

Regression mit Gretl

Eine erste Einführung

Installation:

Gretl für das entsprechende Betriebssystem herunterladen und die Setup-Datei ausführen.

Link: <http://gretl.sourceforge.net/>

Je nach installiertem Betriebssystem "gretl for Windows" oder "gretl on Mac" auswählen.



Download
gretl for Windows
gretl on Mac OS X

Nach dem Programmaufruf sehen Sie dieses Fenster:

Daten importieren

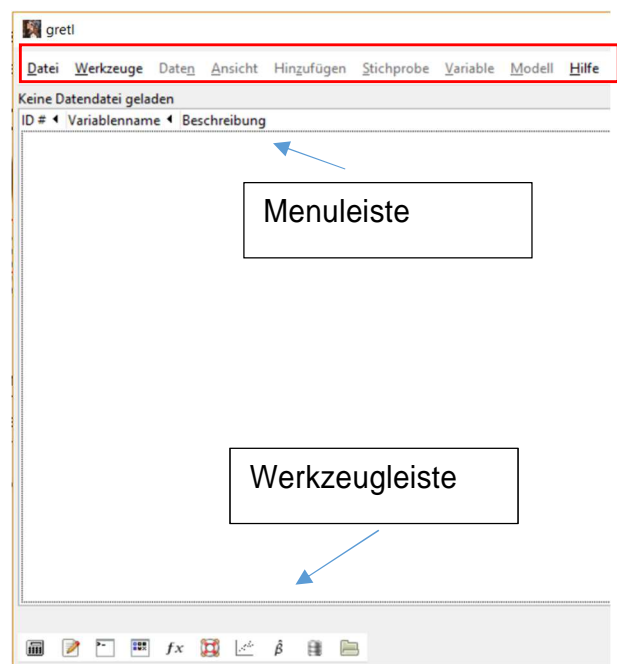
Datei / Öffne Daten / Benutzerdatei

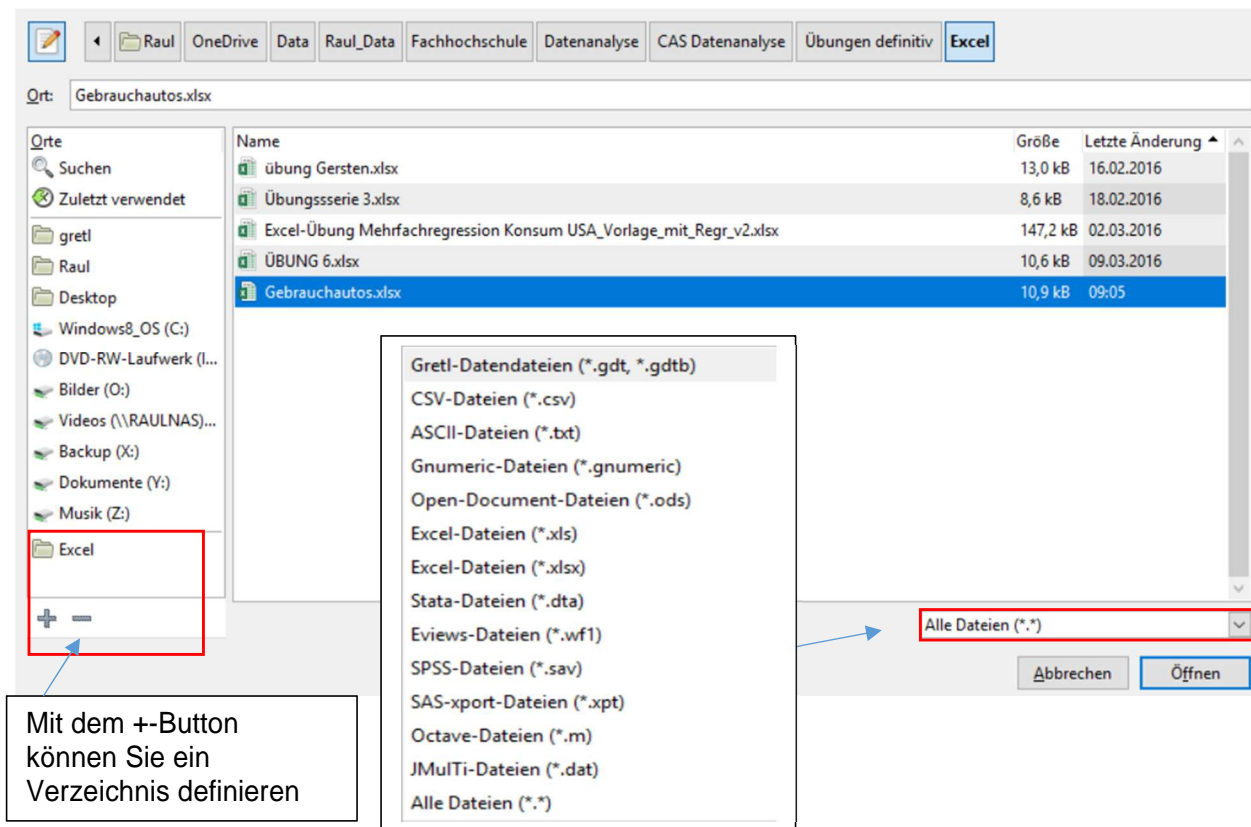
Eine Auswahl von Datenformaten können importiert werden.

Als Standard ist das Gretl-Format .gdt eingestellt.

Am besten wird „Alle Dateien“ ausgewählt, sodass alle Dateien ersichtlich sind.

Im Verzeichnis die gesuchte Datei auswählen.





Konventionen bei der Spreadsheet-Eingabe

Gretl interpretiert die erste Zeile als **Datenbeschriftung**, welche aus bis zu 8 Zeichen bestehen kann.

Gebrauchautos.xlsx - Excel

DATEI START EINFÜGEN SEITENLAY FORMELN DATEN ÜBERPRÜF

Einfügen Schriftart Ausrichtung Zahl Bedingte Formatierung Als Tabelle formatieren Zellenformatvorlagen

Zwischenablage Formatvorlagen

H23

	A	B	C	D	E
1	obs	Preis	Alter	KM	
2	1	16990	2.37	46'200	
3	2	14000	3.38	92'600	
4	3	18000	2.28	35'000	
5	4	14500	3.38	73'000	
6	5	14100	3.46	84'400	
7	6	13480	3.55	137'500	
8	7	18000	2.53	23'000	
9	8	12350	4.23	73'000	
10	9	23900	0.17	1'800	
11	10	15900	2.87	40'500	

Die Datei „Gebrauchautos“ hat vier Spalten. Beim Importieren erscheint ein Pop-up-Fenster.

gretl: Tabellenkalk.-Import

Starte Import bei:

Spalte: 1 Zeile: 1

(A)

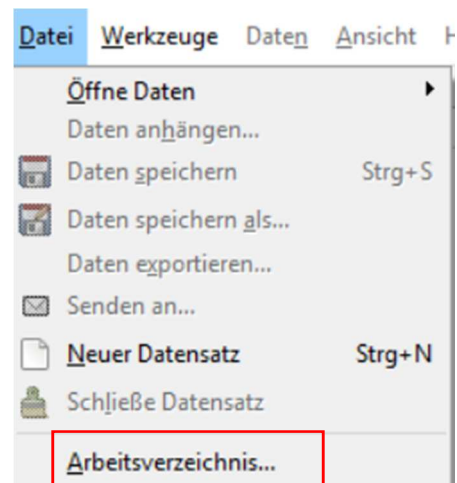
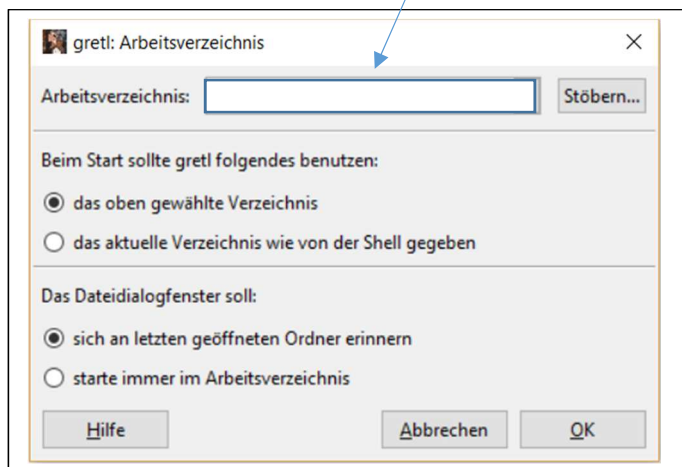
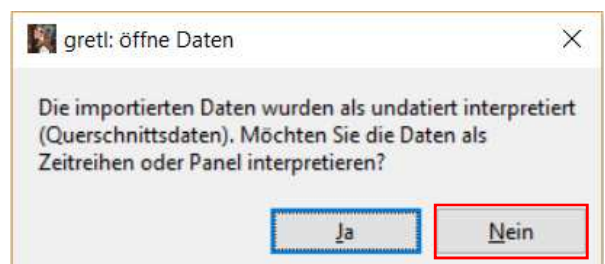
Abbrehen OK

Die importierten Daten sind **nicht** als Zeitreihen zu interpretieren

Man kann gretl beim Import immer zwingen, die zu importierende Tabelle als Zeitreihe zu interpretieren.

