

Unterlagen zum Tutorium der Lehrveranstaltung

MATHEMATIK für NATURWISSENSCHAFTEN (STATISTIK)

Christoph Huber
FB Mathematik
der Universität Salzburg

Stand: 30. April 2008

Inhaltsverzeichnis

1 Einfaches Histogramm	1
1.1 Definieren eines Namens für die Stichprobenwerte	1
1.2 Häufigkeiten zählen	2
1.3 Histogramm erstellen	3
1.4 Editieren eines Histogramms	5
1.5 Alternative zum Häufigkeiten zählen	8
1.6 Histogramm mit der Excel- eigenen Histogramm- Funktion .	10
2 Histogramm mit Klasseneinteilung	12
2.1 Häufigkeiten zählen	12
2.2 Histogramm erstellen	13
3 Stängel-Blatt-Diagramm	14
3.1 Sortieren der Daten	14
3.2 Vorbereiten der Daten für das Stängel-Blatt-Diagramm . . .	15
3.3 Erstellen des Stängel-Blatt-Diagramms	16
4 Empirische Verteilungsfunktion	18
4.1 Vorbereiten der Daten	18
4.2 Empirische Verteilungsfunktion erstellen	21
4.3 Editieren der „Empirischen Verteilungsfunktion“	23
5 Kasten-Bild (Boxplot)	25
5.1 Vorbereiten der Daten für des Kasten-Bild	25
5.2 Erstellen des Kasten-Bildes	27
5.3 Kasten-Bild-Diagrammoption einstellen	30
6 Lineare Regression	32
6.1 Vorbereiten der Daten	32
6.2 Darstellen der Stichproben in einem Koordinatensystem . . .	32
6.3 Darstellung der Regressionsgeraden	33
6.4 Berechnung der Parameter der Regressionsgeraden	35

1 Einfaches Histogramm

Ein Histogramm ist eine graphische Darstellung in Säulenform. Jede Säule entspricht einer Klasse (mehr zu Klassen in Abschnitt 2).

1.1 Definieren eines Namens für die Stichprobenwerte

Zu Beginn geben wir dem Bereich, in dem sich die Stichprobenwerte befinden, einen Namen. Dadurch ist in weiterer Folge ein einfacheres und komfortableres Arbeiten mit den Stichprobenwerten möglich.

Dazu gehen wir wie folgt vor:

1. Den gesamten Bereich, in dem sich die Stichprobenwerte befinden, markieren [hier: (A1:O15)].
2. EINFÜGEN → NAMEN → DEFINIEREN... (Abbildung 1).

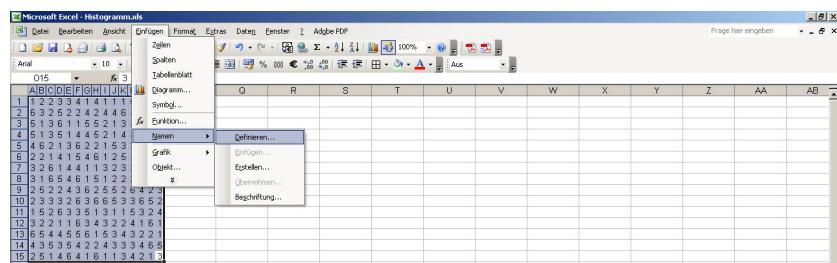


Abbildung 1:

3. Danach geben wir in das Textfeld „Namen in der Arbeitsmappe“ den frei gewählten Namen den wir vergeben möchten [hier: **daten**] und mit **OK** bestätigen (Abbildung 2).

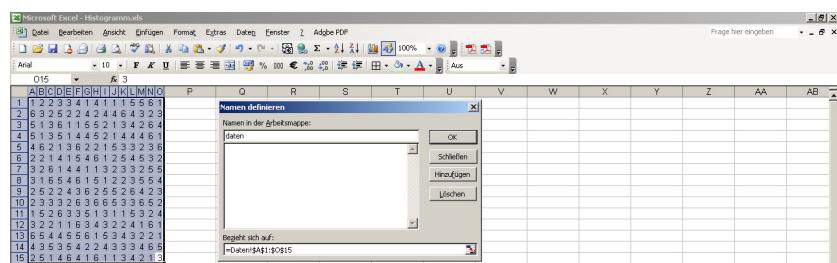


Abbildung 2:

1.2 Häufigkeiten zählen

Jetzt werden wir die Anzahl der jeweils gewürfelten Augenzahlen (absolute Häufigkeiten) zählen.

1. Wir öffnen ein neues Tabellenblatt und geben ihm den Namen „Histogramm“.
2. Danach füllen wir die Zellen entsprechend Abbildung 3 aus.

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	
3	2	
4	3	
5	4	
6	5	
7	6	
8		

Abbildung 3:

3. In die Zelle **B2** tragen wir folgende Formel ein: **=ZÄHLENWENN(daten;A2)** und kopieren die Formel nach unten (Abbildung 4). Das Ergebnis ist in Abbildung 5 zu sehen.

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	=ZÄHLENWENN(daten;A2)
3	2	
4	3	
5	4	
6	5	
7	6	
8		

Abbildung 4:

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	44
4	3	43
5	4	36
6	5	35
7	6	27
8		

Abbildung 5:

4. Jetzt können wir noch überprüfen, ob die Ergebnisse richtig sind, indem wir die Gesamtsumme (Anzahl aller Würfelwürfe) der Versuche bilden, indem wir in die Zelle **B8** folgende Formel eingeben:
=SUMME(B2:B7) (Abbildung 6). Das Ergebnis ist in Abbildung 7 zu sehen.

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	44
4	3	43
5	4	36
6	5	35
7	6	27
8	Summe:	=SUMME(B2:B7)

Abbildung 6:

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	44
4	3	43
5	4	36
6	5	35
7	6	27
8	Summe:	225

Abbildung 7:

1.3 Histogramm erstellen

Nach oben geleisteter Vorarbeit, können wir beginnen das eigentliche Histogramm zu erstellen. Dazu gehen wir wie folgt vor:

1. Wir markieren mit der Maus jene Zahlen, welche als Histogrammbalken dargestellt werden sollen (siehe Abbildung 8).

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	44
4	3	43
5	4	36
6	5	35
7	6	27
8	Summe:	225

Abbildung 8:

2. Danach öffnen wir den Diagramm-Assistenten entweder mittels: **EINFÜGEN → DIAGRAMM...** oder dem entsprechenden Icon  auf der Standard-Toolbar.
3. Auswählen des Diagramm- bzw. Diagrammuntertyps (Abbildung 9). Anschließend auf **Weiter >** klicken, um zu Schritt 2 des Assistenten zu gelangen.

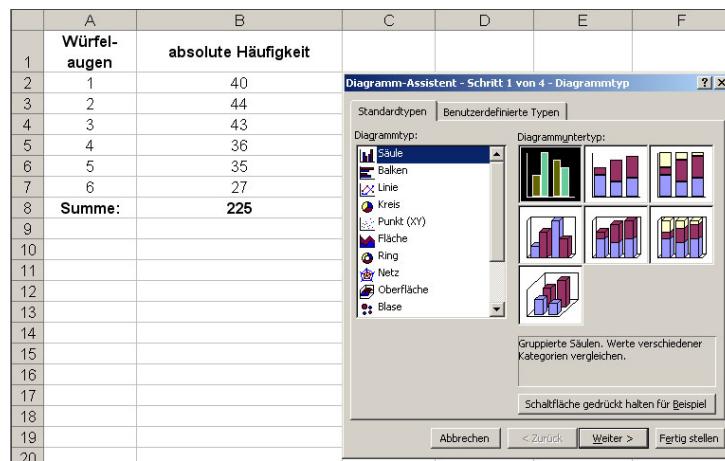


Abbildung 9:

4. Auf Reiter **REIHE** (roter Bereich) klicken (siehe Abbildung 10).
5. Die Schaltfläche ganz unten (grüner Bereich) klicken, um die *Beschriftung der Rubrikenachse* (X-Achse) vorzunehmen (siehe Abbildung 10).
6. Den Bereich (**A2:A7**) mit der Maus markieren und wieder mit dem Icon bestätigen, um zur vorherigen Dialogbox zurückzukehren (siehe Abbildung 10).

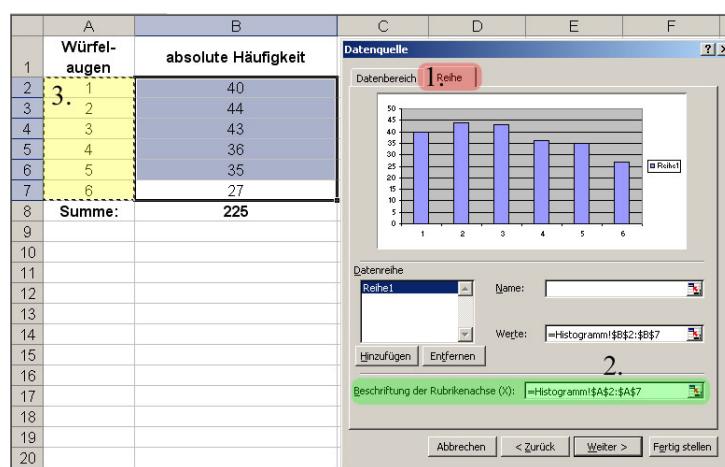


Abbildung 10:

7. Anschließend auf **Weiter >** klicken, um zu Schritt 3 des Assistenten zu gelangen.
8. Festlegen der Überschrift des Histogramms, sowie der Achsenbeschriftung (Abbildung 11). Des weiteren können hier noch verschiedenste Einstellungen für diverse Formatierungen von Achsen, Gitternetzlinien, Legende, Datenbeschriftungen und der Datentabelle gemacht werden (alles optional).

