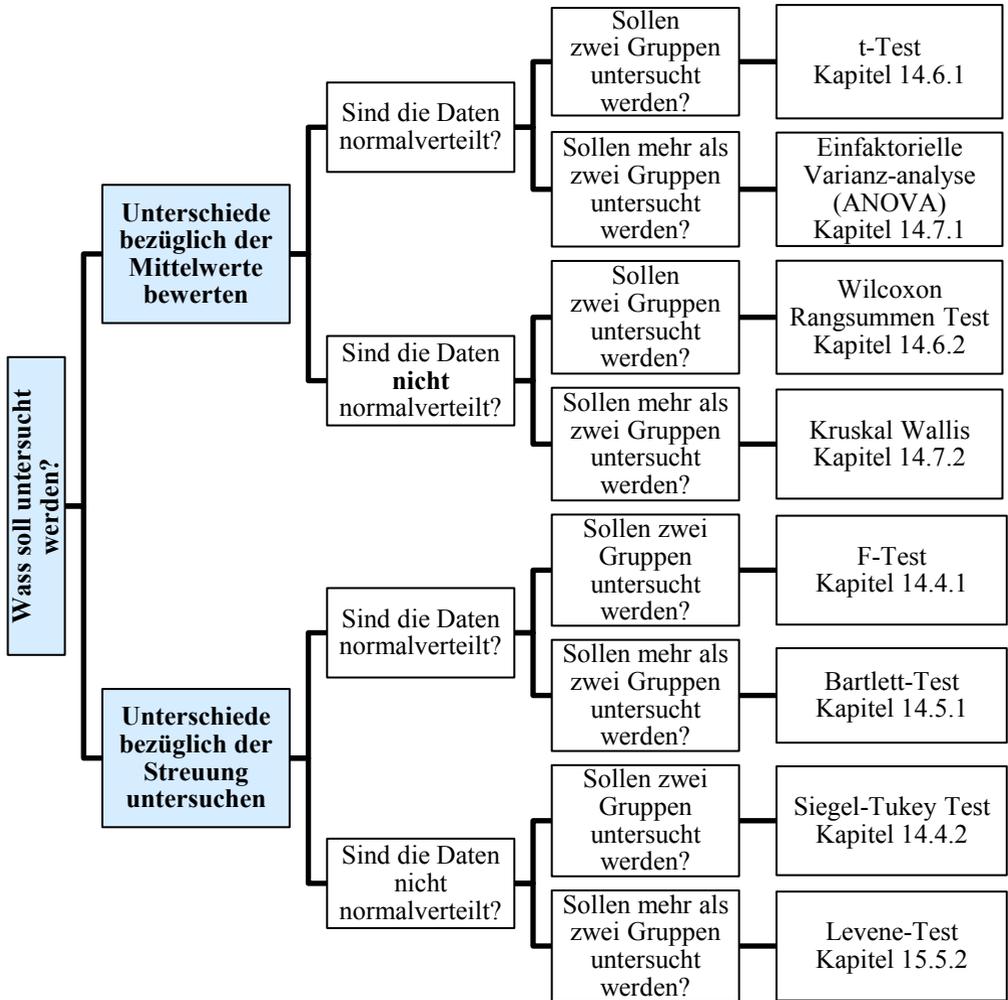


Abbildung 5-4: Assistent zur Auswahl des richtigen statistischen Tests (Teil 1)

