

Erstellung von einfachen linearen Regressionsgleichungen mit grafischer Unterstützung

Mit einfachen Regressionsgleichungen kann man in einem definierten Teilmarkt wertbeeinflussende Abhängigkeiten des Mietwerts von bestimmten Merkmalen erkennen und grafisch aufbereitet für den Gutachtenleser anschaulich sichtbar machen.

Ausgangsbasis ist dazu entweder eine Auskunft aus der Kaufpreis-Sammlung oder eine aufbereitete und validierte Mietpreissammlung.

In dieser Anleitung wird eine Auswertung der Einfamilienhausmieten für Reihenhäuser mit einer Wohnfläche bis 125 m² im Stadtteil Köln-Neuehrenfeld und in vergleichbaren Stadtteilen im Jahr 2013 verwendet. Das zu bewertende Objekt hat 104 m² Wohnfläche, ist ein in den 90er Jahren saniertes Reihemittelhaus in guter Wohnlage aus dem Ursprungsbaujahr 1935. Das mietrechtliche Baujahr wurde mit mietrechtlichen Methoden auf 1983 geschätzt. Die Grundstücksfläche beträgt 165 m². Bewertungsstichtag ist der 01.07.2013. Da aus Neuehrenfeld keine Mieten für Objekte mit vergleichbaren Merkmalen vorhanden waren, wurden aus einer für ganz Köln ausgewerteten Datenbank die folgenden Mietwerte als Stichprobe gezogen:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	S	T	U	V	W	X	Y
1	Datum	STBZ	STB	OT#	LKR	Stadt/PLZ	OT	Straße	Zimmer	ObjTyp	BJ	REN	Bew BJ	WFL	GRU	NKM €/Mo	NKM €/m²	NK €/Mo	Warm €/m²	Interquantil	Konfident	STP	
2	14.01.2013	7	02	06	K	Köln 50997	Rondorf	Am Böding	5	DHH	2004			2004	116 m²	181 m²	1.136,80 €	9,80 €/m²	130,00 €	10,92 €/m²	9,80 €/m²	9,80 €/m²	GG
3	21.01.2013	7	02	06	K	Köln 50997	Rondorf	Waldkauz	5	RH					122 m²	300 m²	1.127,00 €	9,24 €/m²	120,00 €	10,84 €/m²	9,80 €/m²	9,80 €/m²	GG
4	04.02.2013	2	02	04	K	Köln 50968	Raderthal	Dransdorf	5	RH					122 m²	190 m²	1.127,00 €	9,24 €/m²	100,00 €	10,06 €/m²	9,24 €/m²	9,24 €/m²	STP
5	11.02.2013	7	02	10	K	Köln 50999	Sürth	Narzissen	4	DHH	1960	2010	2000	88 m²	278 m²	1.019,20 €	11,58 €/m²	65,00 €	12,32 €/m²	11,58 €/m²	11,58 €/m²	GG	
6	04.03.2013	7	02	09	K	Köln 50999	Weiß		4	DHH	1957	2012	2001	105 m²	278 m²	1.078,00 €	10,27 €/m²	120,00 €	11,41 €/m²	10,27 €/m²	10,27 €/m²	STP	
7	11.03.2013	7	02	10	K	Köln 50999	Sürth		4	EFH	2003		2003	113 m²		1.274,00 €	11,27 €/m²	200,00 €	13,04 €/m²	11,27 €/m²	11,27 €/m²	STP	
8	18.03.2013	7	02	06	K	Köln 50997	Rondorf		4	REH		2012	1996	110 m²	100 m²	1.175,02 €	10,68 €/m²	120,00 €	11,77 €/m²	10,68 €/m²	10,68 €/m²	GG	
9	25.03.2013	7	02	09	K	Köln 50999	Weiß		4	RH	1980		1980	104 m²	157 m²	1.470,00 €	14,13 €/m²	100,00 €	15,10 €/m²			GG	
10	08.04.2013	7	03	09	K	Köln 50859	Widdersdorf		4	RH				110 m²		1.151,50 €	10,47 €/m²	182,00 €	12,12 €/m²	10,47 €/m²	10,47 €/m²	STP	
11	08.04.2013	7	02	10	K	Köln 50999	Sürth	Am Feldra	4	RH	1980		1980	111 m²		1.029,00 €	9,27 €/m²	140,00 €	10,53 €/m²	9,27 €/m²	9,27 €/m²	GG	
12	15.04.2013	3	03	07	K	Köln 5085	Weiden		4	RH				100 m²		1.176,00 €	11,76 €/m²		11,76 €/m²	11,76 €/m²	GG		
13	15.04.2013	3	03	07	K	Köln 5085	Weiden		5	RH	1994	2009	2006	120 m²	350 m²	1.117,20 €	9,31 €/m²	180,00 €	10,81 €/m²	9,31 €/m²	9,31 €/m²	GG	
14	22.04.2013	7	03	09	K	Köln 50859	Widdersdorf		4	RH		2011	1995	102 m²		1.176,00 €	11,53 €/m²	120,00 €	12,71 €/m²	11,53 €/m²	11,53 €/m²	GG	
15	22.04.2013	7	02	09	K	Köln 50999	Weiß		4	EFH				120 m²	300 m²	1.176,00 €	9,80 €/m²	150,00 €	11,05 €/m²	9,80 €/m²	9,80 €/m²	GG	
16	22.04.2013	3	03	07	K	Köln 5085	Weiden		5	EFH				120 m²	350 m²	1.176,00 €	9,80 €/m²		9,80 €/m²	9,80 €/m²	GG		
17	29.04.2013	7	02	10	K	Köln 50999	Sürth		4	EFH	1998		1998	96 m²		1.764,00 €	18,38 €/m²	95,00 €	19,36 €/m²			STP	
18	03.05.2013	7	03	09	K	Köln 50859	Widdersdorf	Unter Lind	4	EFH	2008		2008	114 m²		989,80 €	8,68 €/m²	140,00 €	9,91 €/m²			GG	
19	06.05.2013	7	02	06	K	Köln 50997	Rondorf		5	DHH	1999		1999	124 m²	213 m²	1.205,40 €	9,72 €/m²	170,00 €	11,09 €/m²	9,72 €/m²	9,72 €/m²	STP	
20	13.05.2013	1	01	01	K	Köln 50678	Altstadt-Süd	Rosenstr.	4	EFH				110 m²		1.156,40 €	10,51 €/m²		10,51 €/m²	10,51 €/m²	10,51 €/m²	STP	
21	13.05.2013	7	03	09	K	Köln 50859	Widdersdorf		6	DHH				114 m²	597 m²	1.198,54 €	10,51 €/m²		10,51 €/m²	10,51 €/m²	10,51 €/m²	STP	
22	20.05.2013	3	03	07	K	Köln 5085	Weiden		3	BUN	1966	2004	1996	95 m²		931,00 €	9,80 €/m²	180,00 €	11,69 €/m²	9,80 €/m²	9,80 €/m²	GG	
23	20.05.2013	1	01	03	K	Köln 506	Altstadt-Nord		4	RH	1966	2009	2000	110 m²	290 m²	1.029,00 €	9,35 €/m²		9,35 €/m²	9,35 €/m²	9,35 €/m²	GG	
24	27.05.2013	7	03	09	K	Köln 50859	Widdersdorf		4	EFH	2006		2006	114 m²	300 m²	1.038,80 €	9,11 €/m²	140,00 €	10,34 €/m²	9,11 €/m²	9,11 €/m²	GG	
25	03.06.2013	7	03	09	K	Köln 50859	Widdersdorf		4	EFH				125 m²	125 m²	1.274,00 €	10,19 €/m²	120,00 €	11,15 €/m²	10,19 €/m²	10,19 €/m²	STP	
26	11.06.2013	7	02	09	K	Köln 50999	Weiß	Im Garten	4	RH	1997		1997	120 m²	180 m²	1.254,40 €	10,45 €/m²	120,00 €	11,45 €/m²	10,45 €/m²	10,45 €/m²	GG	
27	24.06.2013	7	03	09	K	Köln 50859	Widdersdorf		4	REH	1962	2013	2003	95 m²	297 m²	1.029,00 €	10,83 €/m²	100,00 €	11,88 €/m²	10,83 €/m²	10,83 €/m²	GG	