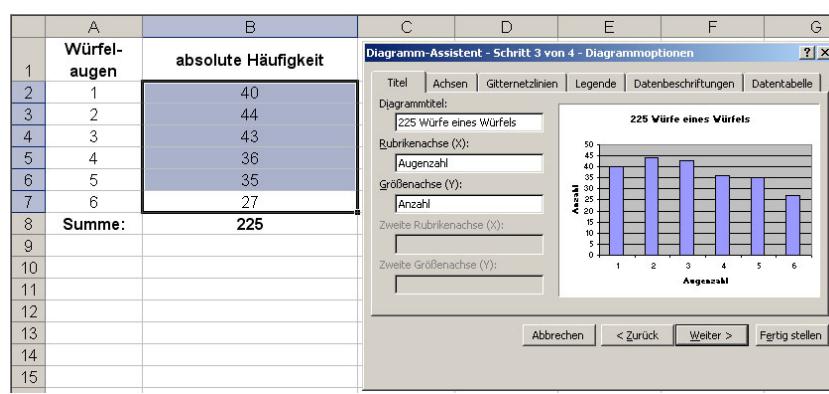


Abbildung 11:

9. Auf **Weiter >** klicken, um zum Letzten Schritt des Assistenten zu gelangen.
10. Hier kann man sich aussuchen, ob das Histogramm auf die aktuelle Tabelle als eigenes Objekt abgelegt werden soll, oder ob man ein eigenes Tabellenblatt mit dem Diagramm erstellen will. Wir wählen *Als Objekt in: Histogramm* und bestätigen mit **Fertigstellen** (Abbildung 12).
11. Damit sollte das Ergebnis wie in Abbildung 13 aussehen.

1.4 Editieren eines Histogramms

Wie man sehen kann, ist das Ergebnis nicht ganz zufriedenstellend, da die Breiten der einzelnen Histogrammbalken nicht über die gesamte „Klassenbreite“ geht. Dieses Problem kann man wie folgt lösen:

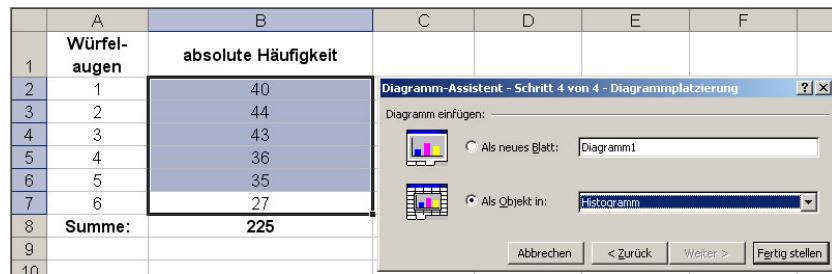


Abbildung 12:

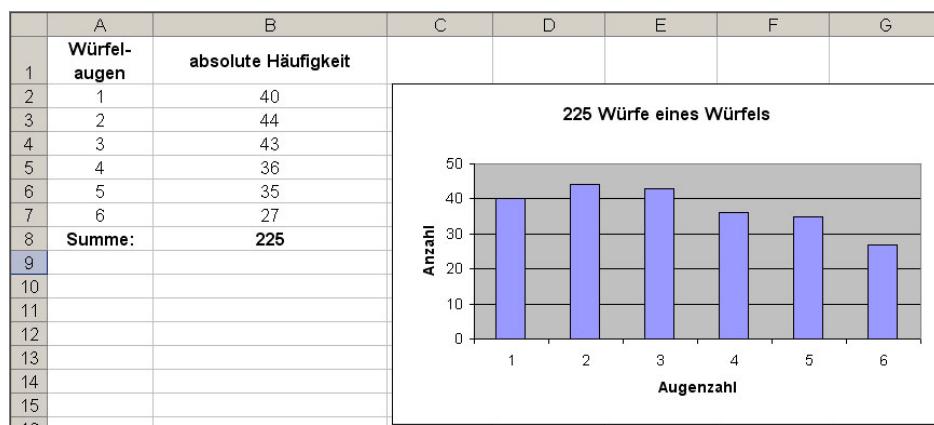


Abbildung 13:

1. Rechtsklick mit der Maus auf einen beliebigen Balken und den Punkt **DATENREIHE FORMATIEREN...** auswählen (siehe Abbildung 14).

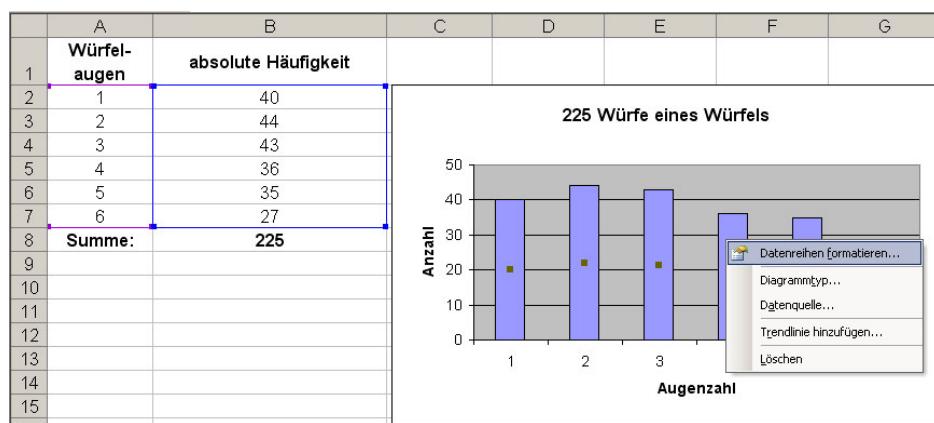


Abbildung 14:

2. Auf Reiter **OPTIONEN** (roter Bereich) klicken und die Werte *Überlappung* und *Abstandsweite* (gelber Bereich) auf „0“ stellen (siehe Abbildung 15).

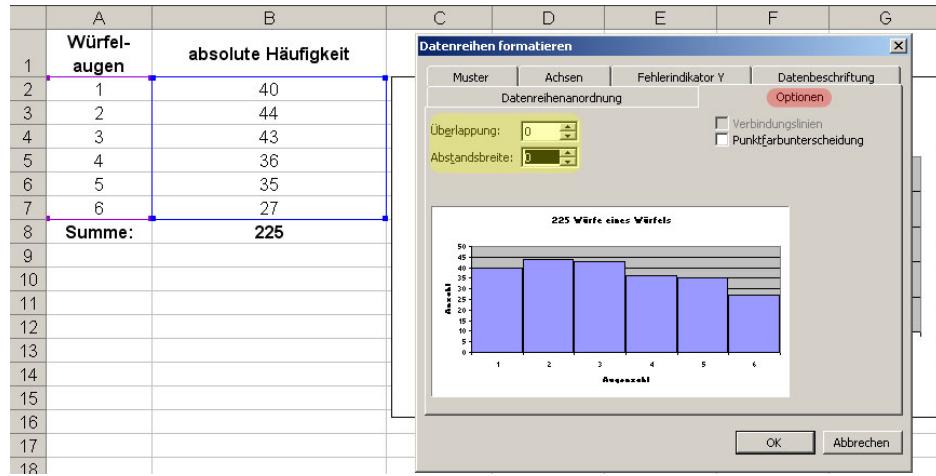


Abbildung 15:

3. Mit **OK** bestätigen. Ergebnis siehe Abbildung 16.

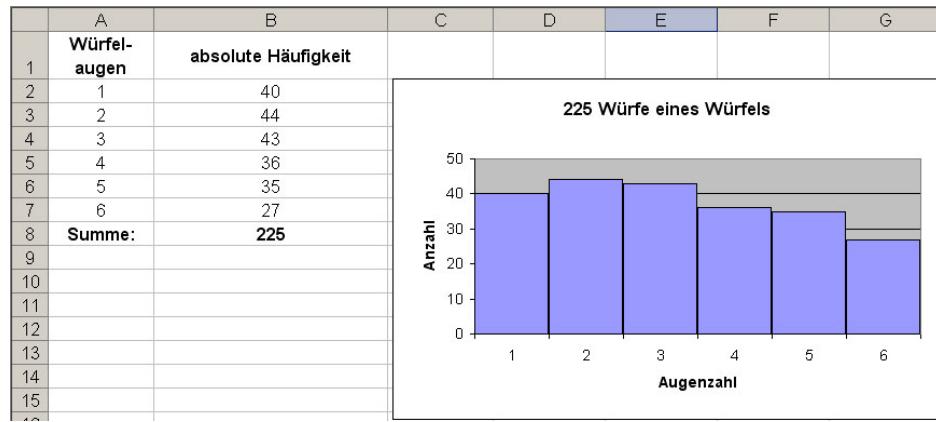


Abbildung 16:

1.5 Alternative zum Häufigkeiten zählen

Wie wir in Abschnitt 1.2 gesehen haben, ist es ein ziemlicher Aufwand die absoluten Häufigkeiten einer Datenmenge zu zählen. Excel bietet dafür wieder eine vordefinierte Formel, die so genannte HÄUFIGKEITEN Funktion. Das Besondere an der HÄUFIGKEIT-Funktion ist nun, dass sie im Gegensatz zu ZÄHLENWENN nicht in jede der Ergebniszellen einzutragen ist, sondern in alle auf einmal eingetragen wird: HÄUFIGKEIT ist eine sog. Matrixformel, ihr Ergebnis ist eine Matrix, die in unserem Beispiel aus sechs Zahlen besteht.

Gehen wir nun aber Schritt für Schritt die Anwendung dieser HÄUFIGKEITEN Funktion durch:

1. Zu Beginn bereiten wir wieder ein Tabellenblatt wie in Abschnitt 1.2 Pkt. 2 vor.
2. Danach tragen wir in die Zelle **B2** folgende Formel ein =HÄUFIGKEI(daten;D2:D7) und bestätigen mit **ENTER** (siehe Abbildung 17).

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	=HÄUFIGKEIT(daten;A2:A7)
3	2	
4	3	
5	4	
6	5	
7	6	
8		

Abbildung 17:

3. Jetzt kommt der eigentliche Unterschied zu einer herkömmlichen Excel Formel. Wir klicken in die Zelle **B2** und ziehen mit gedrückter linker Maustaste den Bereich bis Zelle **B7** (siehe Abbildung 18).
4. Danach drücken wir die Taste **F2** (siehe Abbildung 19).