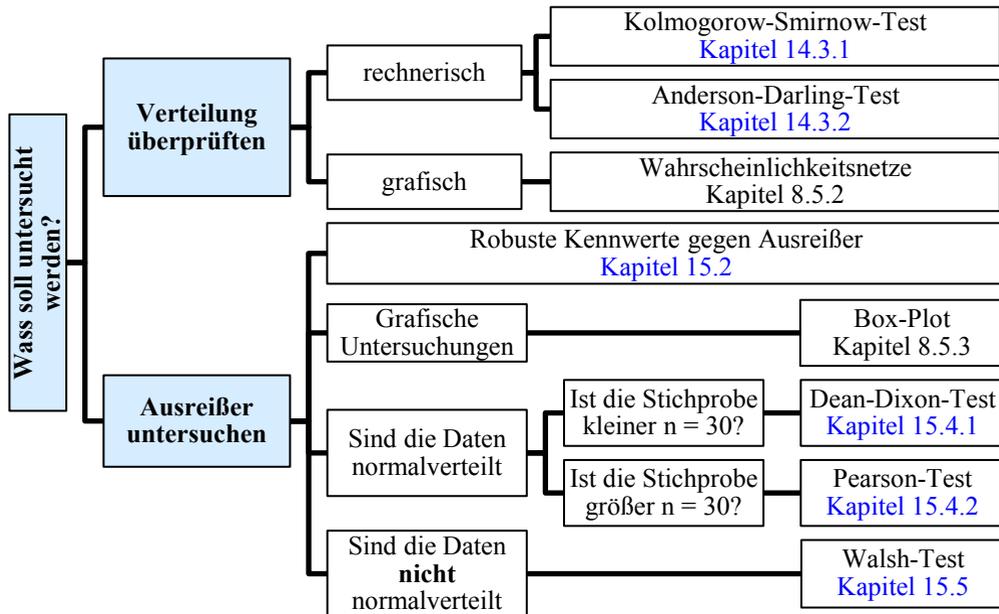


Abbildung 5-5: Assistent zur Auswahl des richtigen statistischen Tests (Teil 2)

In den weiteren Kapiteln finden Sie zusätzliche, detailliertere Assistenten:

- Assistent zur Analyse von Zusammenhängen (Abbildung 14-1, Seite 117)
- Assistent für statistische Tests (Abbildung 15-1, Seite 160)
- Assistent für Verteilungstests (Abbildung 15-4, Seite 169)
- Assistent zur Analyse von Streuungen (Abbildung 15-8, Seite 190)
- Assistent zur Analyse von Mittelwerten (Abbildung 15-14, Seite 213).

Zusätzlich haben wir Ihnen für alle vorgestellten statistischen Tests nützliche Excel-Tools bereitgestellt, mit denen Sie bequem Ihre Daten auswerten können.

6 EINFÜHRUNG IN EXCEL

Die wahrscheinlich größte Stärke von Excel ist, dass praktisch jeder die Software kennt und bereits mit ihr gearbeitet hat. Daneben bietet Excel praktisch beliebige Möglichkeiten Daten schnell und einfach zu analysieren und grafisch aufzubereiten. Auch ist Rechnen mit Excel sehr einfach möglich.

Dieses Kapitel behandelt Excel in der Tiefe, in welcher es für die Inhalte dieses Buches benötigt wird. Sollten Sie bereits viel Erfahrung mit Excel haben, können Sie dieses Kapitel wahrscheinlich überspringen.