Eckwerte dieser Daten nach Stadtteilen

Köln, STBZ Name	STBZ	OT	Mittelwert	Median
101 Altstadt-Süd	01	01	10,08 €/m²	10,43 €/m²
103 Altstadt-Nord	01	03	10,39 €/m²	10,18 €/m²
104 Neustadt-Nord	01	04	9,98 €/m²	9,91 €/m²
203 Raderberg	02	03	10,34 €/m²	10,19 €/m²
205 Zollstock	02	05	10,61 €/m²	10,61 €/m²
206 Rondorf	02	06	10,01 €/m²	9,98 €/m²
207 Hahnwald	02	07	10,65 €/m²	10,45 €/m²
209 Weiß	02	09	9,98 €/m²	9,83 €/m²
210 Sürth	02	10	10,14 €/m²	9,94 €/m²
307 Weiden	03	07	9,85 €/m²	9,80 €/m²
308 Lövenich	03	08	9,54 €/m²	9,36 €/m²
309 Widdersdorf	03	09	9,85 €/m²	9,65 €/m²
402 Neuehrenfeld	04	02	10,21 €/m²	10,29 €/m²
502 Mauenheim	05	02	10,36 €/m²	10,29 €/m²
505 Weidenpesch	05	05	9,75 €/m²	9,66 €/m²

Köln, STBZ Name	STBZ	OT	Mittelwert	Median	Ranking
308 Lövenich	03	08	9,54 €/m²	9,36 €/m²	-6,8%
309 Widdersdorf	03	09	9,85 €/m²	9,65 €/m²	-3,9%
505 Weidenpesch	05	05	9,75 €/m²	9,66 €/m²	-3,8%
307 Weiden	03	07	9,85 €/m²	9,80 €/m²	-2,4%
209 Weiß	02	09	9,98 €/m²	9,83 €/m²	-2,1%
104 Neustadt-Nord	01	04	9,98 €/m²	9,91 €/m²	-1,3%
210 Sürth	02	10	10,14 €/m²	9,94 €/m²	-1,0%
206 Rondorf	02	06	10,01 €/m²	9,98 €/m²	-0,6%
103 Altstadt-Nord	01	03	10,39 €/m²	10,18 €/m²	1,4%
203 Raderberg	02	03	10,34 €/m²	10,19 €/m²	1,5%
402 Neuehrenfeld	04	02	10,21 €/m²	10,29 €/m²	2,5%
502 Mauenheim	05	02	10,36 €/m²	10,29 €/m²	2,5%
101 Altstadt-Süd	01	01	10,08 €/m²	10,43 €/m²	3,9%
207 Hahnwald	02	07	10,65 €/m²	10,45 €/m²	4,1%
205 Zollstock	02	05	10,61 €/m²	10,61 €/m²	5,7%
Mittelwert			10,12 €/m²	10,04 €/m²	0,0%

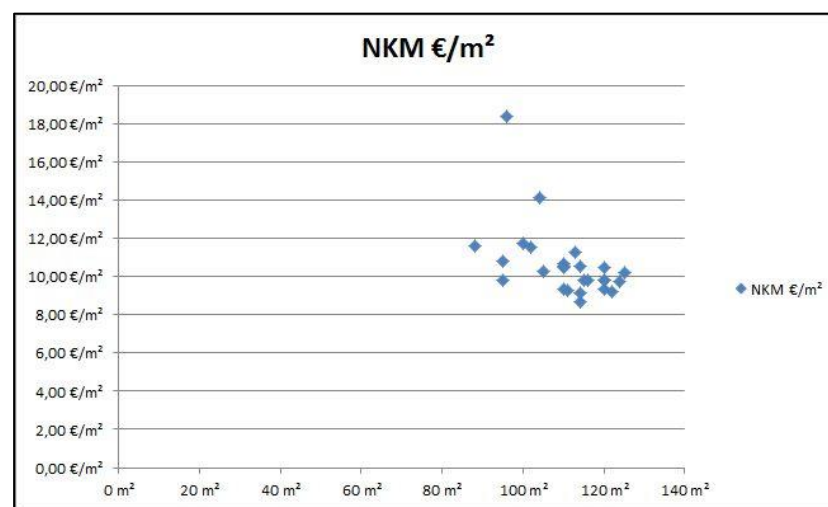
Die Auswahl der vergleichbaren Stadtteile wurde auf der Grundlage aller Mietwerte für Einfamilienhäuser aller Anbauarten vorgenommen. Aus dieser rund 150 Fälle umfassenden Stichprobe wurden dann die vorstehenden 26 Fälle selektiert, die mit den Merkmalen des Objekts am besten vergleichbar waren. Wichtig ist, dass man die Wertepaare von vornherein immer korrekt formatieren, dann werden diese Formatierungen automatisch in die Punktwolke übernommen.

WFL	NKM €/m ²
116 m ²	9,80 €/m ²
115 m ²	9,80 €/m ²
122 m ²	9,24 €/m ²
88 m ²	11,58 €/m ²
105 m ²	10,27 €/m ²
113 m ²	11,27 €/m ²
110 m ²	10,68 €/m ²
104 m ²	14,13 €/m ²
110 m ²	10,47 €/m ²
111 m ²	9,27 €/m ²
100 m ²	11,76 €/m ²
120 m ²	9,31 €/m ²
102 m ²	11,53 €/m ²
120 m ²	9,80 €/m ²
120 m ²	9,80 €/m ²
96 m ²	18,38 €/m ²
114 m ²	8,68 €/m ²
124 m ²	9,72 €/m ²
110 m ²	10,51 €/m ²
114 m ²	10,51 €/m ²
95 m ²	9,80 €/m ²
110 m ²	9,35 €/m ²
114 m ²	9,11 €/m ²
125 m ²	10,19 €/m ²
120 m ²	10,45 €/m ²
95 m ²	10,83 €/m ²

Zur Erläuterung der Vorgehensweise der Erstellung einer einfachen Regressionsanalyse wird das Wertepaar WFL und NKM €/m² verwendet:

Diese Matrix wird mit der Maus markiert und in die Zwischenablage kopiert.