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	
4	3	
5	4	
6	5	
7	6	

Abbildung 18:

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	=HÄUFIGKEIT(daten;A2:A7)
3	2	
4	3	
5	4	
6	5	
7	6	

Abbildung 19:

5. Und zum Abschluß die Tasten **Strg** + **Umschalten** + **ENTER** drücken (siehe Abbildung 20). Diese 3er-Tastenkombination muss genau in dieser Reihenfolge gedrückt werden, ansonsten funktioniert es nicht!

	A	B
1	Würfel-augen	absolute Häufigkeit
2	1	40
3	2	44
4	3	43
5	4	36
6	5	35
7	6	27

Abbildung 20:

6. Jetzt kann man wieder mit Pkt. 4 aus Abschnitt 1.2 fortfahren.

1.6 Histogramm mit der Excel-eigenen Histogrammfunktion

Excel bietet auch die Möglichkeit Histogramme „automatisch“ zu zeichnen. Dazu gehen wir wie folgt vor:

1. Wir gehen in das Tabellenblatt, auf dem unsere Würfelergebnisse (Ausgangsdaten) stehen.
2. Dort schreiben wir in die Zellen **Q1:Q6** die Zahlen von 1 bis 6 (siehe Abbildung 21).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	1	1	2	2	3	3	4	1	4	1	1	1	5	5	6	1	1
2	2	6	3	2	5	2	2	4	2	4	4	6	4	3	2	3	2
3	3	5	1	3	6	1	1	5	5	2	1	3	4	2	6	4	3
4	4	5	1	3	5	1	4	4	5	2	1	4	4	4	6	1	4
5	5	4	6	2	1	3	6	2	2	1	5	3	3	2	3	6	5
6	6	2	2	1	4	1	5	4	6	1	2	5	4	5	3	2	6
7	7	3	2	6	1	4	4	1	1	3	2	3	3	2	5	5	
8	8	3	1	6	5	4	6	1	5	1	2	2	3	5	5	4	
9	9	2	5	2	2	4	3	6	2	5	5	2	6	4	2	3	
10	10	2	3	3	3	2	6	3	6	6	5	3	3	6	5	2	
11	11	1	5	2	6	3	3	5	1	3	1	1	5	3	2	4	
12	12	3	2	2	2	1	1	6	3	4	3	2	2	4	1	6	1
13	13	6	5	4	4	5	5	6	1	5	3	4	3	2	2	1	
14	14	4	3	5	3	5	4	2	2	4	3	3	4	6	5		
15	15	2	5	1	4	6	4	1	6	1	1	3	4	2	1	3	

Abbildung 21:

3. Menü **Extras - Analyse-Funktionen ...**, im Popupfenster wählen wir dann den Punkt **Histogramm** aus und bestätigen mit **OK** (siehe Abbildung 22).



Abbildung 22:

4. Die Schaltfläche neben dem *Eingabebereich* klicken.

5. Den Bereich (**A1:O15**) mit der Maus markieren und wieder mit dem Icon bestätigen, um zur vorherigen Dialogbox zurückzukehren (siehe Abbildung 23).

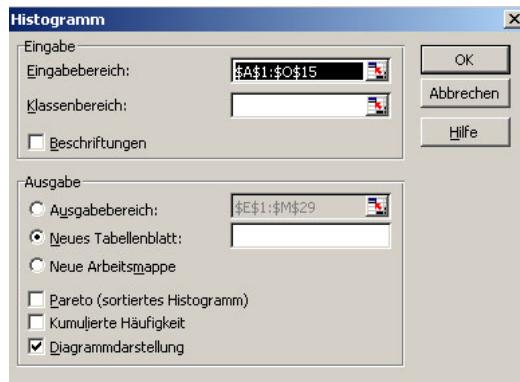


Abbildung 23:

6. Die Schaltfläche neben dem *Klassenbereich* klicken.
7. Den Bereich (**Q1:Q6**) (entspricht der linken Klassengrenze; somit kann diese Funktion auch für ein Histogramm mit Klasseneinteilung, wie in Abschnitt 2 beschrieben, verwendet werden) mit der Maus markieren und wieder mit dem Icon bestätigen, um zur vorherigen Dialogbox zurückzukehren (siehe Abbildung 24).

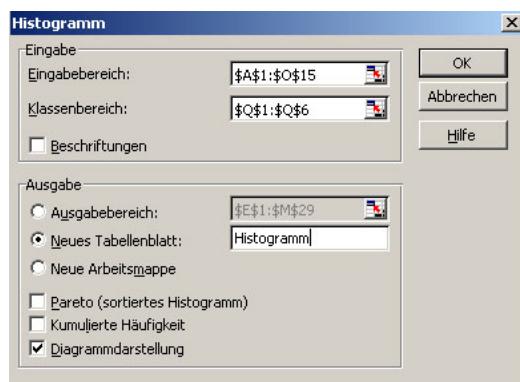


Abbildung 24:

8. Die restlichen Schaltflächen ebenfalls wie in Abbildung 24 anklicken und in das Textfeld *Neues Tabellenblatt* den Namen des neu anzulegenden Tabellenblattes eintragen (hier: *Histogramm*, welches die Auswertung beinhalten wird).
9. Mit **OK** bestätigen und das Histogramm im Tabellenblatt *Histogramm* gemäß Abschnitt 1.4 verändern.

2 Histogramm mit Klasseneinteilung

Eine Klasse ist ein schmales Teilintervall des Bereiches $[x_{min}, x_{max}]$ aller Stichprobenwerte.

2.1 Häufigkeiten zählen

1. Wir öffnen ein neues Tabellenblatt und geben ihm den Namen „*Histogramm mit Klassen*“.
2. Danach füllen wir die Zellen entsprechend Abbildung 25 aus.

	A	B	C	D
1	Klassenname	linke Klassen- grenze	rechte Klassen- grenze	Anzahl
2	{1,2}	1	2	
3	{3,4}	3	4	
4	{5,6}	5	6	
5				

Abbildung 25:

3. In die Zelle D2 tragen wir folgende Formel ein:
 $=\text{ZÄHLENWENN}(\text{daten}; ">=" & B2) - \text{ZÄHLENWENN}(\text{daten}; ">" & C2)$
 und kopieren die Formel nach unten (Abbildung 26). Das Ergebnis ist in Abbildung 27 zu sehen.
4. Jetzt können wir noch überprüfen, ob die Ergebnisse richtig sind, indem wir die Gesamtsumme (Anzahl aller Würfelwüfe) der Versuche bilden, indem wir in die Zelle D5 folgende Formel eingeben:
 $=\text{SUMME}(D2:D4)$. Das Ergebnis ist in Abbildung 28 zu sehen.

=&B2)-ZÄHLENWENN(daten;">=&C2). A callout bubble points to this formula with the text 'Formel nach unten kopieren'. The right table (Abbildung 27) shows the result after copying the formula down, with column D containing the values 84, 79, and 62."/>

	A	B	C	D
1	Klassenname	linke Klassen- grenze	rechte Klassen- grenze	Anzahl
2	{1,2}	1	2	=ZÄHLENWENN(daten;">=&B2)-ZÄHLENWENN(daten;">=&C2)
3	{3,4}	3	4	
4	{5,6}	5	6	
5				
6				
7				

	A	B	C	D
1	Klassenname	linke Klassen- grenze	rechte Klassen- grenze	Anzahl
2	{1,2}	1	2	84
3	{3,4}	3	4	79
4	{5,6}	5	6	62
5				
6				
7				

Abbildung 26:

Abbildung 27:

	A	B	C	D
1	Klassenname	linke Klassen- grenze	rechte Klassen- grenze	Anzahl
2	{1,2}	1	2	84
3	{3,4}	3	4	79
4	{5,6}	5	6	62
5				
6				
7				

Abbildung 28:

2.2 Histogramm erstellen

1. Bei der Histogrammerstellung mit Klasseneinteilung kann man eigentlich in gleicher Weise wie bei der bereits vorgestellten Histogrammerstellung in Abschnitt 1.3 vorgehen. Bei unserem Beispiel markieren wir als darzustellende Balken den Bereich (**D2:D4**).
2. Schritt 2 – Schritt 5 analog Abschnitt 1.3
3. Wie Schritt 6 bei Abschnitt 1.3 nur bei unserem Beispiel den Bereich (**A2:A4**) markieren.
4. Die restlichen Schritte können wieder analog Abschnitt 1.3 und Abschnitt 1.4 durchgeführt werden. Das Ergebniss sollte dann in etwa wie in Abbildung 29 aussehen.

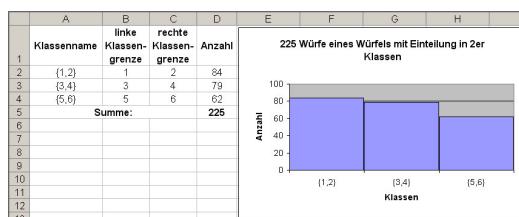


Abbildung 29:

3 Stängel-Blatt-Diagramm

Das Erstellen eines Stängel-Blatt-Diagramms ist in Excel nicht in dem Sinn möglich, wie wir es z.B. für das Erstellen eines Histogramms gewohnt sind. Mit ein paar „Tricks“ ist aber auch dies relativ schnell erledigt.

3.1 Sortieren der Daten

Wie man anhand der vorgegebenen Daten erkennen kann (Abbildung 30), sind diese in beliebiger Reihenfolge angeordnet (meist so, wie die Stichproben auch in der „Natur“ vorkommen). Wir brauchen die Daten aber in sortierter Reihenfolge und gehen dabei wie folgt vor:

	A
1	3178
2	3512
3	3473
4	3090
5	2413
83	2972
84	3424
85	3934
86	3463
87	4110
88	2737

Abbildung 30:

1. Den gesamten Bereich in dem sich die Stichprobenwerte befinden markieren [hier: (A1:A88)].
2. **DATEN → SORTIEREN...** (Abbildung 31).