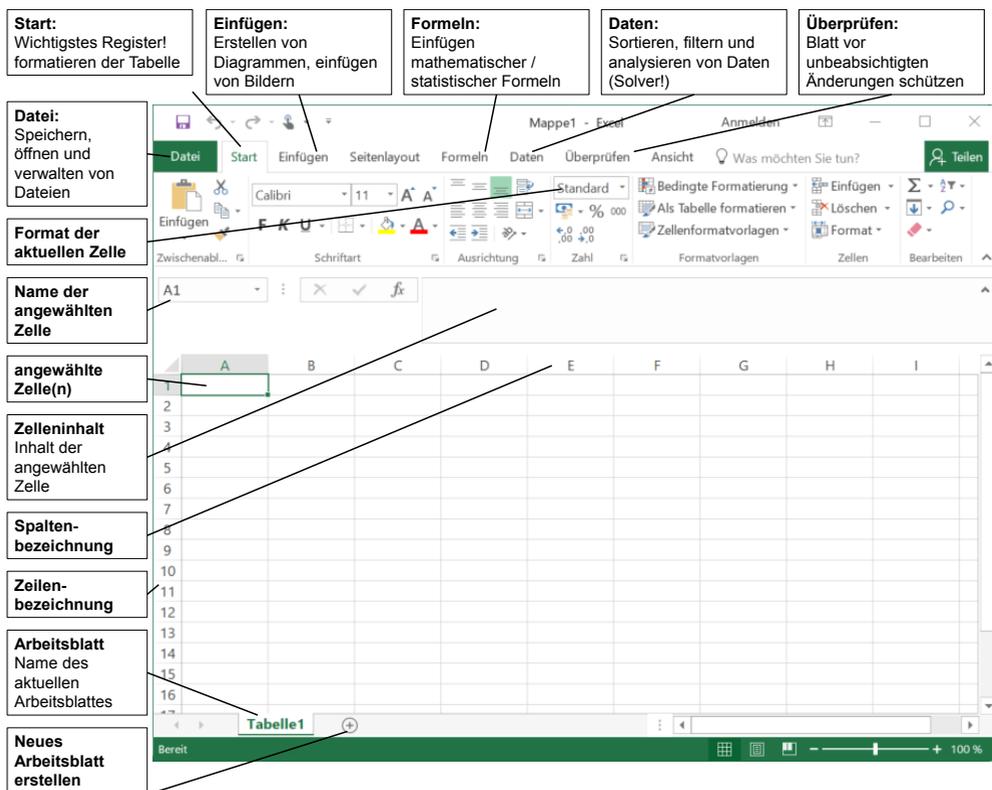


Abbildung 6-1: Startbildschirm von Excel

Wird Excel geöffnet, so erhält man das Bild aus [Abbildung 6-1](#). Dieses kann je nach Version von Excel leicht variieren. Im oberen Teil befindet sich das Menü mit den Reitern Datei, Start, Einfügen usw.. Über dieses Menü werden alle Einstellungen vorgenommen.

Beim Drücken auf einen der Reiter im Menü ändert sich das direkt unter dem Menü befindliche Menüband oder es öffnet sich ein neues Fenster.

Im Menü „Datei“ können Dateien geöffnet, gespeichert, gedruckt oder auch allgemeine Einstellungen vorgenommen werden. Das Menüband ist in Blöcke unterteilt. Wird beispielsweise das Menü „Start“ ausgewählt dann erscheint das Menüband aus [Abbildung 6-2](#). Die angezeigten Blöcke sind dann „Zwischenablage“, „Schriftart“, „Ausrichtung“, ..., „Bearbeiten“.

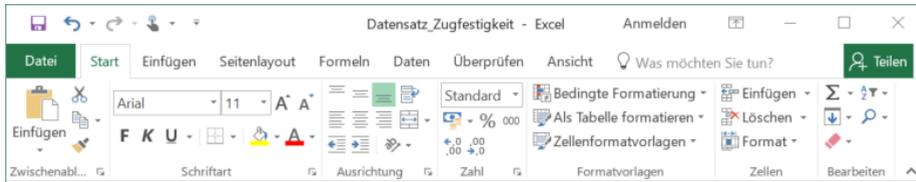


Abbildung 6-2: Das Menüband unterhalb des Menüs „Start“

Soll im Menü „Start“ und im Block „Schriftart“ eine Auswahl getroffen werden, dann wird in diesem Buch folgende Kurzschreibweise verwendet: Start→Schriftart.

Nachfolgend werden die wichtigsten Menüs kurz vorgestellt:

Im Menü „Datei“ können neue Arbeitsblätter aufgerufen (Datei→Neu) werden, bestehende Arbeitsblätter unter dem existierenden Namen (Datei→Speichern) oder unter neuem Namen gespeichert werden (Datei→Speichern unter). Einstellungen für Excel werden in Datei→Optionen vorgenommen.

Unter „Start“ können Exceltabellen zur besseren Lesbarkeit formatiert werden. Es lassen sich Farben der Zellen, Schriften oder Zellenformate ändern, bzw. Rahmenlinien anpassen.

Mit dem Menü „Einfügen“ lassen sich Diagramme erstellen und somit die Daten visualisieren. Das sind z. B. x-y Plots, Balkendiagramme oder Tortendiagramme.

Das Menü „Formeln“ liefert alle nötigen Werkzeuge um mit den vorhandenen Daten zu rechnen. Es können beispielsweise automatisch Mittelwerte von Daten berechnet werden. Auch das Eintragen und Rechnen mit eigenen mathematischen Formeln ist möglich.

