Übungsserie 2: Zeitreihenanalyse

Aufgabe 1: Saisonkomponente

1. Was versteht man unter der Saisonkomponente im Komponentenmodell? Erläutern Sie das Konzept einer saisonbereinigten Zeitreihe. Unterscheiden Sie zwischen dem additiven und multiplikativen Modell.

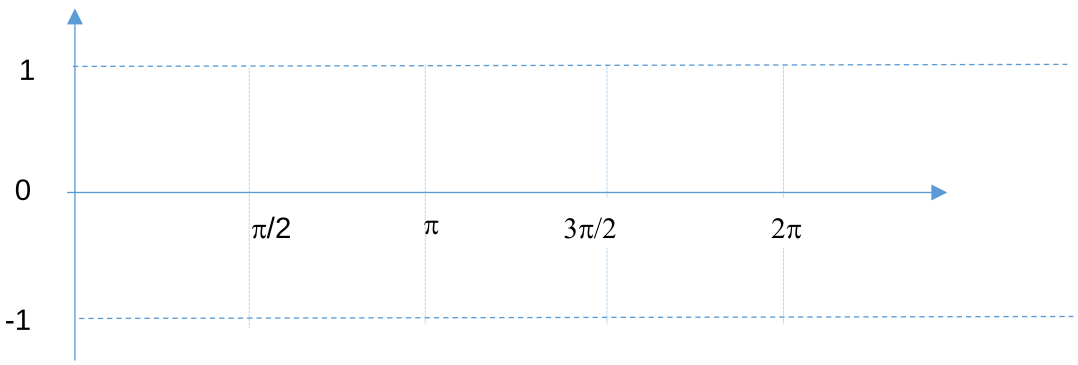
Aufgabe 2: Trigonometrisches Modell für Saisonkomponente

Als Modell für die Saisonschwankungen werden oft trigonometrische Funktionen verwendet.

Die Sinus- (sin) und Kosinusfunktionen (cos) stellen das Grundmodell einer zyklischen Funktion dar. Indem das Argument x mit einem Faktor λ multipliziert oder um einen additiven Term c ergänzt wird, lässt sich eine Vielzahl unterschiedlicher zyklischer Funktionen generieren. Die Multiplikation mit einem konstanten Faktor A erweitert die Palette zusätzlich.

Allgemeine Sinusfunktion: $f(x) = A \sin(\lambda x + c)$

1. Skizzieren Sie die Funktionen sin(x) und cos(x). Nehmen Sie nur 4 Punkte 0, 0.5π, π, 2π



1. Erklären Sie die Auswirkung von A, λ und c auf die generierte Reihe

2. Wie sind die Periodendauer und Frequenz bei Quartalszahlen zu definieren?

Benutzen Sie die Datei USAutos.gdt für Ihre Schätzungen

3. Definieren Sie die neuen Variablen

   Da monatlichen Daten vorhanden sind entspricht P = 12 mit π = 3.1416

   cos1t = cos(time*3.1416/6) →Kosinus-Funktion

   sin1t = sin(time*3.1416/6) →Sinus-Funktion

   cos2t = cos(time*3.1416/3)

   sin2t = sin(time*3.1416/3)

   gretl Hauptfenster: Hinzufügen/ Zeittrend

   Hinzufügen/ Definiere neue Variable

4. Schätzen Sie folgende Modelle

Modell 1: $y_t = \beta_1 + \beta_2 t + \beta_3 t^2 + \beta_4 \cos1t + \beta_5 \sin1t + u$

Modell 2: $y_t = \beta_1 + \beta_2 t + \beta_3 t^2 + \beta_4 \cos1t + \beta_5 \sin1t + \beta_6 \cos2t + \beta_7 \sin2t + u$

Modell 1

Abhängige Variable: y

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	96,5487	0,593874	162,6	5,51e-127	***
time	0,0622203	0,0249030	2,499	0,0140	**
time2	0,00280448	0,000219292	12,79	4,54e-023	***
cos1t	−0,813794	0,274163	−2,968	0,0037	***
sin1t	−1,59936	0,275784	−5,799	7,22e-08	***