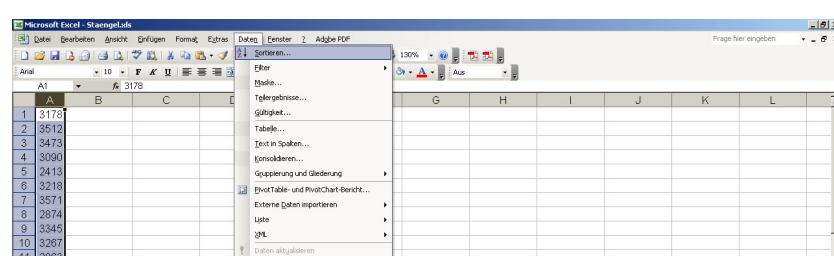


Abbildung 31:

3. Die Dialogbox gemäß Abbildung 32 ausfüllen (um die Daten aufsteigend zu sortieren) und mit **OK** bestätigen. Das Ergebnis der sortierten Daten ist in Abbildung 33 zu sehen.



Abbildung 32:

	A
1	2237
2	2256
3	2305
4	2325
5	2345
83	4022
84	4110
85	4258
86	4346
87	4444
88	4493

Abbildung 33:

3.2 Vorbereiten der Daten für das Stängel-Blatt-Diagramm

1. In die Zelle **B1** tragen wir folgende Formel ein: **=LINKS(A1;2)** und in die Zelle **C1** wird die Formel **=RECHTS(A1;2)** eingetragen (Abbildung 34). Beide Formeln werden nach unten kopiert, um das Ergebnis in Abbildung 35 zu erhalten.

	A	B	C
1	2237	=LINKS(A1;2)	=RECHTS(A1;2)
2	2256		
3	2305		
4	2325		
5	2345		
6	2394		

Abbildung 34:

	A	B	C
1	2237	22	37
2	2256	22	56
3	2305	23	05
4	2325	23	25
5	2345	23	45
6	2394	23	94

Abbildung 35:

2. Wir geben nun den Daten in Spalte **C** den zugehörigen Namen (Namensgebung von Daten siehe Abschnitt 1.1). D.h. wir geben dem Bereich in Spalte **C** für den in Spalte **B** **22** steht, den Namen „**22**“, den Bereich in Spalte **C** für den in Spalte **B** **23** steht, den Namen „**23**“, usw. (Abbildung 36).

The screenshot shows a Microsoft Excel spreadsheet with data in columns A through G. Column A contains row numbers from 77 to 88. Columns B and C contain numerical values. A callout box highlights the 'Name Define' dialog box, which lists names defined in the current workbook, including 'Daten' (with the range \$C\$87:\$C\$88) and other entries like '35', '36', etc. The 'OK' button is visible in the dialog.

	A	B	C	D	E	F	G
77	3816	38	16				
78	3864	38	64				
79	3865	38	65				
80	3895	38	95				
81	3934	39	34				
82	3944	39	44				
83	4022	40	22				
84	4110	41	10				
85	4258	42	58				
86	4346	43	46				
87	4444	44	44				
88	4493	44	93				

Abbildung 36:

3.3 Erstellen des Stängel-Blatt-Diagramms

1. Wir öffnen ein neues Tabellenblatt und geben ihm den Namen „Stängel-Blatt-Diagramm“.
2. Danach sehen wir im Tabellenblatt *Daten* in der Spalte **B** nach, wie viele verschiedene „Stängel“ vorhanden sind [hier die Zahlen von 22 – 44]. Des Weiteren zählen wir, bei welchem „Stängel“ die meisten „Blätter“ angehängt sind [hier beim „Stängel“ 32, 9 „Blätter“] und füllen die Zellen im neuen Tabellenblatt *Stängel-Blatt-Diagramm* entsprechend Abbildung 37 aus.

The screenshot shows a Microsoft Excel spreadsheet with data in columns A through J. Column A contains row numbers from 1 to 24. Column B contains numerical values. Column C contains the header 'Stängel'. Column D contains the value '1'. Columns E through J are empty. A callout box highlights the first few rows of data, showing values 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, and 44.

	A	B	C	D	E	F	G	H	I	J
1	Stängel	1	2	3	4	5	6	7	8	9
2	22									
3	23									
4	24									
5	25									
6	26									
7	27									
8	28									
9	29									
10	30									
11	31									
12	32									
13	33									
14	34									
15	35									
16	36									
17	37									
18	38									
19	39									
20	40									
21	41									
22	42									
23	43									
24	44									

Abbildung 37:

3. In die Zellen **B2:B24** tragen wir folgende Formel ein:
 $=INDEX(Bereichsname;B$1)$, wobei für *Bereichsname* der jeweilige Name des Bereichs in Spalte **A** einzutragen ist (z.B. in Zelle **A2**)

steht 22, dann ist der von uns in Abschnitt 3.2 Punkt 2 vergebene Name für 22 einzutragen, also „22“) (Abbildung 38)

	A	B	C	D	E	F	G	H	I	J
1	Stängel	1	2	3	4	5	6	7	8	9
2	22	=INDEX(_22;B\$1)								
3	23	=INDEX(_23;B\$1)								
4	24	=INDEX(_24;B\$1)								
5	25	=INDEX(_25;B\$1)								
6	26	=INDEX(_26;B\$1)								
7	27	=INDEX(_27;B\$1)								
8	28	=INDEX(_28;B\$1)								
9	29	=INDEX(_29;B\$1)								
10	30	=INDEX(_30;B\$1)								
11	31	=INDEX(_31;B\$1)								
12	32	=INDEX(_32;B\$1)								
13	33	=INDEX(_33;B\$1)								
14	34	=INDEX(_34;B\$1)								
15	35	=INDEX(_35;B\$1)								
16	36	=INDEX(_36;B\$1)								
17	37	=INDEX(_37;B\$1)								
18	38	=INDEX(_38;B\$1)								
19	39	=INDEX(_39;B\$1)								
20	40	=INDEX(_40;B\$1)								
21	41	=INDEX(_41;B\$1)								
22	42	=INDEX(_42;B\$1)								
23	43	=INDEX(_43;B\$1)								
24	44	=INDEX(_44;B\$1)								

Abbildung 38:

4. Diese Formeln werden jetzt bis zur letzten Ziffer in Zelle **A10** kopiert (Abbildung 39) und der Inhalt der Zellen, in denen die Meldung **#BEZUG!** steht, gelöscht (Abbildung 40).

	A	B	C	D	E	F	G	H	I	J
1	Stängel	1	2	3	4	5	6	7	8	9
2	22	37 56	#BEZUG!	## ## ## ## ##						
3	23	05 25	45	94	## ## ## ## ##					
4	24	03 13	53	## ## ## ## ##						
5	25	60	#BEZUG!	## ## ## ## ##						
6	26	59 68	#BEZUG!	## ## ## ## ##						
7	27	08 27	37	37 86 96						
8	28	45 74	96	## ## ## ## ##						
9	29	33 43	53	63 63 72						
10	30	12 21	80	90	## ## ## ## ##					
11	31	00 10	10	20 39 59 78	## ## ## ##					
12	32	18 18	18	37 67 67 77	86 96					
13	33	45 55	75	84 94	## ## ## ## ##					
14	34	24 53	53	63 63 73 83	83 ##					
15	35	12 12	22	41 71 90	## ## ##					
16	36	10 10	40	69 89	## ## ## ## ##					
17	37	08 18	38	87	## ## ## ## ##					
18	38	16 64	65	95	## ## ## ## ##					
19	39	34 44	#BEZUG!	## ## ## ## ##						
20	40	22 #BEZUG!	#BEZUG!	## ## ## ## ##						
21	41	10 #BEZUG!	#BEZUG!	## ## ## ## ##						
22	42	58 #BEZUG!	#BEZUG!	## ## ## ## ##						
23	43	46 #BEZUG!	#BEZUG!	## ## ## ## ##						
24	44	44 93	#BEZUG!	## ## ## ## ##						

Abbildung 39:

Abbildung 40:

4 Empirische Verteilungsfunktion

4.1 Vorbereiten der Daten

1. Die Daten aufsteigend sortieren, falls sie noch nicht sortiert sind (siehe Abschnitt 3.1) (siehe Abbildung 41 und 42).

	A
1	2737
2	2659
3	2953
4	2874
5	2796
26	2237
27	2786
28	2727
29	2845
30	3021

Abbildung 41:

	A
1	2237
2	2256
3	2305
4	2325
5	2345
26	2963
27	2972
28	3012
29	3021
30	3080

Abbildung 42:

2. Wir geben nun den Daten in Spalte A den Namen „daten“ (Namensgebung von Daten siehe Abschnitt 1.1).
3. Den kleinsten und den größten Wert aus den Daten heraussuchen [hier: kleinster = 2237, größer = 3080] und merken.
4. Den kleinsten Wert in die Zelle C1 schreiben und den kleinsten Wert + 1 in Zelle C2 (Abbildung 43).

	A	B	C
1	2237		2237
2	2256		2238
3	2305		
4	2325		
5	2345		
6	2394		

Abbildung 43:

5. Markieren der Zellen C1:C2.

6. Den Mauszeiger in die rechte untere Ecke der Zelle **C2** bewegen. Der Mauszeiger verwandelt sich dabei in ein dünnes Fadenkreuz.
7. Mit gedrückter linker Maustaste die Maus nach unten bewegen, bis im kleinen gelben Fenster der größte Wert der Daten erscheint (Abbildung 44). Ergebnis siehe Abbildung 45.

	A	B	C	D
1	2237		2237	
2	2256		2238	
3	2305			
4	2325			
5	2345			
6	2394			2240
7	2403			
8	2413			
9	2453			
10	2560			

Abbildung 44:

	A	B	C	D
1	2237		2237	
2	2256		2238	
3	2305		2239	
4	2325		2240	
5	2345		2241	
840				3076
841				3077
842				3078
843				3079
844				3080

Abbildung 45:

8. In die Zelle **D1** tragen wir folgende Formel ein: =**ZÄHLENWENN(daten;C1)** und kopieren die Formel nach unten (Abbildung 46).