Dann wird über die Menüpunkte <Einfügen><Diagramme><Punktwolke> die erste Punktwolken-Vorlage ausgewählt:

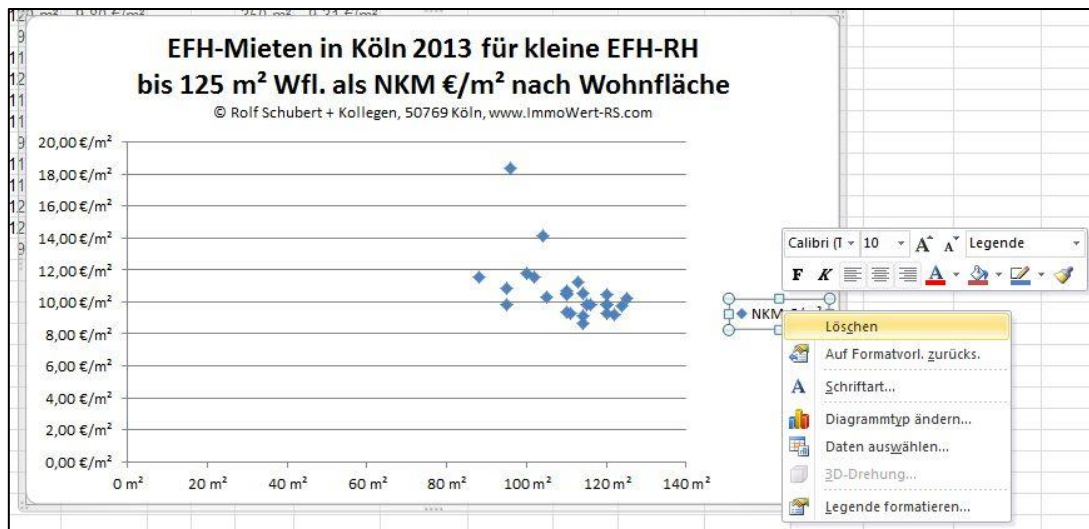


Nun sollte man zunächst die Überschrift erklärend festlegen. Es hat sich bewährt auch immer einen Copyright-Vermerk anzubringen, weil man

a) zu seinen Auswertungen stehen und

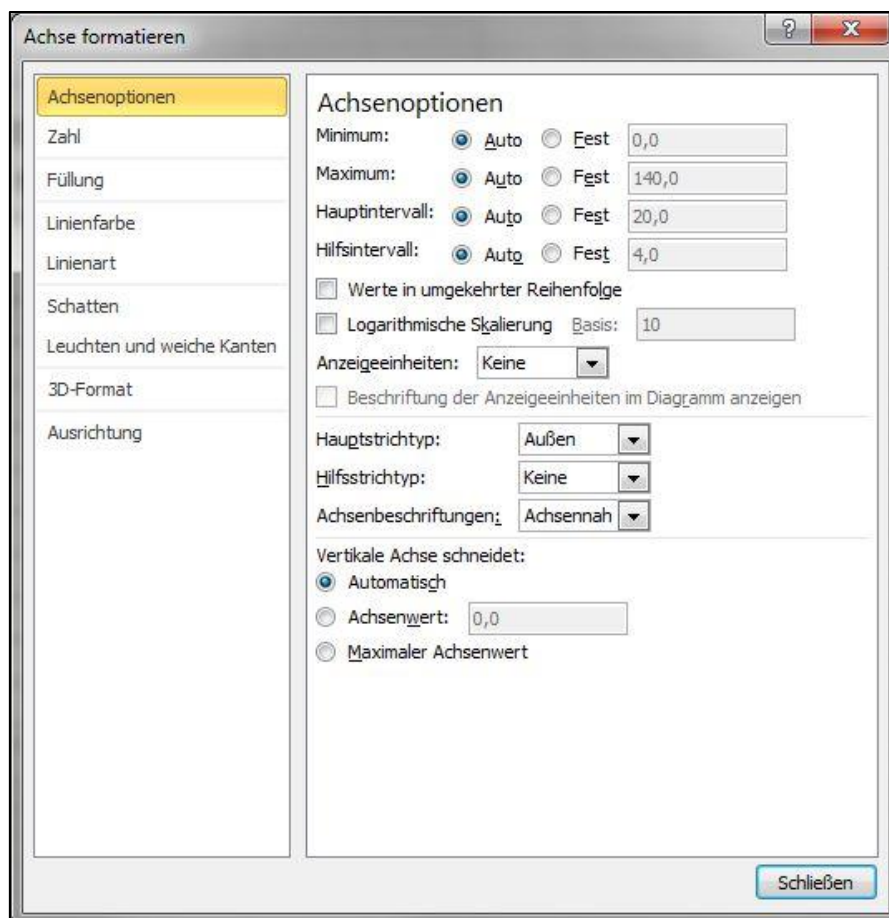
b) verhindern möchte, dass jemand diese Informationen aus dem Zusammenhang gerissen an anderer Stelle verwendet.

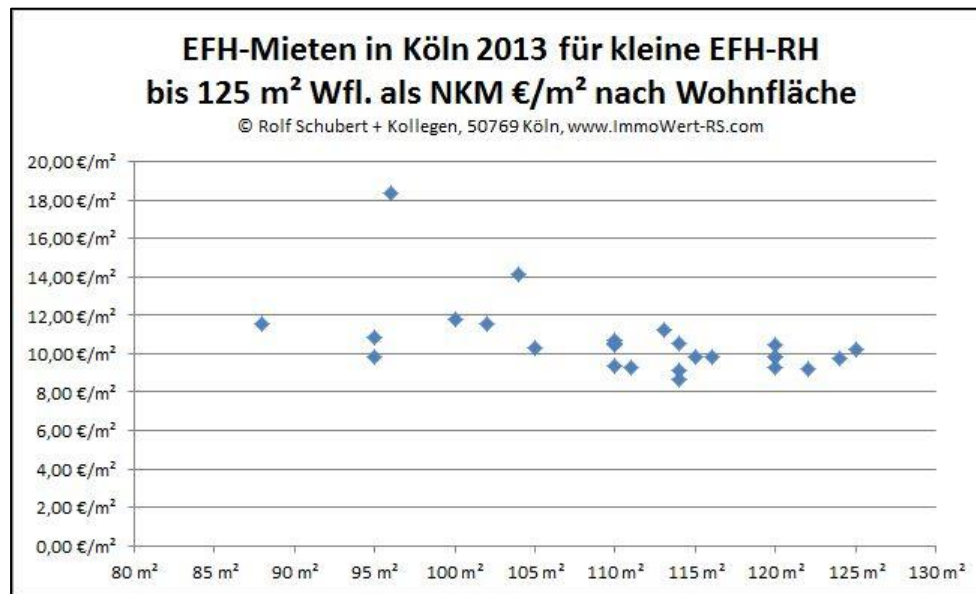
Die Aussagen einer solchen Regressionsgleichung gelten generell nur für die selektierte Stichprobe und können in anderen Zusammenhängen unzutreffend sein.



Wenn die Überschrift aussagekräftig ist kann man die Legende löschen. Dazu klickt man mit der rechten Maustaste auf die Legende und dann im Pop-Up-Menü auf <Löschen>.

Da es für das Verständnis ganz praktisch ist, wenn unnötiger Leerraum entfernt wird, kann man die WFL-Skala entsprechend straffen. Dazu klickt man mit der rechten Maustaste auf die untere WFL-Skala und wählt im Pop-UP-Menü <Achse formatieren> aus. Es öffnet sich der Dialog <Achse formatieren>. Hier trägt man bei Minimum den Festwert 80 ein. Das ist auch keine Verzerrung, denn RH mit wesentlich kleineren WFL gibt es nicht.