Mittel d. abh. Var.	111,2171	Stdabw. d. abh. Var.	12,25795
Summe d. quad. Res.	426,6441	Stdfehler d. Regress.	2,025425
R-Quadrat	0,973709	Korrigiertes R-Quadrat	0,972698
F(4, 104)	962,9342	P-Wert(F)	3,49e-81
Log-Likelihood	−229,0351	Akaike-Kriterium	468,0702

Modell 2

Abhängige Variable: y

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	96,5004	0,540015	178,7	3,54e-129	***
time	0,0628231	0,0226416	2,775	0,0066	***
time2	0,00280864	0,000199375	14,09	1,14e-025	***
cos1t	−0,813308	0,249172	−3,264	0,0015	***
sin1t	−1,59451	0,250628	−6,362	5,72e-09	***
cos2t	−1,13639	0,250091	−4,544	1,52e-05	***
sin2t	0,460389	0,248941	1,849	0,0673	*

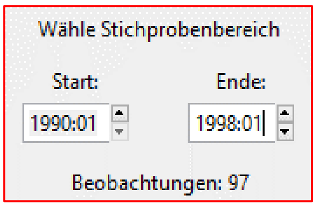
Mittel d. abh. Var.	111,2171	Stdabw. d. abh. Var.	12,25795
Summe d. quad. Res.	345,5576	Stdfehler d. Regress.	1,840603
R-Quadrat	0,978706	Korrigiertes R-Quadrat	0,977453
F(6, 102)	781,3406	P-Wert(F)	7,28e-83
Log-Likelihood	−217,5470	Akaike-Kriterium	449,0941

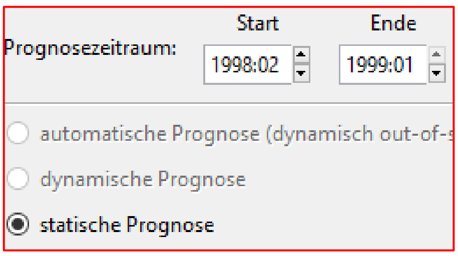
5. Erstellen Sie die Grafik der originären Zeitreihe mit der angepassten Daten

6. Welches Modell würden Sie vorziehen?

7. Erstellen Sie mittels Modell 2 Prognosen für den Prognosezeitraum 1998:2 – 1999:01

Hinweis: Zuerst muss die Stichprobe reduziert werden und Modell 2 neu geschätzt werden. Anschliessend den Prognosezeitraum definieren: 1998:02 -1999:01

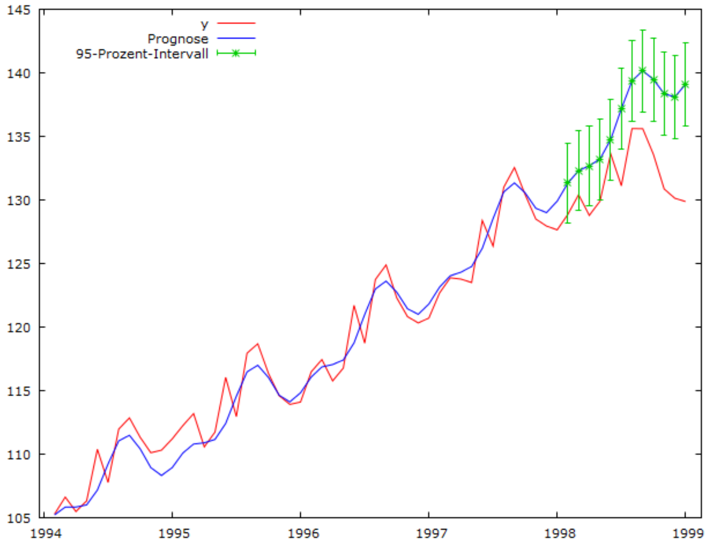




Analyse LaTeX
Zeige tatsächl
Prognosen...