	A	B	C	D
1	2237		2237	=ZÄHLENWENN(daten;C1)
2	2256		2238	
3	2305		2239	
4	2325		2240	
5	2345		2241	

Abbildung 46:

9. In die Zell **E1** tragen wir folgende Formel ein: =**D1/ANZAHL2(daten)** (Abbildung 47) und bestätigen mit **ENTER**. Dann markieren wir die Zelle **E1** mit der Maus und wählen in der Formattoolbox das Prozentformat **%** aus. Jetzt kopieren wir die Formel nach unten (Abbildung 48).
10. In die Zelle **F1** tragen wir folgende Formel ein: =**E1** und bestätigen mit **ENTER**. In die darunterliegende Zelle **F2** tragen wir dann folgende

	A	B	C	D	E
1	2237		2237	1	=D1/ANZAHL2(daten)
2	2256		2238	0	
3	2305		2239	0	
4	2325		2240	0	
5	2345		2241	0	

Abbildung 47:

	A	B	C	D	E
1	2237		2237	1	3,33%
2	2256		2238	0	0,00%
3	2305		2239	0	0,00%
4	2325		2240	0	0,00%
5	2345		2241	0	0,00%

Abbildung 48:

Formel ein: **=F1+E2** und kopieren die Formel von **E2** nach unten (Abbildung 49). Das Ergebnis ist in Abbildung 50 zu sehen.

	A	B	C	D	E	F
1	2237		2237	1	3,33%	=E1
2	2256		2238	0	0,00%	=F1+E2
3	2305		2239	0	0,00%	
4	2325		2240	0	0,00%	
5	2345		2241	0	0,00%	
840				3076	0	0,00%
841				3077	0	0,00%
842				3078	0	0,00%
843				3079	0	0,00%
844				3080	1	3,33%
						100,00%

Abbildung 49:

Abbildung 50:

4.2 Empirische Verteilungsfunktion erstellen

1. Wir markieren die Spalte C und mit gedrückter STRG-Taste die Spalte F (Abbildung 51).

	A	B	C	D	E	F
1	2237		2237	1	3,33%	3,33%
2	2256		2238	0	0,00%	3,33%
3	2305		2239	0	0,00%	3,33%
4	2325		2240	0	0,00%	3,33%
5	2345		2241	0	0,00%	3,33%

Abbildung 51:

2. Danach öffnen wir den Diagramm-Assistenten entweder mittels: **EINFÜGEN → DIAGRAMM...** oder entsprechenden Icon  auf der Standard-Toolbar.
3. Auswählen des Diagramm- bzw. Diagrammuntertyps siehe Abbildung 52. Anschließend zweimal auf **Weiter >** klicken, um zu Schritt 3 des

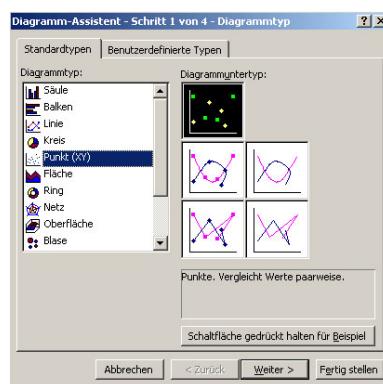


Abbildung 52:

Assistenten zu gelangen.

4. Festlegen der Überschrift des Diagramms, sowie der ev. Achsenbeschriftung.

5. Auf **Weiter >** klicken, um zum Letzten Schritt des Assistenten zu gelangen.
6. Hier kann man sich aussuchen, ob das Histogramm auf die aktuelle Tabelle als eigenes Objekt abgelegt werden soll, oder ob man ein eigenes Tabellenblatt mit dem Diagramm erstellen will. Wir wählen *Als neues Blatt: Empirische Verteilungsfunktion* und bestätigen mit **Fertigstellen** (Abbildung 53). Ergebnis siehe Abbildung 54.

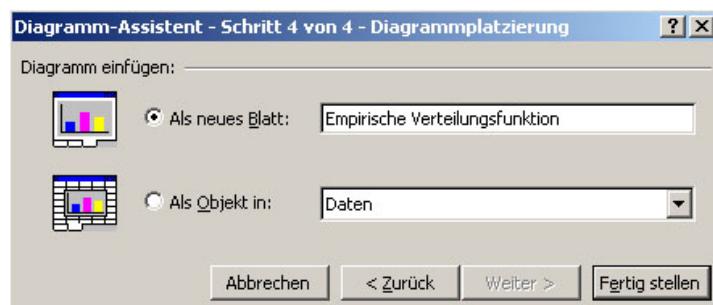


Abbildung 53:

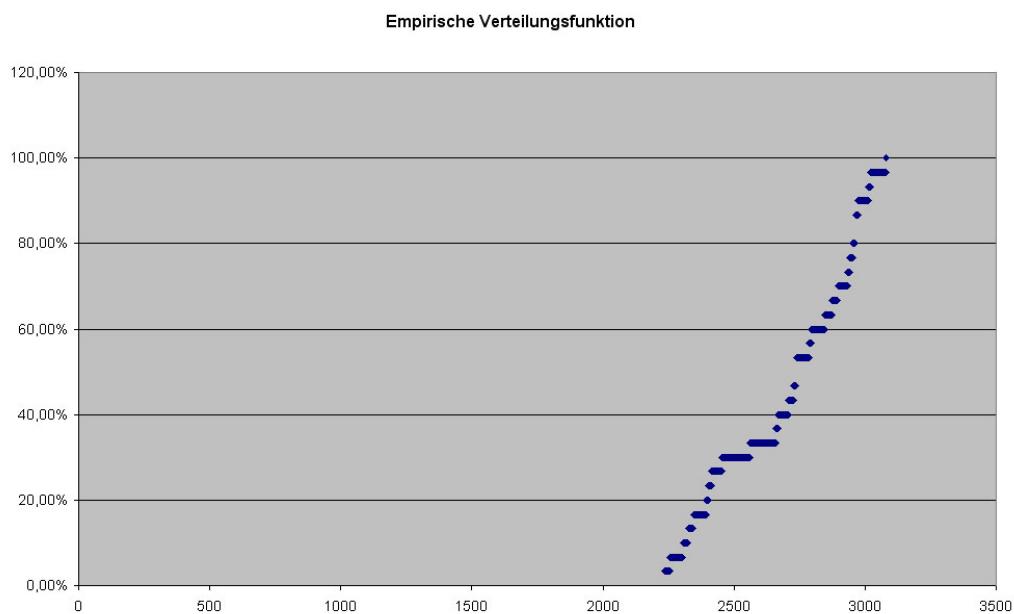


Abbildung 54:

4.3 Editieren der „Empirischen Verteilungsfunktion“

Wie man sehen kann, ist das Ergebnis nicht ganz zufriedenstellend, da die Skalierung der X- bzw. der Y-Achse nicht optimal ist, sowie die Symbole für die Datenpunkte geändert werden sollten.

1. Wir machen einen Doppelklick auf eine Beschriftung der X-Achse (z. B. 2200) und klicken beim Aufgehenden Dialogfenster „Achse formatieren“ auf den Reiter **SKALIEREN**. Wir stellen das *Minimum* auf „2200“ und das *Maximum* auf „3100“ ein und bestätigen mit **OK** (Abbildung 55). Das Ergebnis ist in Abbildung 56 zu sehen.

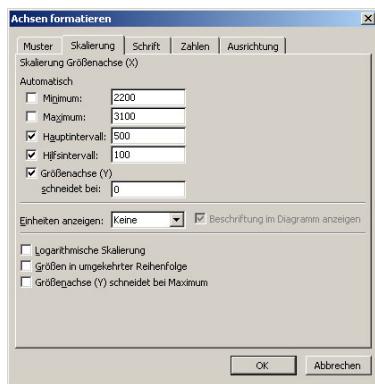


Abbildung 55:

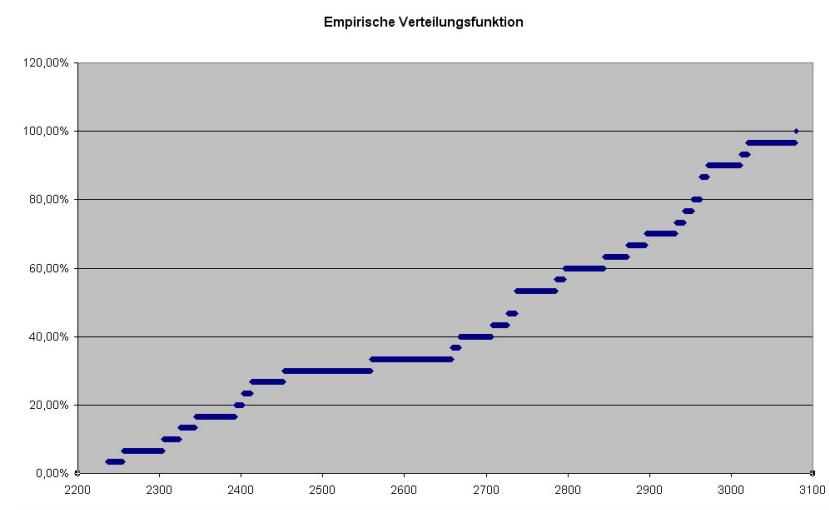


Abbildung 56:

2. Jetzt ändern wir noch die Symbole der Datenpunkte, indem wir einen Doppelklick auf einen Datenpunkt machen und beim Aufgehenden Dialogfenster „*Datenreihe formatieren*“ auf den Reiter **MUSTER** klicken. Wir stellen bei *Markierung* die *Art* auf den „–“ und bei *Größe* auf „2“ und bestätigen mit **OK** (Abbildung 57). Das Ergebnis ist in Abbildung 58 zu sehen.