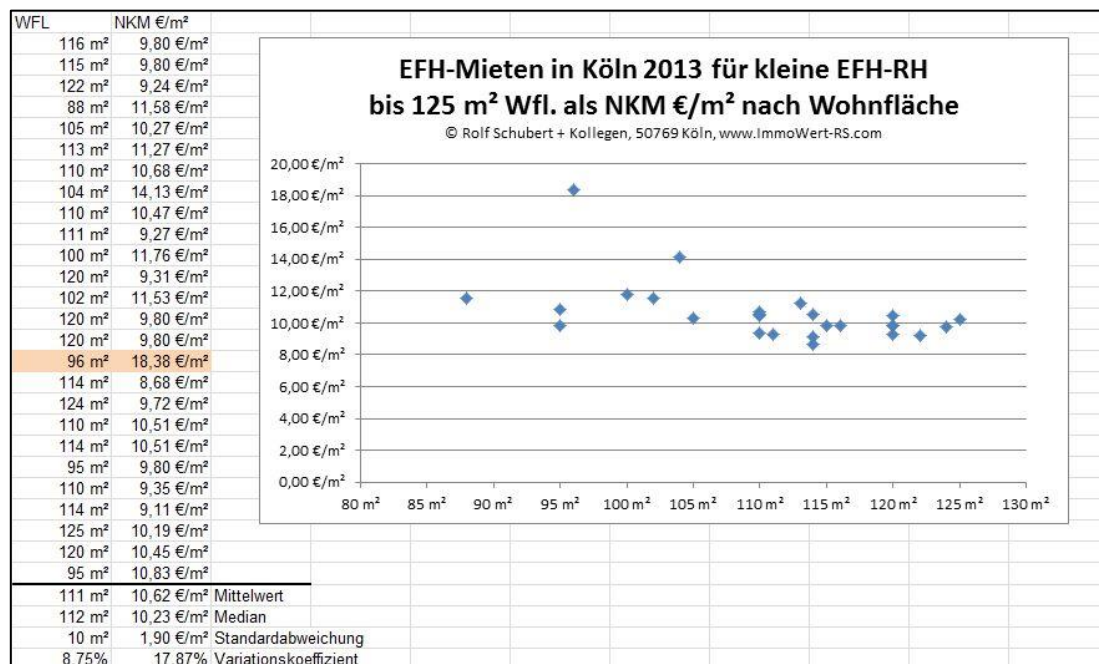




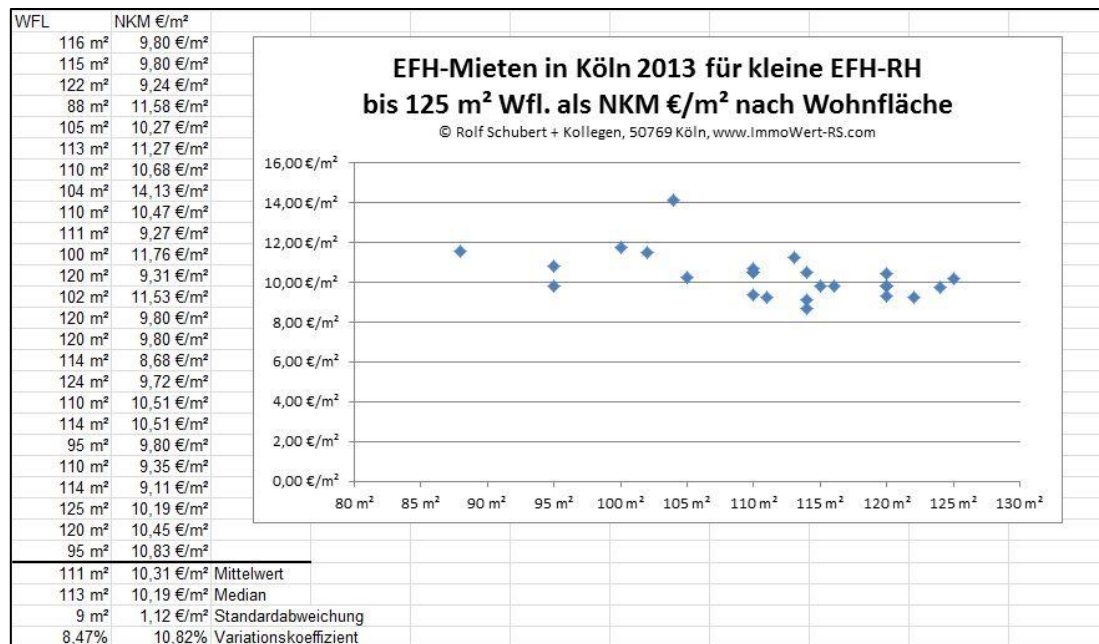
Anders sieht es bei der vertikalen Skala aus. Hier sollte man, wenn man den Skalenbereich schon ändern möchte, im gesamten Gutachten den gleichen Bereich einheitlich ändern.

Wenn man diese Punktwolke sachverständig betrachtet, dann fällt einem sofort auf, dass ein Punkt erheblich außerhalb der anderen Punkte liegt. Das kann korrekt sein, deutet aber auf einen wahrscheinlichen Ausreißer hin. Es empfiehlt sich, diesen einen Wert aus der Stichprobe zu entfernen.

Dazu eignet sich die Darstellung einer Punktwolke sehr gut. Damit man die beiden Stichproben miteinander vergleichen können, fügt man unter den Wertepaaren die statistischen Kennwerte Mittelwert, Median, STABW.S und den Variationskoeffizienten an.



Der gesamte Bereich einschließlich der Datengrundlage wird in die Zwischenablage kopiert und unterhalb erneut in die Tabelle eingefügt. Dadurch bleiben die Bezüge erhalten. Man braucht nur das eine Wertepaar zu entfernen und erhält folgende neue Grafik:

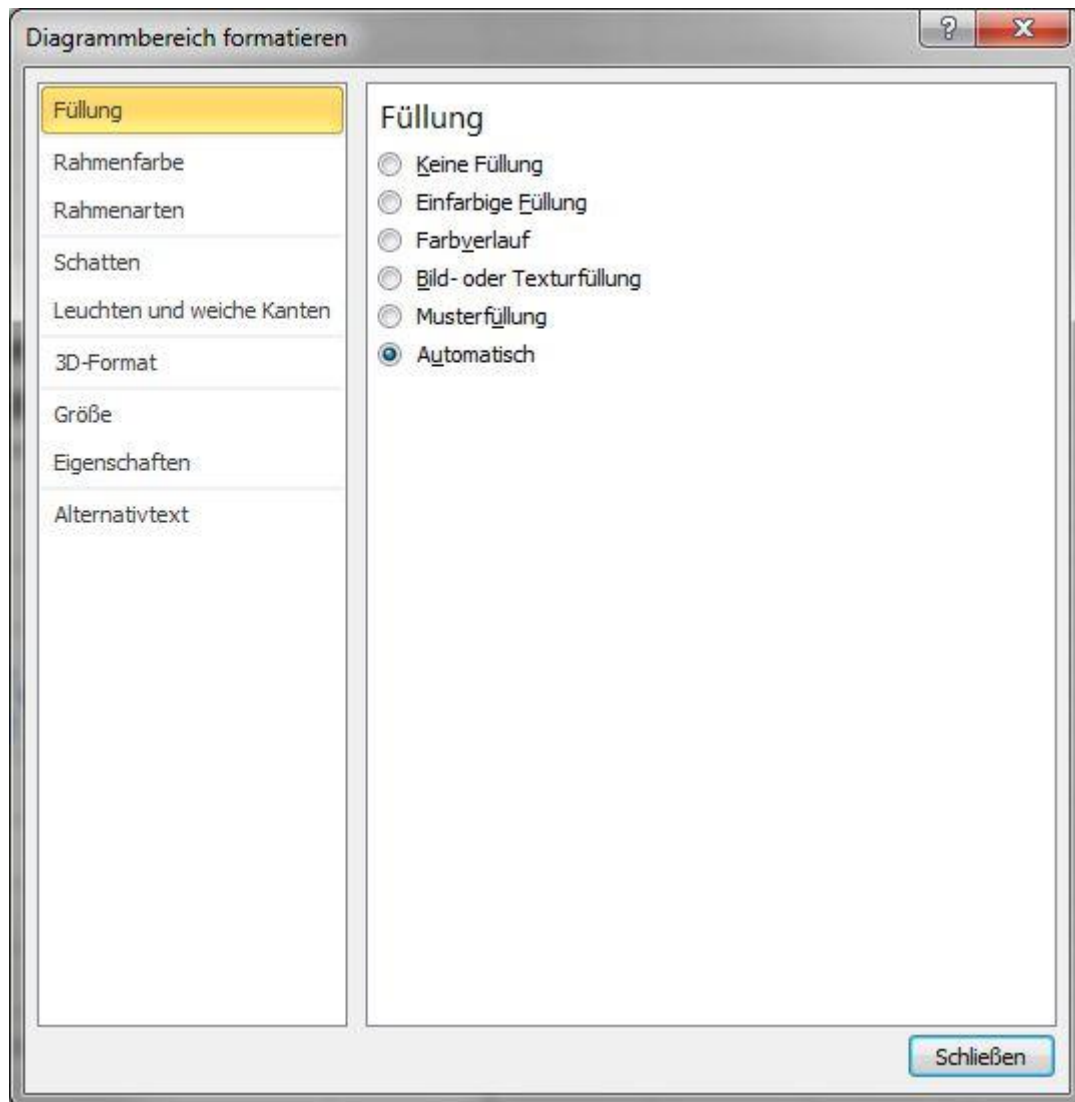


Das Ergebnis dieser Bereinigung ist offensichtlich: Das Bild der Punktwolke ist deutlich homogener, der Mittelwert hat sich von 10,62 auf 10,31 reduziert, das sind immerhin 3 %. Der Median, der nicht so stark auf Ausreißer reagiert hat sich von 10,23 €/m² auf 10,19 €/m² reduziert, das sind 0,4 %. Die Standardabweichung hat sich von 1,90 €/m² auf 1,12 €/m² reduziert, das sind über 40 % weniger und der Variationskoeffizient hat sich von 17,87 % auf 10,82 % verbessert, das sind ebenfalls etwa 40 %.

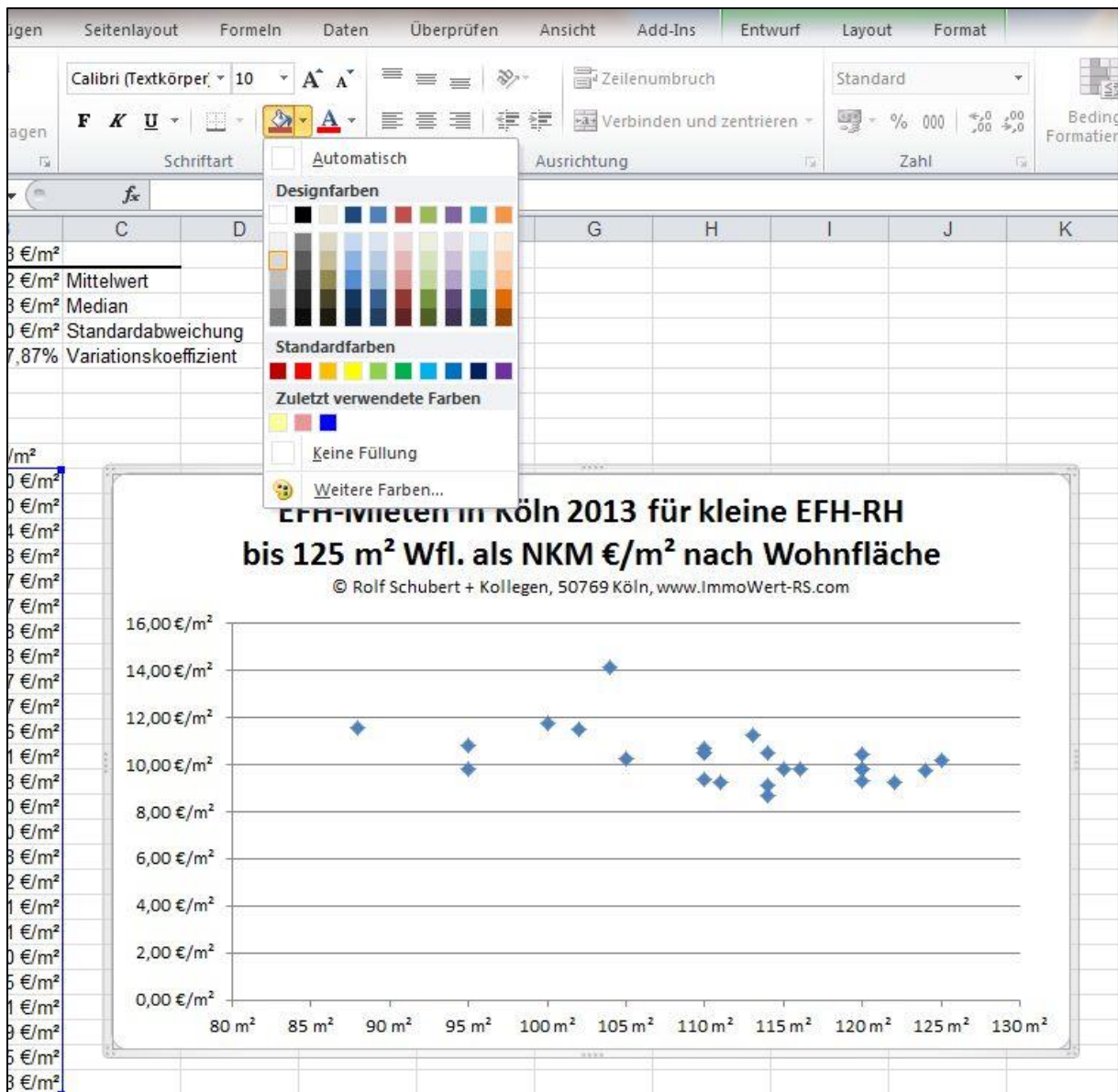
Das zeigt anschaulich, wie stark ein einzelner abweichender Wert die Aussage einer Statistik verzerren kann.

Nun geht es mit der bereinigten Stichprobe weiter.

Aus optischen Gründen fügt man einen leichten grauen Hintergrund ein. Dazu klickt man in den weißen Hintergrund links neben der Überschrift und wählt den Dialog <Diagrammbereich formatieren> aus, dann den Punkt <Füllung> und wählt schließlich eine einfarbige Füllung mit 25 % grau aus.

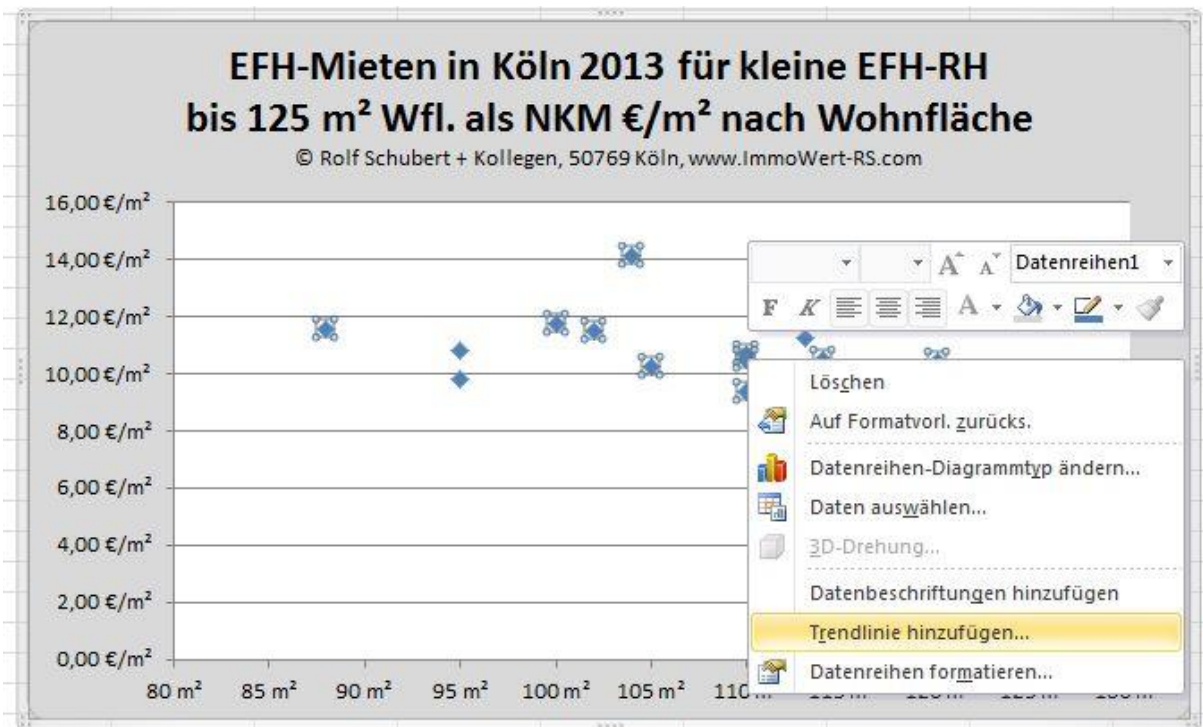


Alternativ reicht es auch, nach Klick in den Hintergrund der Grafik im Startmenü den Farbeimer anzuklicken, auch dort kann die helle Graufäche auswählen und fertig ist die Grafik.



Es ist in Excel möglich, in dieser Punktwolke eine Trendkurve zu erzeugen, d.h. eine lineare Regressionskurve zu erzeugen, einschließlich der Formel und des Bestimmtheitsgrads.

Dazu klickt man mitten in der Punktwolke mit der rechten Maustaste auf einen der Punkte und erhält ein Pop-Up-Menü. Dort wählt man <Trendlinie hinzufügen> aus.



Es erscheint der Dialog <Trendlinie formatieren>, der uns mehrere Möglichkeiten eröffnet.

Trendlinie formatieren

Trendlinienoptionen

Linienfarbe
Linienart
Schatten
Leuchten und weiche Kanten

Trendlinienoptionen

Trend-/Regressionstyp

Exponential
 Linear
 Logarithmisch
 Polynomisch Reihenfolge: 2
 Potenz
 Gleitender Durchschnitt Zeitraum: 2

Name der Trendlinie

Automatisch: Linear (Datenreihen1)
 Benutzerdefiniert:

Prognose

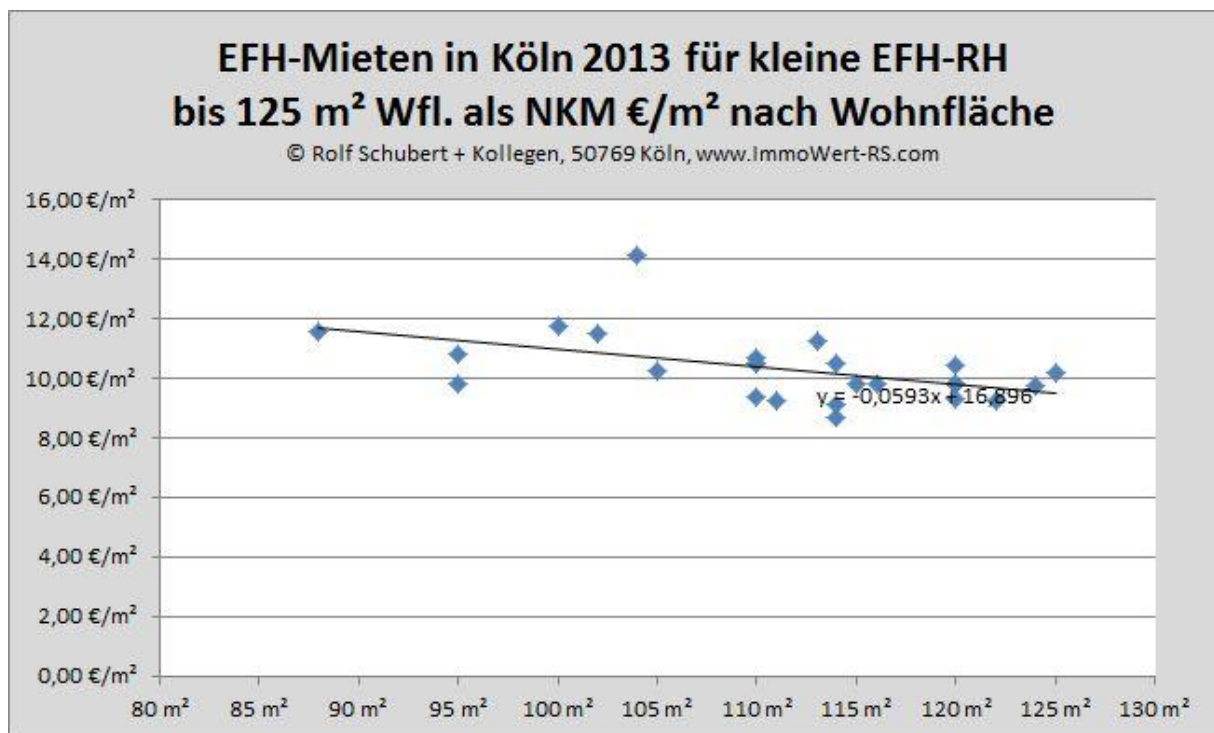
Weiter: Punkte
Zurück: Punkte

Schnittpunkt =
 Formel im Diagramm anzeigen
 Bestimmtheitsmaß im Diagramm darstellen