Datei Bearbeiten Tests Speichern Graphen Analyse LaTeX

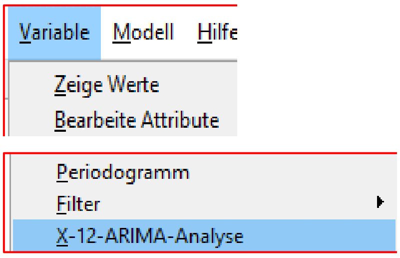
```
Modell 5: KQ, benutze die Beobachtungen 1990:01-1998:01 (T = 97)
Abhängige Variable: y

             Koeffizient   Std.-fehler   t-Quotient    p-Wert
  ------------------------------------------------------------
  const       97,6569      0,461383       211,7       3,30e-123 ***
  time        -0,0232989   0,0217074       -1,073     0,2860
  time2        0,00383172  0,000214560     17,86      2,47e-031 ***
  cos1t       -0,666921    0,212304        -3,141     0,0023    ***
  sin1t       -1,65236     0,213729        -7,731     1,46e-011 ***
  cos2t       -1,06010     0,213148        -4,974     3,13e-06  ***
  sin2t        0,451619    0,212046         2,130     0,0359    **
```



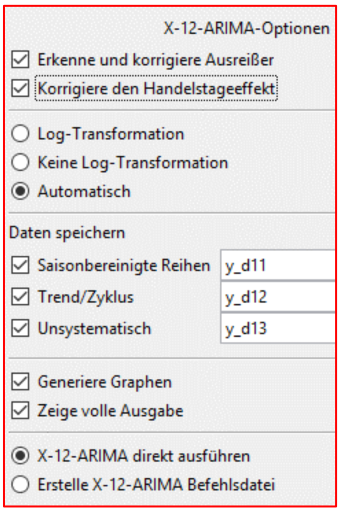
Aufgabe 3: Census-Verfahren X-12-Arima

Das Census-Verfahren wird unter anderem vom U.S. Bureau of the Census, der OECD, von der Europäischen Zentralbank und vielen nationalen statistischen Behörden verwendet.
Neben eigenständigen Modulen zur Erkennung und Berücksichtigung extremer Werte (Ausreisser) und zur Bereinigung um Kalendereinflüsse besteht das Verfahren im Kern aus einer iterativen Prozedur zur Bestimmung der Trendkomponente und der daraus resultierenden Saisonfaktoren.
Die Bestimmung der Trendkomponente basiert im Wesentlichen auf gleitenden Durchschnitten.
Aus den Originalwerten y_t und den gewichteten gleitenden Durchschnitten $y_t(\varnothing 13)$ werden für jeden Monat (jedes Quartal) Saisonfaktoren s_t berechnet, für die wiederum gleitende Durchschnitte $s_t(\varnothing)$ gebildet werden. Nachdem die Ursprungsreihe um den Einfluss dieser Saisonfaktoren bereinigt wurde, wird das Verfahren erneut auf die verbleibende Restgrösse angewandt.
Gehen Sie auf die Webseite http://gretl.sourceforge.net/win32/ und installieren Sie das Paket X-12-ARIMA

Für Mac: http://gretl.sourceforge.net/osx.html
Für Linux: http://gretl.sourceforge.net/#dl



gretl Datei: USAutos.gdt



1. Benutzen Sie die gretl Funktion X-12-ARIMA für die Bereinigung der Zeitreihe der registrierten Autos.

   gretl Hauptfenster: Variable / X-12-ARIMA-Analyse

Aufgabe 4: Hodrick-Prescott Filter (HP-Filter)

2. Erklären Sie kurz die Idee der Methode

   Die Idee der Methode besteht darin, die Abwägung zwischen einer möglichst guten Anpassung der vorhandenen Daten einerseits und einer möglichst glatten Trendkomponente andererseits explizit vorzugeben.