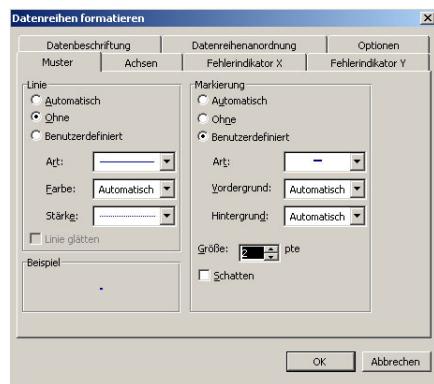


Abbildung 57:

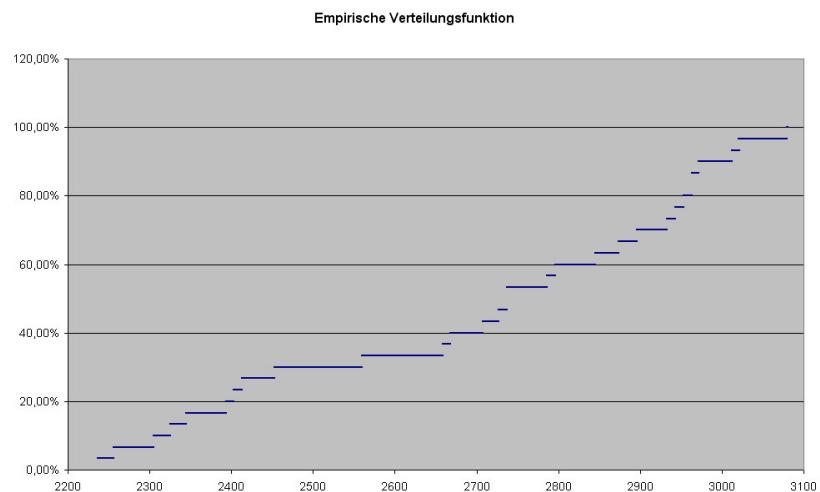


Abbildung 58:

5 Kasten-Bild (Boxplot)

Microsoft Excel beinhaltet keine Kasten-Bild Diagrammtypen. Anhand der folgenden Schritte wird aber gezeigt, wie eine „vereinfachte“ Form mit Hilfe des Kurs-Diagrammtyps erstellt werden kann. Siehe Abbildung 59, die einen schematisches Kasten-Bild darstellt.

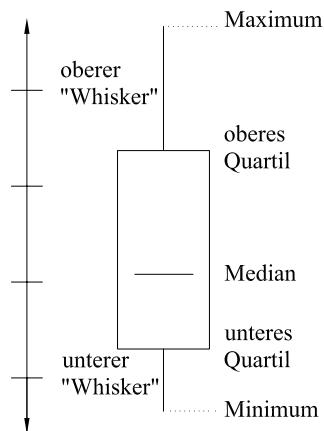


Abbildung 59:

5.1 Vorbereiten der Daten für des Kasten-Bild

Wir gehen davon aus, dass das Arbeitsblatt „BoxPlot.xls“ geöffnet vor uns am Bildschirm angezeigt wird. Im Tabellenblatt „Daten“ finden wir unsere Messdaten, welche Temperaturen im Jänner der Jahre 2005 – 2007 repräsentieren.

1. Wir geben den Daten im Bereich **B3:B33** den Namen „**temp_2005**“, dem Bereich **C3:C33** den Namen „**temp_2006**“ und dem Bereich **D3:D33** den Namen „**temp_2007**“. (Namensgebung von Daten siehe Abschnitt 1.1).
2. Wir öffnen ein neues Tabellenblatt und geben ihm den Namen „**Box-Plot**“.
3. Danach füllen wir die Zellen entsprechend Abbildung 60 aus.

	A	B	C	D
1		2005	2006	2007
2	Median			
3	1. Quartile			
4	Minimum (unterer "Whisker")			
5	Maximum (oberer "Whisker")			
6	3. Quartile			

Abbildung 60:

4. Für die Berechnung der einzelnen Zellen geben wir nun folgende Formeln ein:

- Zelle **B2**: $=\text{MEDIAN}(\text{temp_2005})$, welche den Median der Messdaten ermittelt.
- Zelle **B3**: $=\text{QUARTILE}(\text{temp_2005};1)$, welche das 1. Quartil der Messdaten ermittelt.
- Zelle **B4**: $=\text{MIN}(\text{temp_2005})$, welche das Minimum der Messdaten ermittelt.
- Zelle **B5**: $=\text{MAX}(\text{temp_2005})$, welche das Maximum der Messdaten ermittelt.
- Zelle **B6**: $=\text{QUARTILE}(\text{temp_2005};3)$, welche das 3. Quartil der Messdaten ermittelt.

Für den Bereich (**C2:D6**) der Jahre 2006 und 2007 gehen wir analog oben genannter Punkte vor, nur anstatt der Messdaten „temp_2005“ werden die Messdaten „temp_2006“ und „temp_2007“ verwendet. (Abbildung 61 und 62)

	A	B	C	D
1		2005	2006	2007
2	Median	$=\text{MEDIAN}(\text{temp_2005})$		
3	1. Quartile	$=\text{QUARTILE}(\text{temp_2005};1)$		
4	Minimum (unterer "Whisker")	$=\text{MIN}(\text{temp_2005})$		
5	Maximum (oberer "Whisker")	$=\text{MAX}(\text{temp_2005})$		
6	3. Quartile	$=\text{QUARTILE}(\text{temp_2005};3)$		

Abbildung 61:

	A	B	C	D
1		2005	2006	2007
2	Median	2,20	4,10	5,20
3	1. Quartile	-1,80	-7,20	3,05
4	Minimum (unterer "Whisker")	-8,90	-14,90	4,50
5	Maximum (oberer "Whisker")	8,80	2,40	11,70
6	3. Quartile	4,65	-0,30	8,90

Abbildung 62:

5.2 Erstellen des Kasten-Bildes

1. Wir markieren den Bereich **A1:D6** und öffnen den Diagramm-Assistenten entweder mittels: **EINFÜGEN → DIAGRAMM...** oder mit Hilfe des entsprechenden Icon  auf der Standard - Toolbar.
2. Als **Diagrammtyp** klicken wir **Kurs** an und als **Diagrammuntertyp** wählen wir den Vierten aus. Die folgende Erklärung wird unterhalb des Diagrammuntertyps angezeigt: **Volumen-Öffnung-Hoch-Tief-Geschlossen. Erfordert fünf Werte als Serie in dieser Reihenfolge** (Abbildung 63).

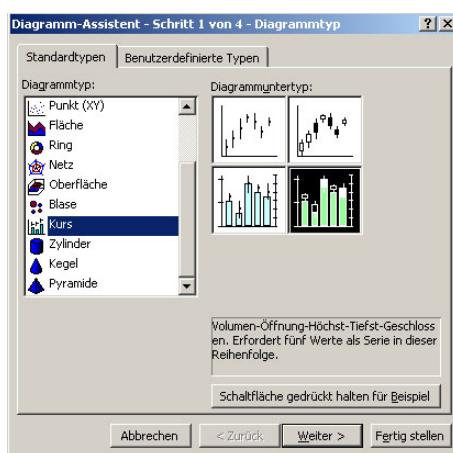


Abbildung 63:

3. Anschließend klicken wir auf **Weiter >**.

4. Jetzt klicken wir auf der Registerkarte **Datenbereich** unter **Reihe in** auf **Zeilen**, und klicken dann auf **Weiter >** (Abbildung 64).

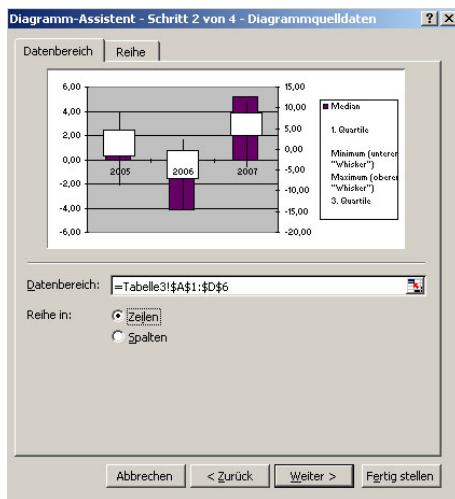


Abbildung 64:

5. In der Registerkarte **Legende** deaktivieren wir das Kontrollkästchen **Legende anzeigen** und in der Registerkarte **Achsen** unter **Sekundärachse** das Kontrollkästchen **Größenachse (Y)**. Danach klicken wir auf **Fertig stellen** (Abbildung 65).

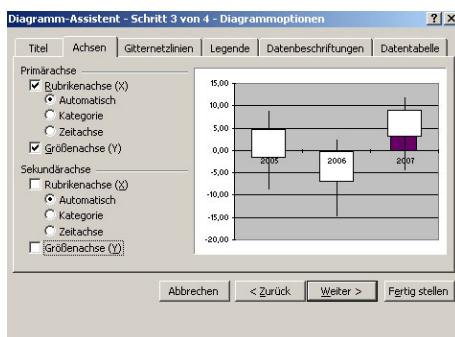


Abbildung 65:

6. Nun klicken wir einmal auf eine der farbigen Spalten, um die Datenreihe auszuwählen (Keine weiße Spalte auswählen!). Mit der rechten

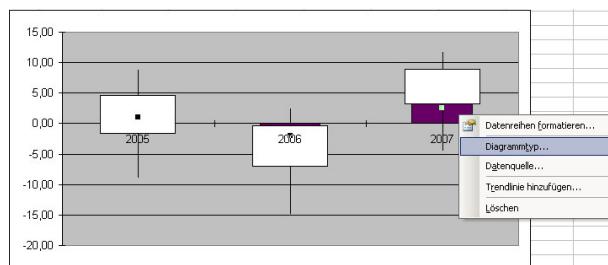


Abbildung 66:

Maustaste öffnen wir das Kontextmenü und wählen den **Diagrammtyp ...** aus (Abbildung 66).