Schließen

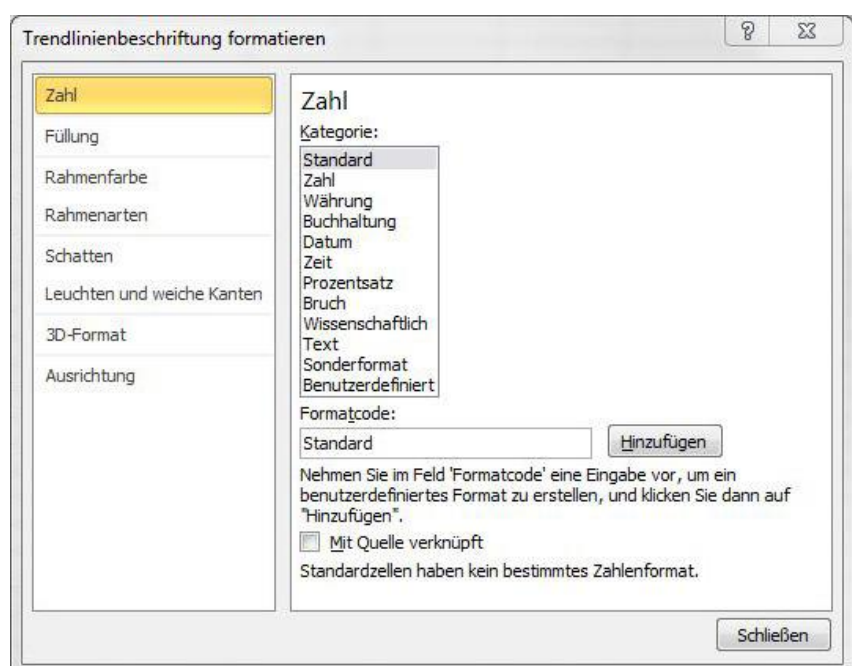
Standardmäßig ist die lineare Trendlinie, d.h. die Kurve für die einfache lineare Regressionsgleichung, d.h. der ausgleichende Mittelwert aller Punkte in Abhängigkeit von der jeweiligen Wohnfläche als Gerade angezeigt. Wir erkennen eine wertmäßige Abhängigkeit, wonach größere Wohnflächen zu niedrigeren relativen Mietwerten führen.

Damit man diese „Gesetzmäßigkeit“ später im Gutachten verwenden kann, lässt man die Formel im Diagramm anzeigen. Dazu klickt man auf das Häkchen vor „Formel im Diagramm anzeigen“. Die Formel wird immer ganz nahe an die Trendkurve gelegt. Das sieht dann so aus:



Man kann die Formel anklicken und in einen freien Bereich der Grafik ziehen. Dabei hat sich ein weiterer Dialog geöffnet, der die Möglichkeit der Formatierung der Trendlinienbeschriftung eröffnet.

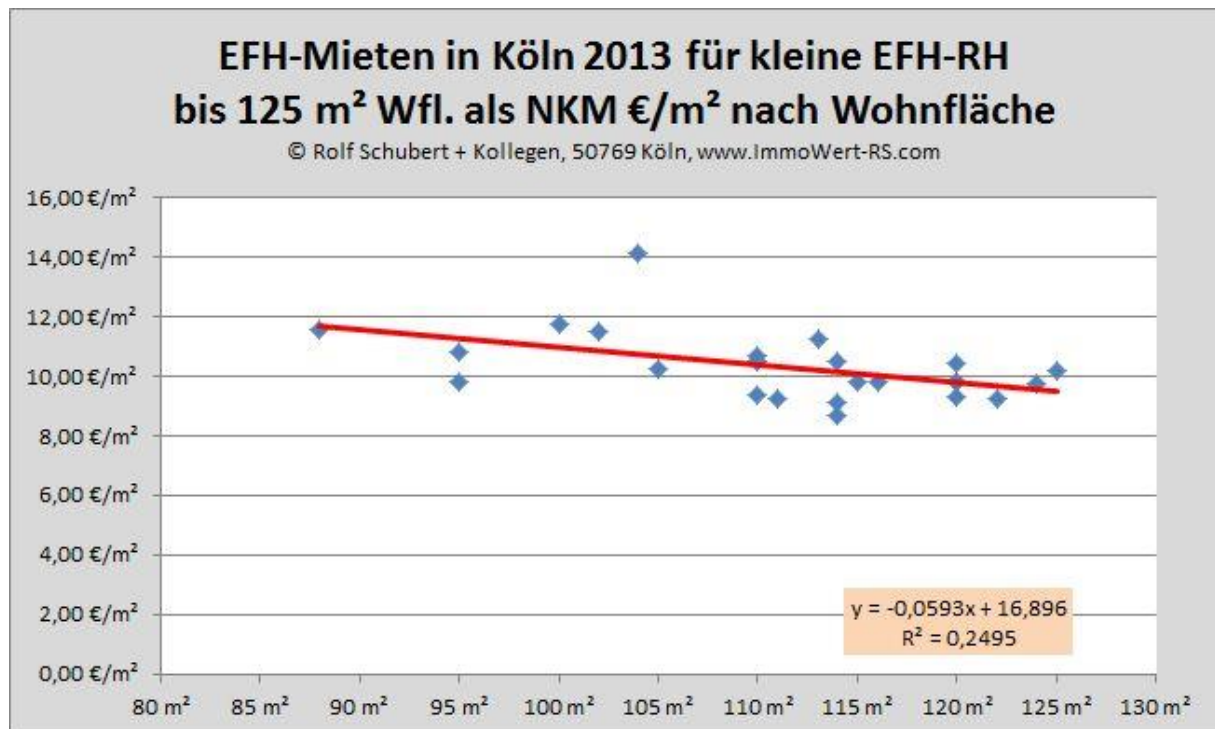
Hier wählt man Füllung aus und fügt passend zur Farbe der Trendlinie einen hellroten Hintergrund ein.



Wenn man diese Trendlinie erneut anklickt öffnet sich wieder der bereits bekannte Dialog <Trendlinie formatieren>. Hier klickt man nun den untersten Haken an <Bestimmtheitsmaß im Dialog darstellen>. Das Bestimmtheitsmaß könnte man benutzen, um die Aussagekraft dieser Trendlinie zu beurteilen. Siehe dazu das entsprechende Seminarscript des Verfassers.

Über den Dialog <Trendlinie formatieren> kann man die Linienfarbe, z.B. in einfarbig rot und die Linienart, z.B. in 2 pt. ändern.

Die fertige Grafik sieht dann so aus:

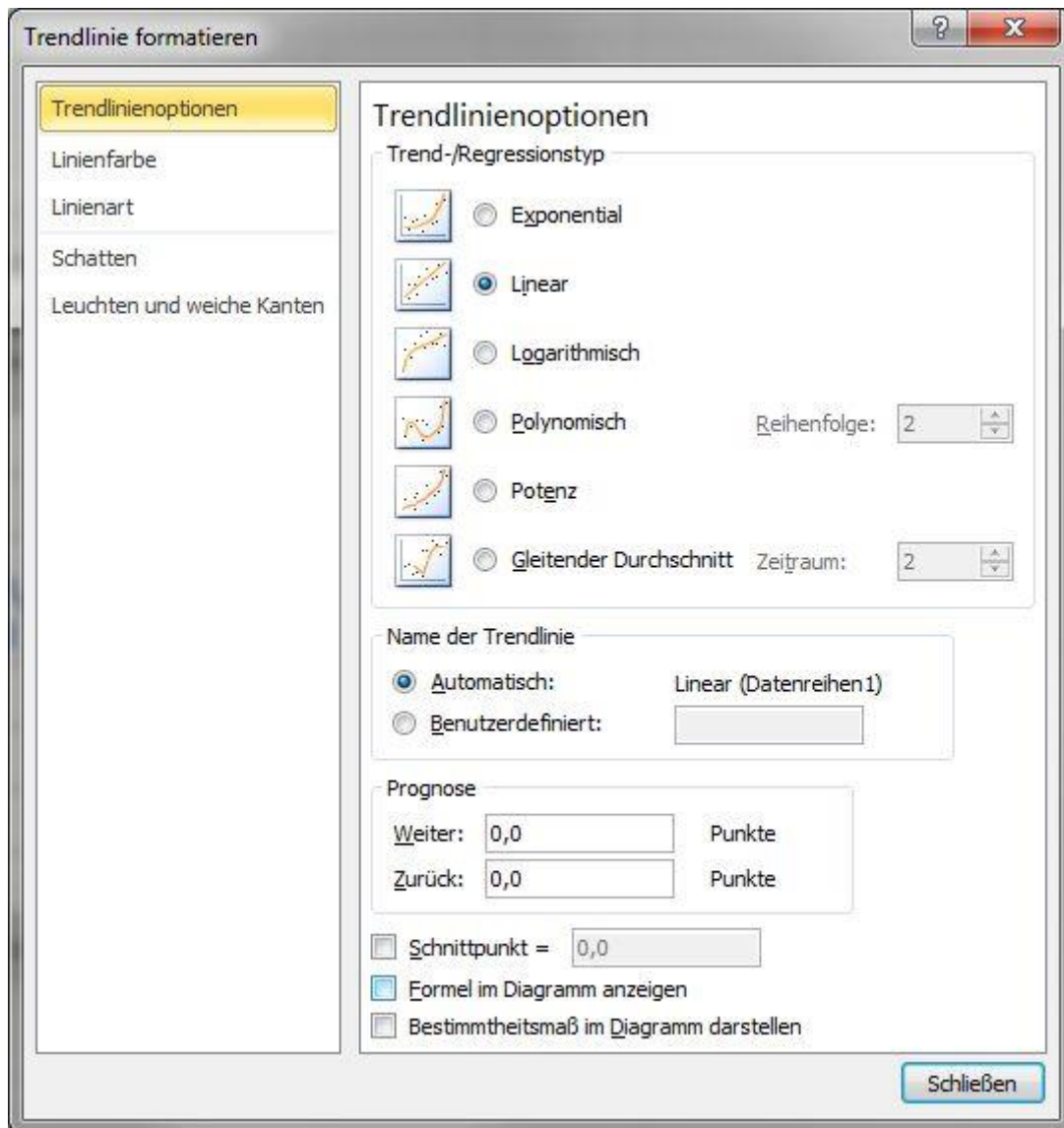


So lassen sich aus dieser Punktwolke mit wenigen einfachen Berechnungen Anpassungsparameter für die Bewertung ableiten:

	F	G	H	I	J	K	L
60							
61		Anpassung des Mittelwerts an das Bewertungsobjekt				Formeln	
62		Mittelwert	111 m ²	10,31 €/m ²		-0,0593*H62+16,896	
63		Objekt	104 m ²	10,73 €/m ²		-0,0593*H63+16,896	
64		Differenz		0,42 €/m ²		I63-I62	
65		prozentual		4,1%		I64/I62	
66		Faktor		1,041		I63/I62	
67							

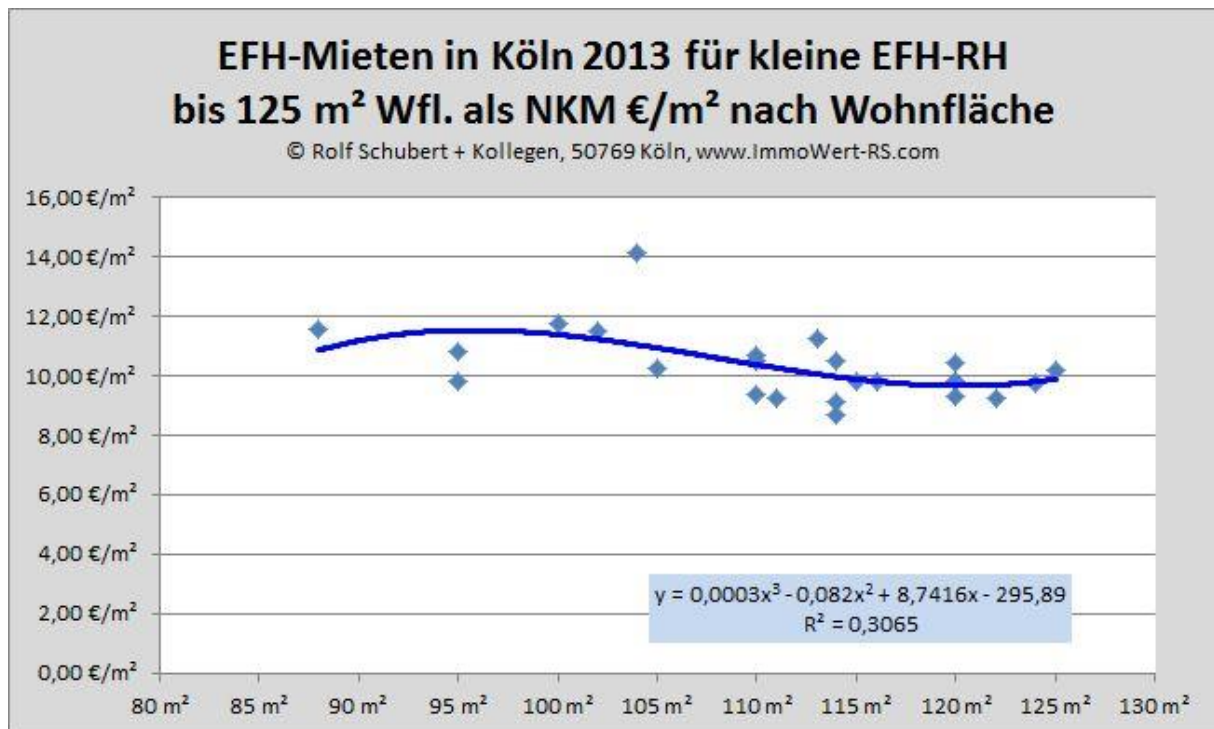
Man braucht nun nur noch diese Anpassungsparameter oder Faktoren auf die Verhältnisse des Bewertungs-Objekts anwenden.

In dem Dialog <Trendlinie formatieren> stecken aber noch weitere gut nutzbare Optionen:



Statt der linearen Trendlinienoption kann man auch polynomische Trendlinien erzeugen, die mit 2 bis 6 Graden abrufbar sind. Man kann auch einen gleitenden Durchschnitt erzeugen oder exponentielle, logarithmische oder Potenz-Funktionen erzeugen. Nachfolgend noch eine polynomische Formel 3. Grades:

Die polynomischen Trendkurven zeigen durchweg bessere Bestimmtheitsmaße und erlauben differenziertere Beurteilungen, sind aber nicht immer zielführend. Es kommt auf die beabsichtigte Aussage an.



Es ist auch möglich mit Datumsangaben zu rechnen. Dabei muss man beachten, dass Excel das Datum in einen Festwert umrechnet, wobei der 01.01.1900 den Wert 0 hat. Seit diesem Datum werden sämtliche Tage fortlaufend gezählt.

Wichtig ist bei der Festlegung des Intervalls, zu beachten, dass 1 Jahr 365,25 Tage, ein halbes Jahr 182,625 und ein Vierteljahr 91,3125 Tage hat. Man sollte das Datum immer formatieren am besten als „Mrz.01“. Denn da die einzelnen Monate unterschiedlich lang sind, stößt Excel sonst an Grenzen.