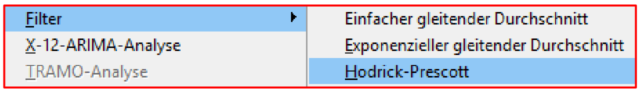
3. Erklären Sie kurz beide Komponente der Zielfunktion für den HP-Filter:

$$\text{Min} \sum_{t=1}^{T}(y_t-\tau_t)^2+\lambda\sum_{t=1}^{T}\left[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})\right]^2$$

4. Erklären Sie kurz die Auswirkung auf die Glättung für kleine und grosse Gewichtungsparameter λ.

   Hinweis: gretl erkennt die Frequenz der Daten und setzt den entsprechenden Gewichtungsparameter automatisch ein.

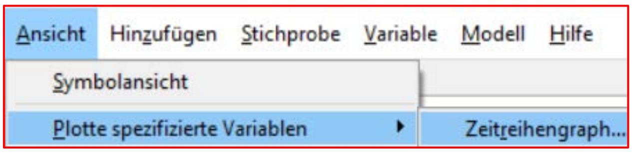
5. Glätten Sie die Zeitreihe der registrierten Autos anhand des HP-Filters. Anschliessend wiederholen Sie die Glättung mit λ = 100. Was stellen Sie fest?

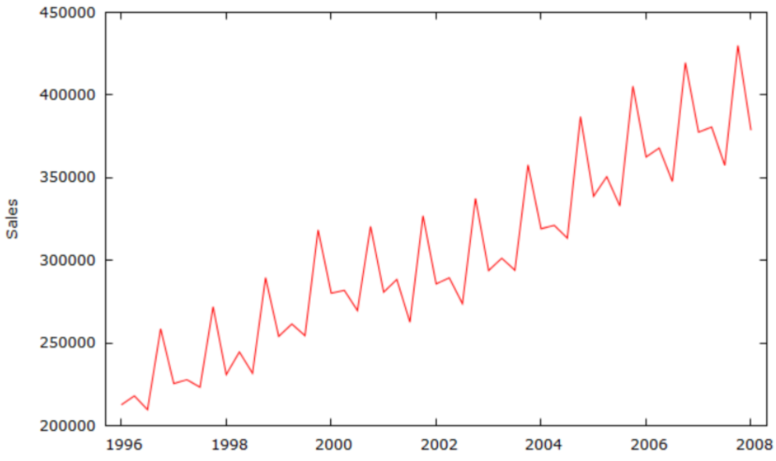


Aufgabe 5: Saisondummies

Benutzen Sie die gretl-Datei sales.gdt. Die Daten stellen die US retail & food services sales dar, entsprechen den Einzelhandelsumsätzen für die Periode 1996:Q1 bis 2008:Q1.

1. Erstellen Sie das Zeitreihendiagram. Was stellen Sie fest?





2. Schätzen Sie das Trendmodell $y_t = \beta_1 + \beta_2 \mathbf{t} + u_t$

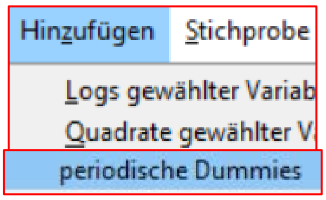
```
Abhängige Variable: Sales

             Koeffizient   Std.-fehler   t-Quotient    p-Wert
  ------------------------------------------------------------
  const        211461        6876,48        30,75      8,78e-033 ***
  time           3678,10      239,408       15,36      5,99e-020 ***

Mittel d. abh. Var.      303413,3   Stdabw. d. abh. Var.    57550,40
Summe d. quad. Res.      2,64e+10   Stdfehler d. Regress.   23700,20
R-Quadrat                0,833940   Korrigiertes R-Quadrat  0,830407
```

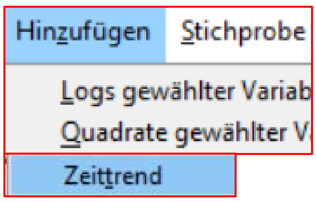
3. Erklären Sie kurz was Saisondummies sind.