Erstellen Sie zuerst ein Verzeichnis, wo Sie alle Excel-Dateien für dieses Modul speichern werden.

Sie können jetzt dieses Arbeitsverzeichnis in gretl festlegen.



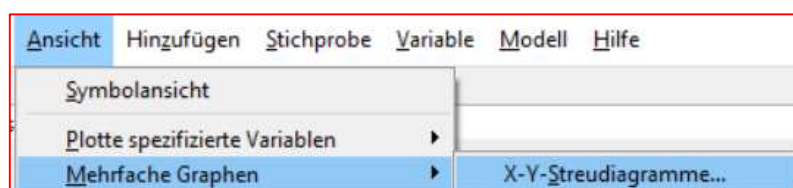
Einfache Regression

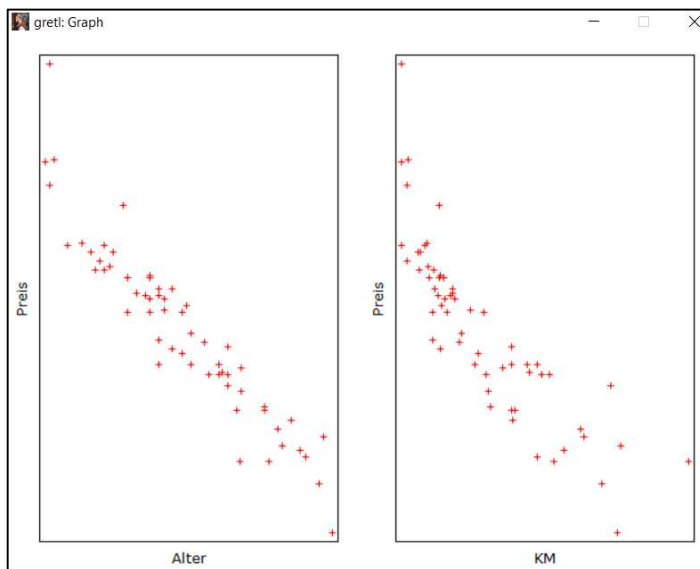
Beim Einlesen der Daten importiert gretl die Variablennamen. Kurz danach sieht man die Variablennamen im Hauptfenster.

Gretl generiert automatisch eine Konstante, da eine Regressionsgleichung ein konstantes Glied enthält.

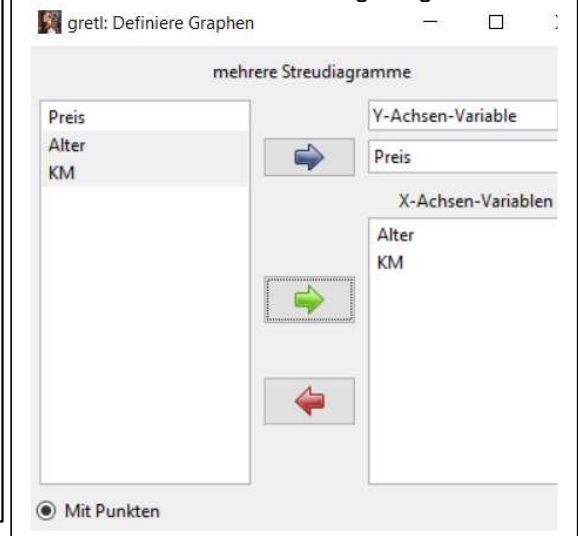


In einem ersten Schritt sollte man sich die Daten ansehen, um eine Idee zu bekommen, wie der Zusammenhang zwischen den Variablen aussieht.





Die unabhängige Variable (Preis) wird immer auf der Y-Achse abgetragen.

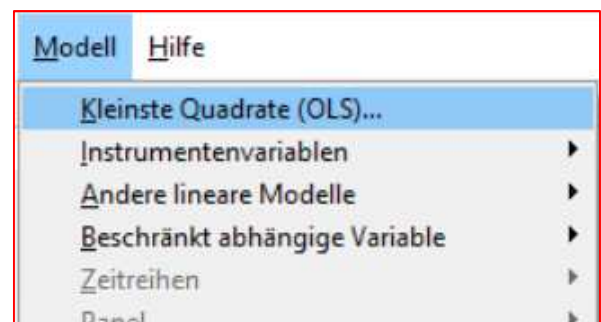


Im nächsten Schritt ist ein Regressionsmodell auszuwählen.

Wenn man auf Modell klickt, fordert gretl einen auf, das Modell zu spezifizieren.