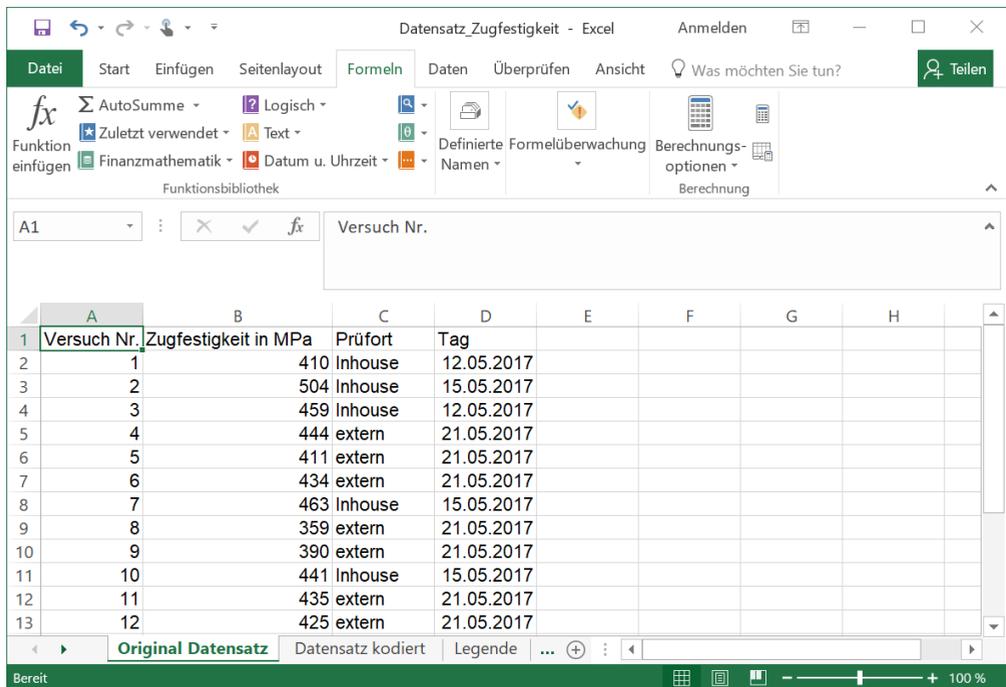
Das Menü „Daten“ erlaubt im Block „Daten→Analyse“ eine umfangreiche statistische Analyse der Daten. Daneben bietet sich die Möglichkeit Daten zu sortieren und zu filtern (um beispielsweise nur die extern gemessenen Zugfestigkeiten zu sehen).

6.1 UMGANG MIT DATEN 🚀

Excel ist als Matrix aufgebaut. Das bedeutet, dass die Daten in Spalten und Zeilen gespeichert werden. Spalten werden mit Großbuchstaben wie A, B, ... und Zeilen mit Zahlen wie 1, 2, 3, ... beschriftet. Jede Zelle erhält einen eindeutigen Namen. Dieser besteht aus dem Spaltennamen und dem Zeilennamen. Die oberste linke Zelle erhält also den Zellennamen A1.

Üblicherweise werden die Datensätze in Excel so aufgebaut, dass in einer Spalte immer die Variablen stehen (also z. B. die Zugfestigkeit) und in den Zeilen dann die Werte. Die erste Spalte erhält dann oftmals eine fortlaufende Nummer zur eindeutigen Identifizierung eines jeden Datensatzes. In [Abbildung 6-3](#) ist ein Beispiel für einen typischen Datensatz gegeben.

In diesem Fall wurden Zugversuche ermittelt. Alle vorhandenen Informationen zu den Zugfestigkeiten, den Prüforten und dem Prüftag sind in dem Bereich von Spalte B und Zeile 2 (kurz: Zelle B2) und Spalte D und Zeile 13 (kurz: D13) abgelegt. Der Versuch Nr. 10 lieferte eine Zugfestigkeit von 441 MPa, wurde Inhouse am 15.05.2017 durchgeführt.



	A	B	C	D	E	F	G	H
1	Versuch Nr.	Zugfestigkeit in MPa	Prüfort	Tag				
2	1	410	Inhouse	12.05.2017				
3	2	504	Inhouse	15.05.2017				
4	3	459	Inhouse	12.05.2017				
5	4	444	extern	21.05.2017				
6	5	411	extern	21.05.2017				
7	6	434	extern	21.05.2017				
8	7	463	Inhouse	15.05.2017				
9	8	359	extern	21.05.2017				
10	9	390	extern	21.05.2017				
11	10	441	Inhouse	15.05.2017				
12	11	435	extern	21.05.2017				
13	12	425	extern	21.05.2017				

Abbildung 6-3: Aufbau eines typischen Datensatzes

In jede Zelle können prinzipiell zwei Dinge eingetragen werden. Das sind Daten oder Formeln. Daten lassen sich bei Excel noch einmal in numerische (also Zahlen) und nichtnumerische (also z. B. Buchstaben) Daten unterteilen. Mit numerischen Werten kann gerechnet werden. Nichtnumerische Werte lassen dies nicht zu. Excel kennzeichnet das jeweilige Format in der Darstellung (siehe [Abbildung 6-4](#)).