4. Schätzen Sie das Modell 2 mit den entsprechenden Saisondummies.

   Modell 2: $y_t = \beta_1 + \beta_2 t + \beta_3 D_2 + \beta_4 D_3 + \beta_5 D_4 + u$

   Fügen Sie zuerst die Dummyvariablen für die Saisons sowie eine Trendvariable mittels gretl Menu:
   gretl Hauptfenster: Hinzufügen / periodische Dummies *Hinzufügen / Zeittrend*





Abhängige Variable: Sales

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	205459	3416,85	60,13	7,62e-041	***
time	3608,74	100,337	35,97	4,42e-032	***
dq2	1754,41	3644,95	0,4813	0,6329	
dq3	-14503,9	3643,57	-3,981	0,0003	***
dq4	43639,5	3644,95	11,97	8,41e-015	***

Mittel d. abh. Var.	296011,2	Stdabw. d. abh. Var.	53501,94
Summe d. quad. Res.	3,05e+09	Stdfehler d. Regress.	8728,700
R-Quadrat	0,975803	Korrigiertes R-Quadrat	0,973383
F(4, 40)	403,2687	P-Wert(F)	9,71e-32
Log-Likelihood	-469,5488	Akaike-Kriterium	949,0977

Abhängige Variable: Sales

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	295406	15007,7	19,68	9,65e-024	***
dq2	-968,327	21661,8	-0,04470	0,9645	
dq3	-14491,7	21661,8	-0,6690	0,5069	
dq4	48156,2	21661,8	2,223	0,0313	**

Mittel d. abh. Var.	303413,3	Stdabw. d. abh. Var.	57550,40
Summe d. quad. Res.	1,32e+11	Stdfehler d. Regress.	54111,09
R-Quadrat	0,171205	Korrigiertes R-Quadrat	0,115952

5. Welche implizite Annahme legt dieser Spezifikation mit Dummyvariablen zugrunde? Welches Quartal ist das Referenzquartal?

6. Welches Modell weist die beste Anpassungsgüte auf?

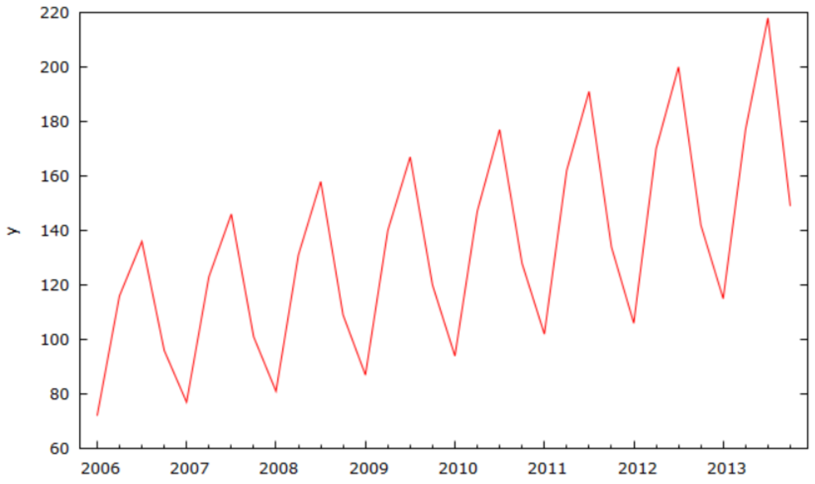
7. Berechnen Sie die normierten Saisonfaktoren anhand der Regressionsergebnisse. (H57:H60)

8. Interpretieren Sie den Saisonfaktor S_3

Aufgabe 6: Holt-Winters Modell

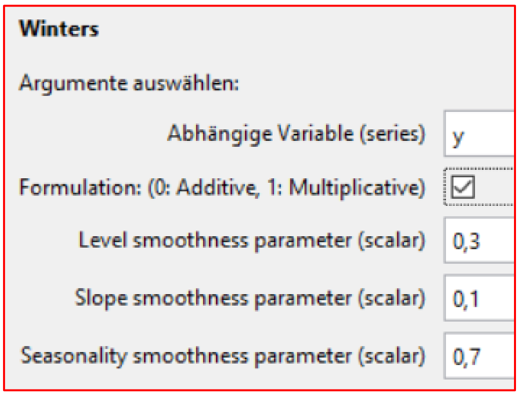
Benutzen Sie die Datei Sportgetränke.gdt