Kleinste Quadrate =

OLS: ordinary least squares

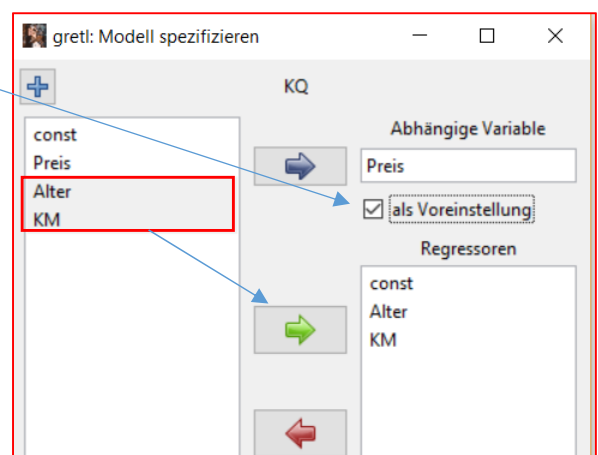


Wie haben eine multiple Regression mit zwei Regressoren: KM und Alter

Mit den Pfeiltasten können Variablen hinzugefügt oder von dort entfernt werden, wenn man einen anderen Zusammenhang testen will.

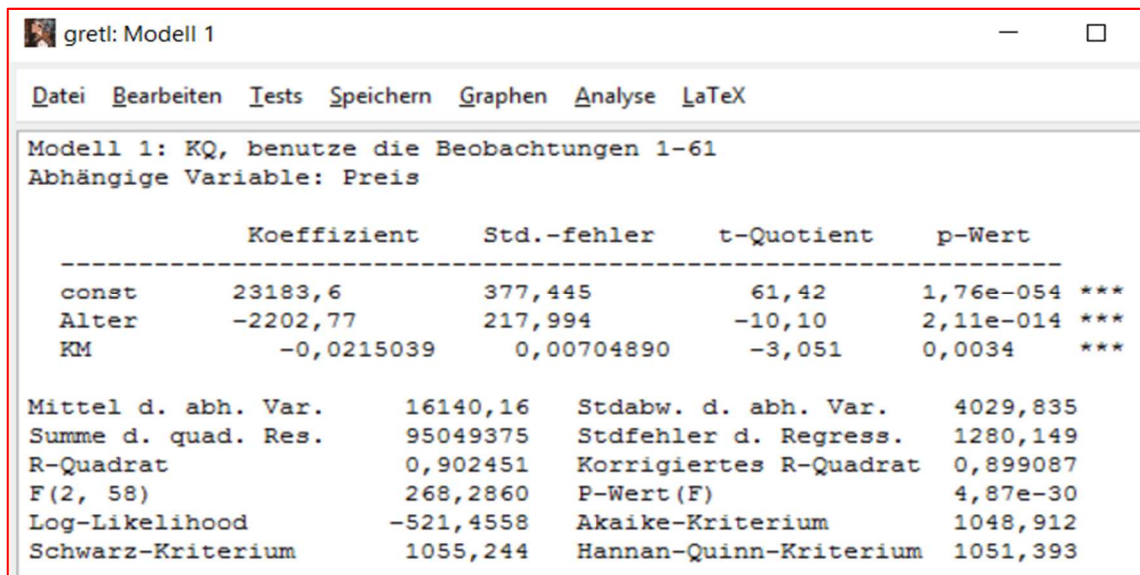
Zuerst wird die abhängige Variable ausgewählt und mit dem grünen Pfeil bestätigt. Ein Haken sollte für die Voreinstellung vorhanden sein, sodass die Variable *Preis* immer als abhängige Variable erscheint

Mehrere unabhängige Variablen können mit der Shift-Taste ausgewählt werden und anschliessend mit dem grünen Pfeil bestätigt werden.



Ein Klick auf OK und es erscheint die Auswertungstabelle in einem neuen Fenster, dem **Modellfenster**

1



	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	23183,6	377,445	61,42	1,76e-054 ***
Alter	-2202,77	217,994	-10,10	2,11e-014 ***
KM	-0,0215039	0,00704890	-3,051	0,0034 ***

Mittel d. abh. Var.	16140,16	Stdabw. d. abh. Var.	4029,835
Summe d. quad. Res.	95049375	Stdfehler d. Regress.	1280,149
R-Quadrat	0,902451	Korrigiertes R-Quadrat	0,899087
F(2, 58)	268,2860	P-Wert(F)	4,87e-30
Log-Likelihood	-521,4558	Akaike-Kriterium	1048,912
Schwarz-Kriterium	1055,244	Hannan-Quinn-Kriterium	1051,393

Hier kommt automatisch eine Reihe von Statistiken, welche wir in der Klasse besprechen werden.