Alle numerischen Daten werden in Excel rechtsbündig dargestellt. Dazu zählen Daten wie Zahlen, Uhrzeiten, ein Datum, Währungen oder prozentuale Werte. Die nichtnumerischen Werte werden dagegen linksbündig dargestellt. Das sind beispielsweise Buchstaben, beliebige Zeichenfolgen, unsinnige Datumsangaben, eine Mischung aus Zahlen und Buchstaben oder Zahlen mit dem Punkt als Dezimaltrennzeichen.

	A	B
1	Zahlen (numerisch)	Daten (nichtnumerisch)
2	2246	Stuttgart
3	2254,77	2245.4
4	45%	45 Prozent
5	254,47 €	254,47 £
6	31.01.2018	32.01.18
7	07:07:45	

Abbildung 6-4: Darstellung von Daten in Excel

Es ist immer einfacher mit Zahlen zu arbeiten als mit Texten. Deswegen werden Datensätze häufig so modifiziert, dass anstelle von Texten mit Zahlen gearbeitet wird. Das Vorgehen wird als Kodierung des Datensatzes bezeichnet. Im vorliegenden Beispiel kodieren wir den Prüfort. Es gilt Inhouse = 1 und extern = 2. Der Datensatz gleicht dann [Abbildung 6-5](#).

	A	B	C	D	E	F	G	H
1	Versuch Nr.	Zugfestigkeit in MPa	Prüfort	Tag				
2	1	410	1	12.05.2017				
3	2	504	1	15.05.2017				
4	3	459	1	12.05.2017				
5	4	444	2	21.05.2017				
6	5	411	2	21.05.2017				
7	6	434	2	21.05.2017				
8	7	463	1	15.05.2017				
9	8	359	2	21.05.2017				
10	9	390	2	21.05.2017				
11	10	441	1	15.05.2017				
12	11	435	2	21.05.2017				
13	12	425	2	21.05.2017				

Abbildung 6-5: Kodierter Datensatz bzgl. des Prüfortes

Um die Kodierungen und den Datensatz in Summe auch später noch zu verstehen, wird eine Legende benötigt ([Abbildung 6-6](#)). In diese Legende werden alle wichtigen Informationen zu den Variablen, der Skala der Daten eingetragen (siehe [Kapitel 7](#) auf Seite 35). Die Skalen von Daten sind entscheidend für die statistischen Tests. Manchmal sind Datensätze unvollständig, weil Messwerte unplausibel waren, die Informationen nicht bekannt sind oder aus sonstigen

Gründen. In diesen Fällen werden die fehlenden Werte durch einen Fantasiewert ersetzt. Einzige Bedingung für diesen Wert ist, dass dieser nicht im Datensatz vorkommen darf. Im Beispiel von [Abbildung 6-6](#) werden fehlende Werte mit * gekennzeichnet.

Variable	Einheit	Beschreibung der Variable	Werte	Fehlende Werte	Skala
Versuch Nr.	-	fortlaufende Nummer des Versuches	fortlaufend	*	ordinal
Zugfestigkeit	MPa	gemessene Zugfestigkeit	in MPa	*	metrisch
Prüfort	-	Ort an dem die Prüfung durchgeführt wurde	1=Inhouse 2=extern	*	nominal
Tag	Datum	Tag der Prüfung	Datum	*	metrisch

Abbildung 6-6: Legende eines Datensatzes

6.2 DATEN EINFÜGEN

Es gibt prinzipiell drei Arten wie Daten nach Excel kommen. Die einfachste Art ist das Eintragen der Daten per Hand. Leider ist dies auch die fehleranfälligste Art.

Eine weitere Möglichkeit besteht darin, die Daten per Copy Paste in Excel einzufügen. Der gewünschte Datensatz wird markiert und mittels der Tastenkombination „Strg + C“ in die Zwischenablage kopiert. In Excel wird dann die Zelle markiert, an welcher der Datensatz eingefügt wird. Durch die Tastenkombination „Strg + V“ wird der Datensatz anschließend aus der Zwischenablage in diese Zelle kopiert.

Die dritte Möglichkeit besteht im Datenimport. Diesen finden Sie unter „Datei → Externe Daten abrufen“. Hier können Sie dann wählen woher die Daten kommen, z. B. aus Textdateien oder andern externen Quellen.

Hierbei sei angemerkt, dass öfter Probleme mit der Formatierung oder dem Datenformat auftreten. Excel erkennt manchmal nicht die Trennungen zwischen Daten oder interpretiert Trennzeichen wie den „;“ gerne anders als der Nutzer. Richten Sie sich beim Aufbereiten der Daten auf einen relativ großen Zeitraum ein!

6.3 MIT DATEN RECHNEN (FORMELN)

Eine statistische Analyse von Daten bedeutet, dass mit den Daten gerechnet wird. Dazu werden Formeln benötigt. In Excel ist dies sehr bequem möglich. Es gibt zwei Möglichkeiten mit Formeln zu rechnen. Zum Einen kann die Formel direkt und händisch in die Zelle eingetragen werden. Soll beispielsweise der Inhalt von Zelle B2 mit dem Inhalt von Zelle B3 addiert werden und das Ergebnis in Zelle B14 ausgegeben werden, dann schreibt man in Zelle B14 folgende Formel: = B2 + B3

Diese Formeln können beliebig kompliziert werden. Es sind alle gängigen mathematischen Operationen (z. B. *, /, +, -) möglich. Auch Klammerausdrücke oder Potenzen (^) können eingegeben werden.

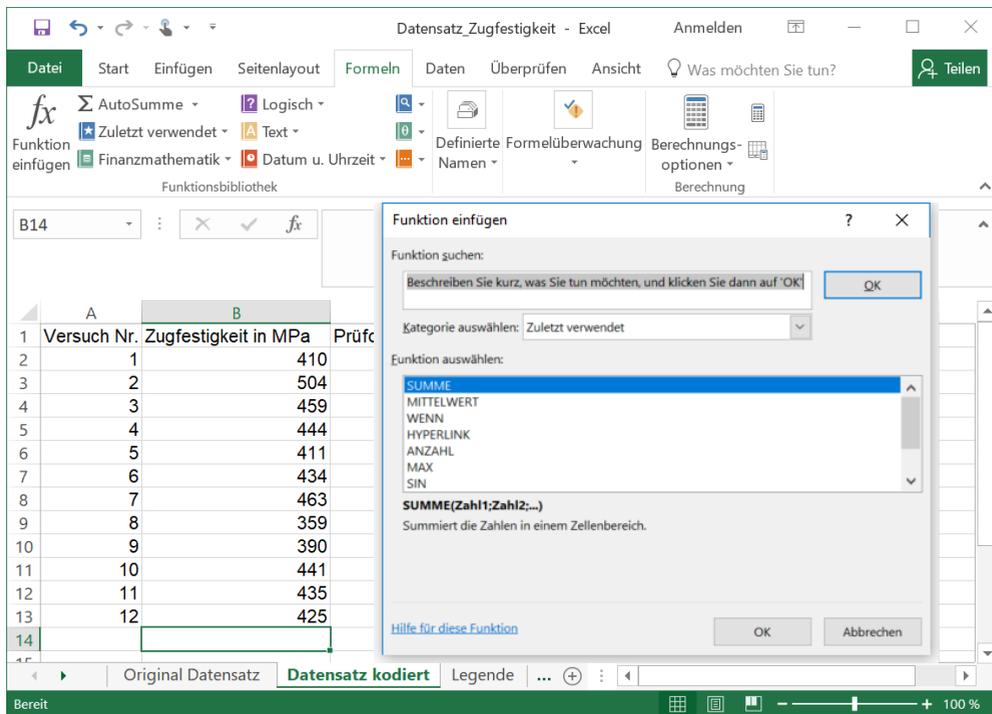


Abbildung 6-7: Formeln in Excel einfügen

Zum Anderen kann für typische mathematische Funktionen wie die Berechnung des Mittelwertes oder der Quadratwurzel auf bereits vorhandene Gleichungen in Excel zurückgegriffen werden. Angenommen, es soll der Mittelwert der Zugfestigkeiten aus obigem Beispiel berechnet und in die Zelle B14 ausgegeben werden. Dazu wird die Zelle B14 markiert und unter „Datei→Formeln“ die Schaltfläche „Funktion einfügen“ (Abbildung 6-7) gewählt. Es öffnet

sich dann das Fenster „Funktionen einfügen“. Hier kann aus einer Vielzahl von Funktionen ausgewählt werden.

Für die Berechnung des Mittelwertes wird die Funktion „MITTELWERT“ gewählt und mit „OK“ die Auswahl bestätigt. Danach können die Zellen ausgewählt werden, für welche der Mittelwert berechnet werden soll. Durch drücken und halten der „Strg“ Taste und gleichzeitiges Auswählen mit der linken Maustaste können einzelne Zellen gewählt werden. Wird die „Strg“ Taste nicht gedrückt, dann kann durch gedrückt halten des linken Maustaste ein ganzer zusammenhängender Bereich von Zellen ausgewählt werden. Für das Beispiel werden alle Zellen von B2 bis B13 ausgewählt.

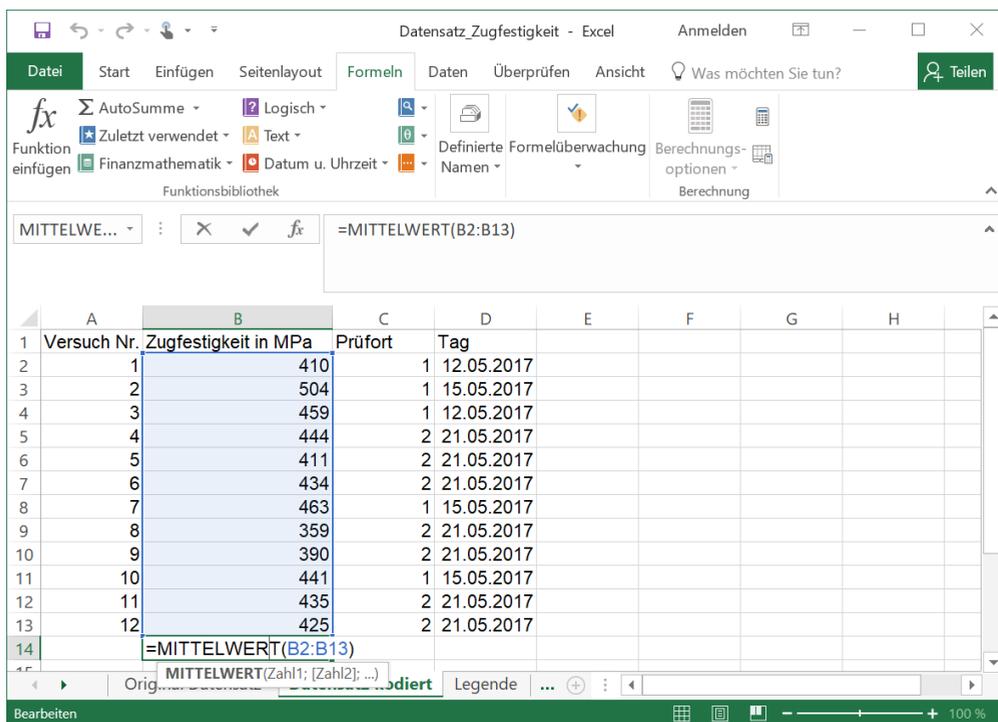


Abbildung 6-8: Ergebnis der Berechnung des Mittelwertes

Abbildung 6-8 zeigt das Ergebnis, wenn die Zelle B14 durch einen Doppelklick ausgewählt wird. Es wird die Funktion angezeigt, die von Excel auf Grund unserer Wahl automatisch eingetragen wurde. Diese lautet: =MITTELWERT(B2:B13).

Theoretisch hätten wir auch direkt diese Formel in die Zelle B14 eintragen können. In diesem Buch werden alle Formeln für die in Excel Funktionen existieren wie folgt angegeben: EXCEL: =MITTELWERT(B2:B13). Je erfahrener Sie werden, umso eher werden Sie auf die direkte Eingabe übergehen. Für eine bequeme Auswahl der Formeln ist auch die Hilfefunktion von Excel sehr wertvoll.