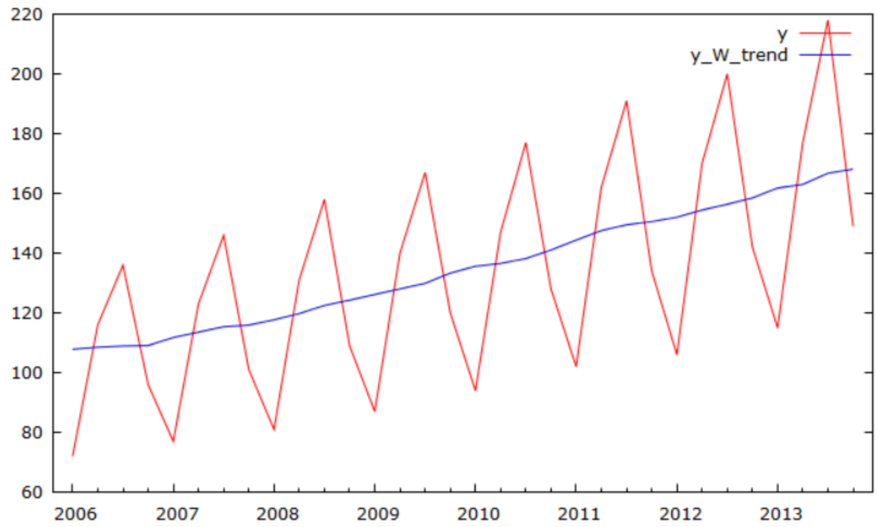
1. Erstellen Sie das Zeitreihendiagram. Was stellen Sie fest?

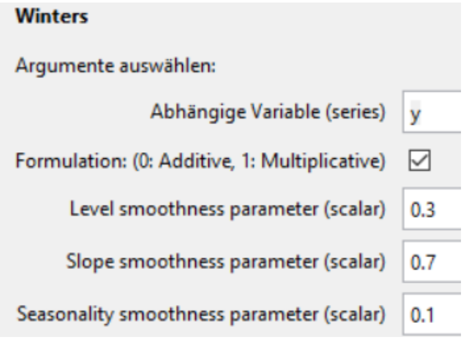


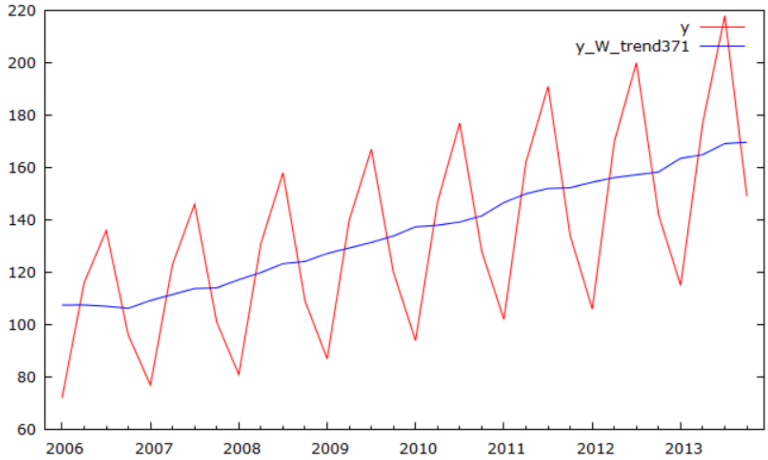
2. Wann wird das Winters Verfahren angewendet?

3. Glätten Sie mittels Winters-Methode die Zeitreihe y mit den Parametern $\alpha = 0.3$, $\gamma = 0.1$ und $\delta = 0.7$?




y
y_W_trend
220
200
180
160
140
120
100
80
60
2006
2007
2008
2009
2010
2011
2012
2013


Winters
Argumente auswählen:
Abhängige Variable (series)
y
Formulation: (0: Additive, 1: Multiplicative)
Level smoothness parameter (scalar)
0.3
Slope smoothness parameter (scalar)
0.7
Seasonality smoothness parameter (scalar)
0.1


y
y_W_trend371
220
200
180
160
140
120
100
80
60
2006
2007
2008
2009
2010
2011
2012
2013

Aufgabe 7: Multiplikatives Modell

Sie erhalten folgende Tabelle mit den Quartalsumsätzen von Traktoren eines Unternehmens. Alle Zahlen sind in Millionen Euros ausgedrückt. Ein multiplikatives Modell für die Saisonbereinigung wurde angewandt.

1	2	3	4	5	6	7	8	9
		t	y_t	GD	S_{ij}	S*	y*	Trend
2005	Q1	1	362			0.960519635	376.879334	342.8
	Q2	2	385			1.022149385	376.657273	360.6
	Q3	3	432	382.5	1.12941176	1.140020527	378.940545	378.4
	Q4	4	341	388	0.87886598	0.877310453	388.687948	396.2
2006	Q1	5	382	399.25	0.95679399	0.960519635	397.701396	414
	Q2	6	409	413.25	0.98971567	1.022149385	400.137207	431.8
	Q3	7	498	430.375	1.15713041	1.140020527	436.834239	449.6
	Q4	8	387	A	B	C	D	E
2007	Q1	9	473	478.25	0.98902248	0.960519635	492.441782	485.2
	Q2	10	513	499.625	1.02677008	1.022149385	501.883587	503
	Q3	11	582	519.375	1.12057762	1.140020527	510.517123	520.8
	Q4	12	474	536.875	0.88288708	0.877310453	540.287647	538.6
2008	Q1	13	544	557.875	0.97512884	0.960519635	566.360104	556.4
	Q2	14	582	580.625	1.00236814	1.022149385	569.388397	574.2
	Q3	15	681	601.5	1.13216958	1.140020527	597.357665	592
	Q4	16	557	627.625	0.88747262	0.877310453	634.894977	609.8
2009	Q1	17	628	654.75	0.95914471	0.960519635	653.812767	627.6
	Q2	18	707	670.625	1.05424045	1.022149385	691.67972	645.4
	Q3	19	773	674.875	1.1453973	1.140020527	678.057966	663.2
	Q4	20	592	677	0.87444609	0.877310453	674.789634	681
2010	Q1	21	627	689.375	0.90951949	0.960519635	652.771664	698.8
	Q2	22	725	708.125	1.02383054	1.022149385	709.28967	716.6
	Q3	23	854			1.140020527	749.109318	734.4
	Q4	24	661			0.877310453	753.439102	752.2

Spalte 5: gleitende Durchschnitte (GD)

Spalte 6: Saisonfaktoren (S_{ij})

Spalte 7: Normierte Saisonfaktoren S*

Spalte 8: Saisonbereinigte Werte y*

Spalte 9: Trendkomponente

Leider sind die Werte für **2006:Q4** verloren gegangen

Die unnormierten Saisonfaktoren wurden wie folgt berechnet:

Q1	0.95792
Q2	1.01938
Q3	1.13693
Q4	0.87493

Summe = 3.98918

1. Schreiben Sie das multiplikative Komponentenmodell auf und erklären Sie wann es angewendet werden sollte.

2. Bestimmen Sie den Wert A

3. Bestimmen Sie den Wert B

4. Berechnen Sie den Normalisierungsfaktor

5. Berechnen Sie den normierten Saisonfaktor für das 4. Quartal, C.

6. Erklären Sie kurz warum normierte Saisonkomponente für Q1 und Q2 2005 vorhanden sind, obwohl keine Werte für die gleitenden Durchschnitte berechnet wurden.

7. Berechnen Sie den saisonbereinigten Wert D

8. Berechnen Sie den geschätzten Trendwert E.

9. Berechnen Sie den entsprechenden Prognosefehler