7. Als **Diagrammtyp** wählen wir **Linie**, und klicken anschließend auf **OK** (Abbildung 67). Im Diagramm wird jetzt eine Linie angezeigt, die die drei weißen Boxen verbindet.

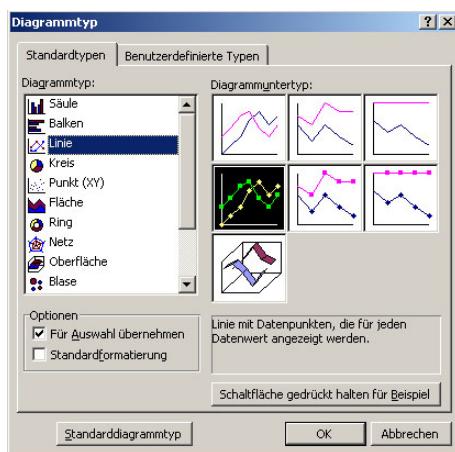


Abbildung 67:

8. Wir klicken einmal auf die Linie und mit der rechten Maustaste öffnen wir das Kontextmenü, indem wir den Punkt **Datenreihe formatieren ...** auswählen (Abbildung 68).
9. In der Registerkarte **Muster** klicken wir unter **Linie** auf **Ohne** und unter **Markierung** auf **Benutzerdefiniert**. In der Liste **Art** wählen wir den längeren horizontalen Strich aus (–) (Abbildung 69).

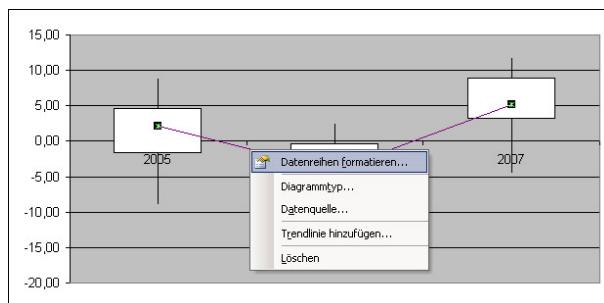


Abbildung 68:

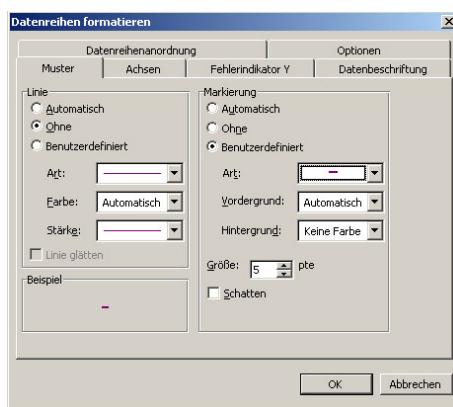


Abbildung 69:

5.3 Kasten-Bild-Diagrammoption einstellen

Nun wollen wir noch mit Hilfe des Add-In-Managers das erstellte Diagramm für zukünftige Verwendungen in den Diagramm-Assistenten aufnehmen.

1. Wir klicken auf das Diagramm mit der rechten Maustaste und wählen im aufgeklappten Kontextmenü **Diagrammtyp** aus (Abbildung 70).
2. Im Reiter **benutzerdefinierte Typen** wählen wir unter **Auswählen aus** den Punkt **Benutzerdefiniert** an und klicken auf **Hinzufügen** (Abbildung 71).
3. In die Dialogbox geben wir einen sinnvollen Namen sowie eine Bezeichnung ein - siehe Abbildung 72.

Nachdem der neue Diagrammtyp gespeichert wurde, kann man ihn wie einen herkömmlichen Excel Diagrammtyp verwenden.

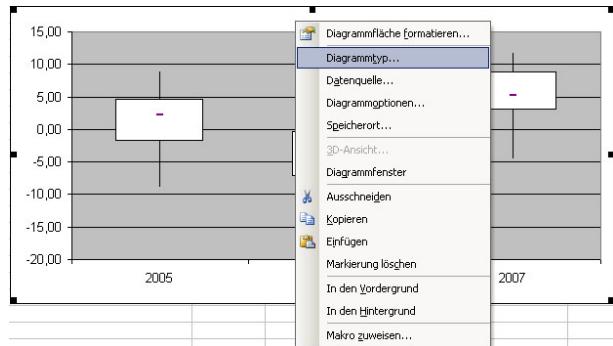


Abbildung 70:

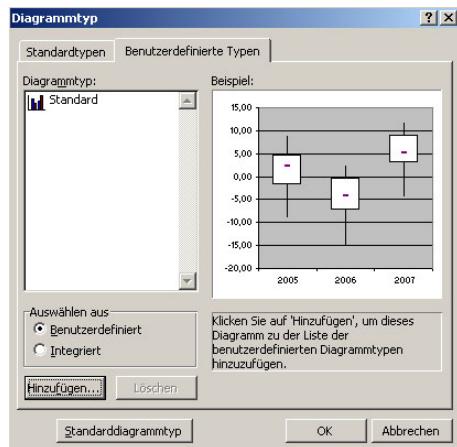


Abbildung 71:

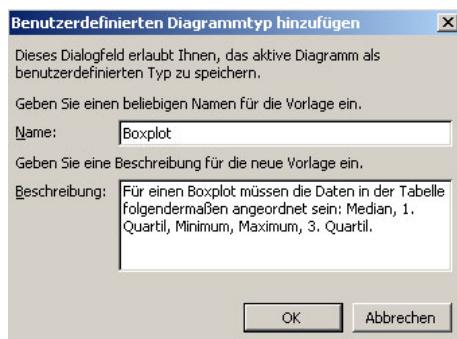


Abbildung 72:

6 Lineare Regression

6.1 Vorbereiten der Daten

Wir gehen davon aus, dass Arbeitsblatt „LineareRegression.xls“ geöffnet vor uns am Bildschirm angezeigt wird. Im Tabellenblatt „Daten“ finden wir unsere zwei verbundenen Stichproben.

1. Wir geben den Daten im Bereich **A2:A12** den Namen „**datenX**“ und dem Bereich **B2:B12** den Namen „**datenY**“ (Namensgebung von Daten siehe Abschnitt 1.1).

6.2 Darstellen der Stichproben in einem Koordinatensystem

1. Wir markieren den Bereich **A2:B13** und öffnen den Diagramm-Assistenten entweder mittels: **EINFÜGEN → DIAGRAMM...** oder mit Hilfe des entsprechenden Icon  auf der Standard - Toolbar.
2. Als **Diagrammtyp** klicken wir **Punkt (XY)** an und als **Diagrammuntertyp** wählen wir den Ersten aus. Die folgende Erklärung wird unterhalb des Diagrammuntertyps angezeigt: **Punkte. Vergleicht Werte paarweise** (Abbildung 73).

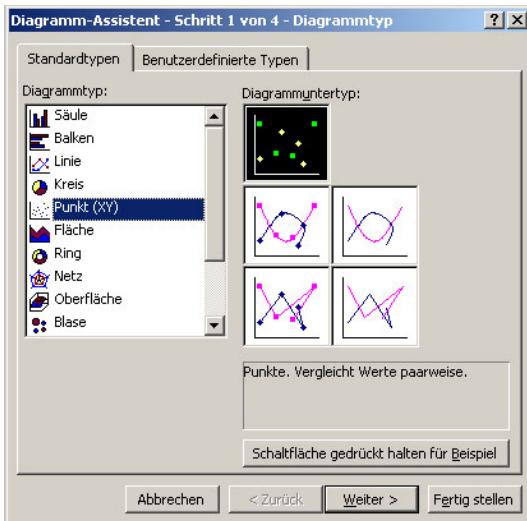


Abbildung 73:

3. Anschließend zweimal auf **Weiter >** klicken.

4. Festlegen der Überschrift sowie der Achsenbeschriftungen.
5. Danach klicken wir auf **Fertig stellen** (siehe Abbildung 74).

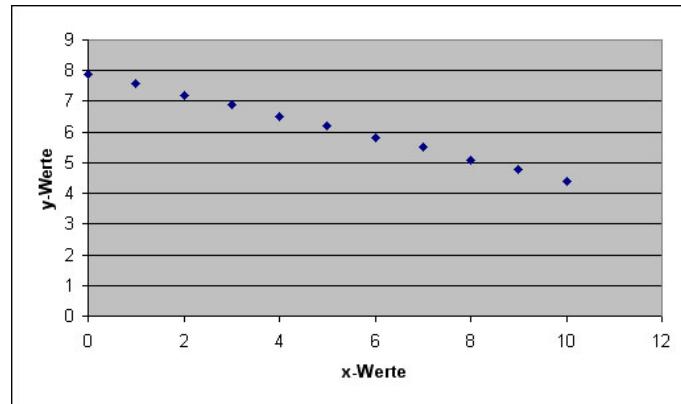


Abbildung 74:

6.3 Darstellung der Regressionsgeraden

1. Wir klicken einmal mit der linken Maustaste auf einen Punkt der Datenreihe, um ihn auszuwählen. Mit der rechten Maustaste öffnen wir das Kontextmenü und wählen **Trendlinie hinzufügen ...** aus (Abbildung 75).

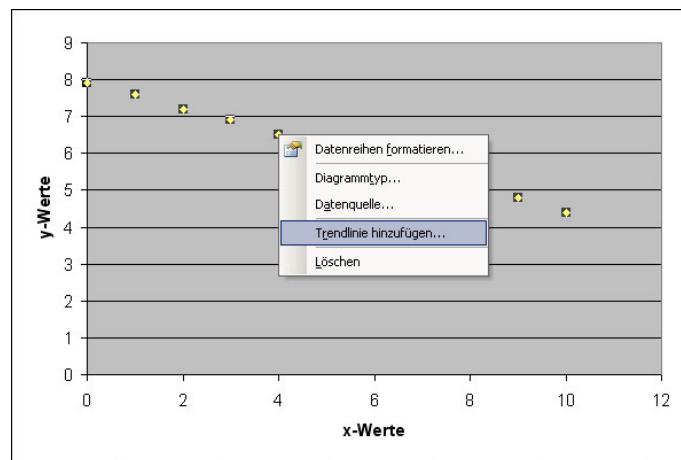


Abbildung 75:

2. Als Trend-/Regressionstyp wählen wir den Linearen (Abbildung 76) und klicken auf die Registerkarte **Optionen**.

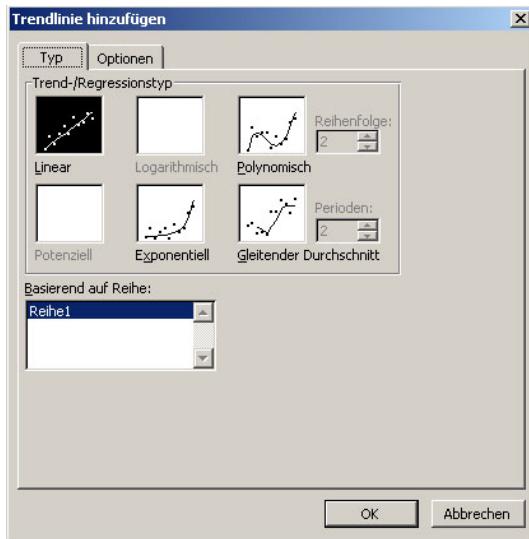


Abbildung 76:

3. In der Registerkarte **Optionen** markieren wir das Kästchen *Gleichung in Diagramm darstellen* und *Bestimmtheitsmaß im Diagramm darstellen* (Abbildung 77).

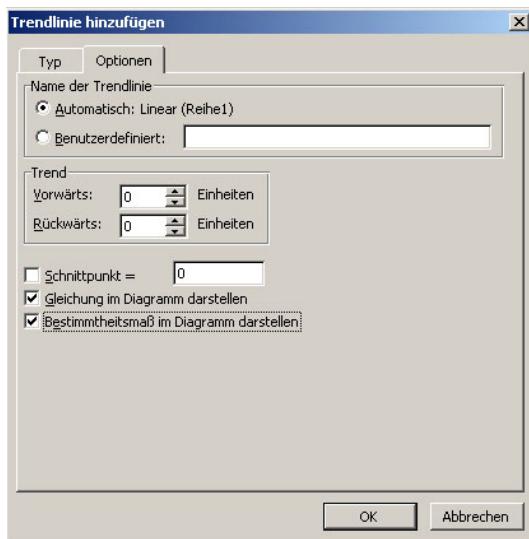


Abbildung 77:

4. Anschließend klicken wir auf **OK** und können die Trendlinie samt Gleichung und Bestimmtheitsmaß im Diagramm betrachten (Abbildung 78).

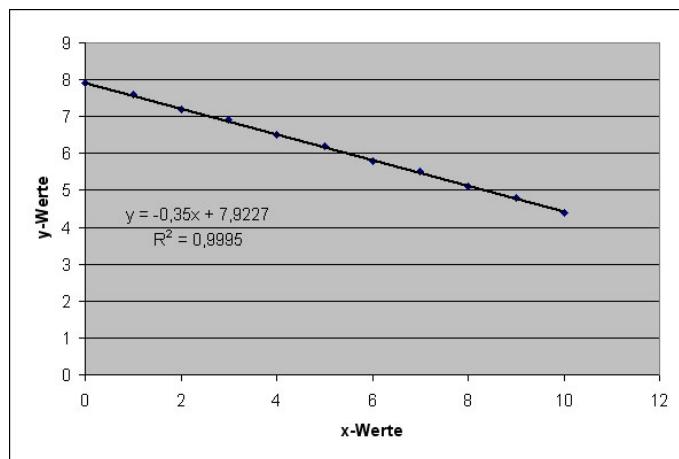


Abbildung 78:

6.4 Berechnung der Parameter der Regressionsgeraden

1. Wir öffnen das Tabellenblatt „Auswertung“.
2. Für die Berechnung der einzelnen Zellen geben wir nun folgende Formeln ein:
 - Zelle **C2**: $=\text{MITTELWERT}(\text{datenX})$, welche den Mittelwert \bar{x}_n berechnet.
 - Zelle **C3**: $=\text{MITTELWERT}(\text{datenY})$, welche den Mittelwert \bar{y}_n berechnet.
 - Zelle **C4**: $=\text{VARIANZ}(\text{datenX})$, welche die Stichprobenvarianz s_x^2 berechnet.
 - Zelle **C5**: $=\text{VARIANZ}(\text{datenY})$, welche die Stichprobenvarianz s_y^2 berechnet.
 - Zelle **C6**: $=\text{KOVAR}(\text{datenX}; \text{datenY})$, welche die Kovarianz s_{xy} der Stichprobe berechnet.
 - Zelle **C8**: $=\text{BESTIMMTHEITSMASS}(\text{datenY}; \text{datenX})$, welche das Bestimmtheitsmaß \tilde{r}^2 berechnet.
 - Zelle **C9**: $=\text{KORREL}(\text{datenX}; \text{datenY})$, welche den Korrelationskoeffizienten \tilde{r} berechnet.

(Abbildung 79). Durch den hohen Wert des Bestimmtheitsmaßes \tilde{r}^2 von 0,999 in Zelle C8, kann man einen linearen Zusammenhang zwischen den Zufallsgrößen X und Y vermuten.

	A	B	C
1	Bezeichnung		EXCEL
2	Mittelwert der x-Werte:	\bar{x}_x	=MITTELWERT(datenX)
3	Mittelwert der y-Werte:	\bar{y}_y	=MITTELWERT(datenY)
4	Stichprobenvarianz der x-Werte:	s_x^2	=VARIANZ(datenX)
5	Stichprobenvarianz der y-Werte:	s_y^2	=VARIANZ(datenY)
6	Kovarianz der Stichprobe:	s_{xy}	=KOVAR(datenX;datenY)
7			
8	Bestimmtheitsmaß:	\tilde{r}^2	=BESTIMMTHEITSMASS(datenY;datenX)
9	Korrelationskoeffizient:	\tilde{r}	=KORREL(datenX;datenY)
10			
11	Annahme:	$Y = kX + d$ (d.h. Y unabhängig, X abhängige Variable)	
12		d	
13	Achsenabschnitt	d	
14	Steigung	k	
15			
16	Annahme:	$X = k'Y + d'$ (d.h. X unabhängig, Y abhängige Variable)	
17		d'	
18	Achsenabschnitt	d'	
19	Steigung	k'	

Abbildung 79:

3. Nun können wir zwei Fälle unterscheiden:

1. Fall: X ist die unabhängige Variable und Y die abhängige Variable.

Dann geht man wie folgt vor: In die Zelle C13 schreiben wir folgende Formel $=\text{ACHSENABSCHNITT}(\text{datenY};\text{datenX})$, um den Ordinatenabschnitt (Achsenabschnitt) d zu erhalten und in Zelle C14 schreiben wir $=\text{STEIGUNG}(\text{datenY};\text{datenX})$, um die Steigung k der Regressionsgerade zu erhalten.

Die Regressionsgerade lautet somit: $y(x) = -0,35x + 7,923$ (Abbildung 80).

	A	B	C
1	Bezeichnung		EXCEL
2	Mittelwert der x-Werte:	\bar{x}_x	5,000
3	Mittelwert der y-Werte:	\bar{y}_y	6,173
4	Stichprobenvarianz der x-Werte:	s_x^2	11,000
5	Stichprobenvarianz der y-Werte:	s_y^2	1,348
6	Kovarianz der Stichprobe:	s_{xy}	3,500
7			
8	Bestimmtheitsmaß:	\tilde{r}^2	0,999
9	Korrelationskoeffizient:	\tilde{r}	1,000
10			
11	Annahme: $Y = kX + d$ (d.h. X unabhängig, Y abhängige Variable)		
12			
13	Achsenabschnitt	d	=ACHSENABSCHNITT(datenY;datenX)
14	Steigung	k	=STEIGUNG(datenY;datenX)
15			
16	Annahme: $X = k'Y + d'$ (d.h. Y unabhängig, X abhängige Variable)		
17			
18	Achsenabschnitt	d'	
19	Steigung	k'	

Abbildung 80:

- 2. Fall:** Y ist die unabhängige Variable und X die abhängige Variable.
 Dann geht man wie folgt vor: In die Zelle **C18** schreiben wir folgende Formel $=\text{ACHSENABSCHNITT}(\text{datenX};\text{datenY})$, um den Ordinatenabschnitt (Achsenabschnitt) d' zu erhalten und in Zelle **C19** schreiben wir $=\text{STEIGUNG}(\text{datenX};\text{datenY})$, um die Steigung k' der Regressionsgerade zu erhalten.
 Die Regressionsgerade lautet somit: $y(x) = 2,856x + 22,627$ (Abbildung 81).

	A	B	C
1	Bezeichnung		EXCEL
2	Mittelwert der x-Werte:	\bar{x}_s	5,000
3	Mittelwert der y-Werte:	\bar{y}_s	6,173
4	Stichprobenvarianz der x-Werte:	s_x^2	11,000
5	Stichprobenvarianz der y-Werte:	s_y^2	1,348
6	Kovarianz der Stichprobe:	s_{xy}	3,500
7			
8	Bestimmtheitsmaß:	\tilde{r}^2	0,999
9	Korrelationskoeffizient:	\tilde{r}	1,000
10			
11	Annahme: $\tilde{Y} = k\tilde{X} + d$ (d.h. X unabhängig, Y abhängige Variable)		
12			
13	Achsenabschnitt	d	7,923
14	Steigung	k	0,350
15			
16	Annahme: $X = k'Y + d'$ (d.h. Y unabhängig, X abhängige Variable)		
17			
18	Achsenabschnitt	d'	$=\text{ACHSENABSCHNITT}(\text{datenX};\text{datenY})$
19	Steigung	k'	$=\text{STEIGUNG}(\text{datenX};\text{datenY})$

Abbildung 81: