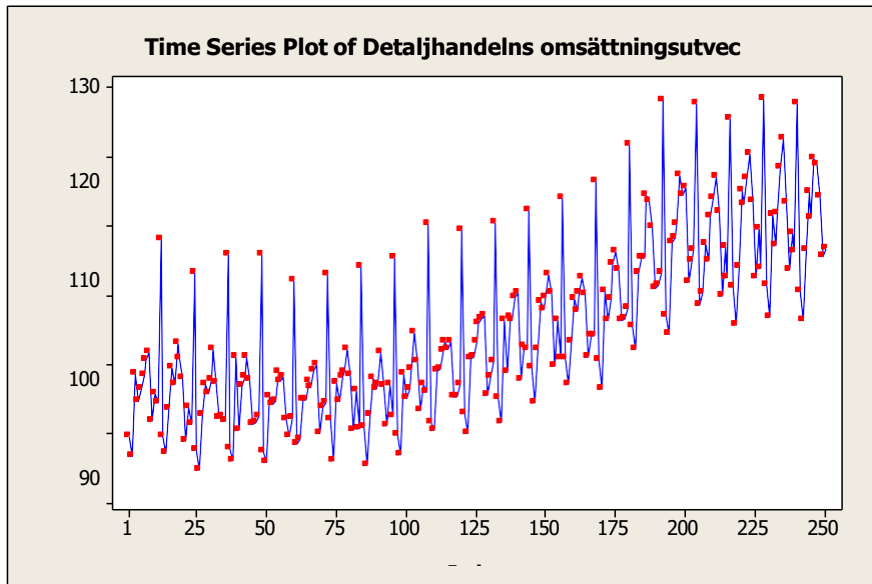


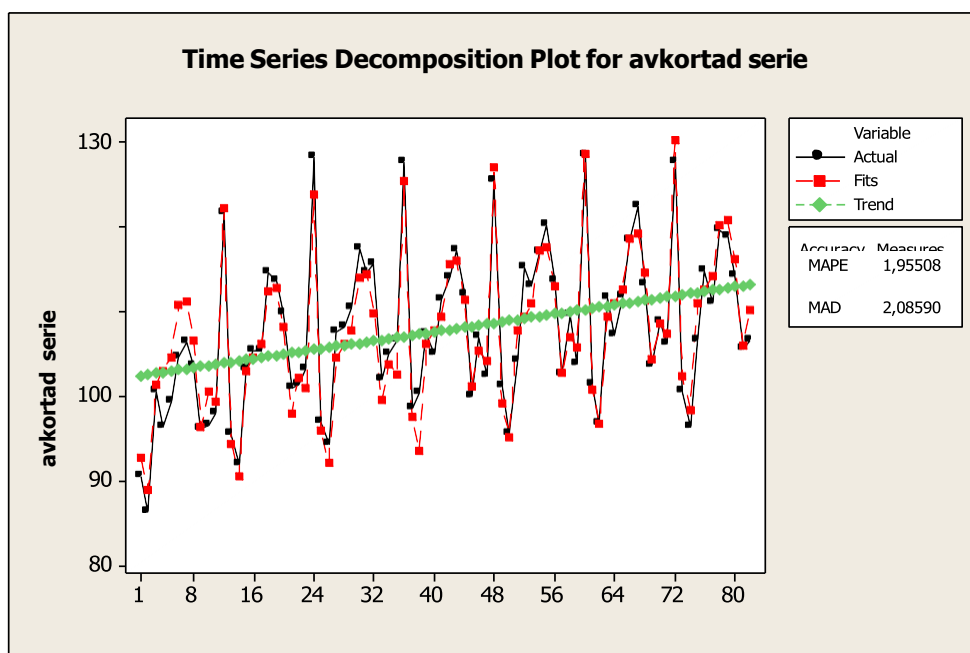
Lektion 2

1

På SCB's hemsida har hämtats data: **Detaljhandelns omsättningsutveckling efter näringsgren, 2005=100. Månad 1991M01-2011M10, Fasta priser, dagligvaror.**



Man kan konstatera att serien har ett tydligt säsongsmönster men inte en linjär trend. Eftersom det är mer intressant med nyare värden så har serien kortats av från och med januari 2005. Nedan är två modeller anpassade för den avkortade serien. Uppgifterna kommer efter alla utskrifter. Modell 1



Time Series Decomposition for avkortad serie

Additive Model

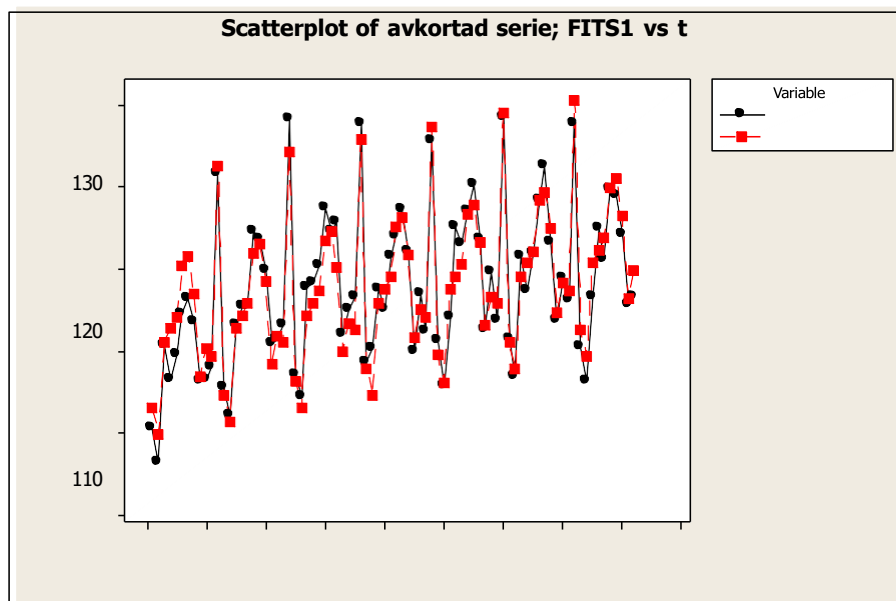
Data avkortad serie
Length 82
NMissing 0
Fitted Trend Equation
 $Y_t = 102,295 + 0,133 \cdot t$
Seasonal Indices

Period	Index
1	-9,6417
2	-13,6833
3	-1,3521
4	0,1875
5	1,6208
6	7,6500
7	8,0479
8	3,3062
9	-7,1458
10	-3,0271
11	-4,3167
12	18,3542

Accuracy Measures

MAPE 1,95508
MAD 2,08590
MSD 6,91477

Modell 2



Regression Analysis: avkortad serie versus t; Jan; ...

The regression equation is

avkortad serie = 89,5 + 0,133 t + 3,39 Jan + 11,2 mars + 12,6 april + 14,0 maj
+ 20,0 juni + 21,0 juli + 16,3 aug + 6,05 sep + 9,35 okt
+ 8,39 nov + 31,5 dec

Predictor	Coef	SE Coef	T	P
Constant	89,474	1,184	75,60	0,000
t	0,13302	0,01330	10,00	0,000
Jan	3,390	1,514	2,24	0,028
mars	11,167	1,514	7,38	0,000
april	12,634	1,514	8,35	0,000
maj	13,972	1,514	9,23	0,000
juni	19,968	1,515	13,18	0,000
juli	20,963	1,515	13,84	0,000
aug	16,345	1,516	10,78	0,000
sep	6,055	1,516	3,99	0,000
okt	9,350	1,517	6,16	0,000
nov	8,389	1,576	5,32	0,000
dec	31,456	1,576	19,96	0,000

S = 2,83164 R-Sq = 91,9% R-Sq(adj) = 90,5%

Analysis of Variance

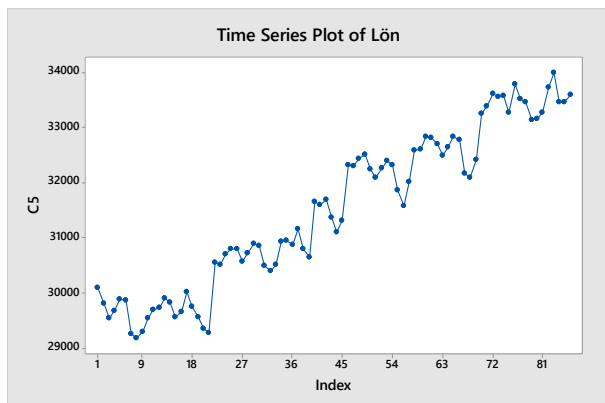
Source	DF	SS	MS	F	P
Regression	12	6287,47	523,96	65,35	0,000
Residual Error	69	553,25	8,02		
Total	81	6840,73			

Durbin-Watson statistic = 1,28747

- I vilken månad är omsättningen lägst enligt modell 1?
- Hur stor är ökningen i omsättning per månad och per år i snitt enligt någon av de två modellerna?
- Tolka säsongskomponenten för december i båda modellerna.
- Beräkna prognos för november och december 2011 samt december 2012. Dvs tre prognoser. Gör detta dels för modell 1 och dels för modell 2.
- Är det rimligt att anta att residualerna är tidsberoende i modell 2?

2

Nu ska genomsnittlig månadslön för polisväsendet analyseras



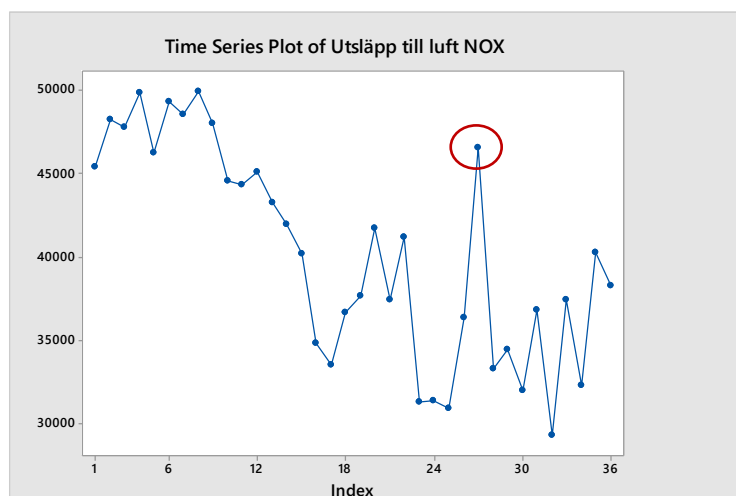
De sista 13 värdena i tidsserien lön är givna nedan

År,månad	Lön
2016M02	33563
2016M03	33270
2016M04	33779
2016M05	33521
2016M06	33463
2016M07	33134
2016M08	33161
2016M09	33268
2016M10	33731
2016M11	33986
2016M12	33461
2017M01	33449
2017M02	33587

- Räkna om Lön till fastbasindex från månad 2016M10 till månad 2017M02 men där basmånad är 2016M08, dvs index för 2016M08 ska vara 100.
- Om man vill jämföra hur den genomsnittliga månadslönen för polisväsendet ligger i förhållande till den allmänna prisutvecklingen i samhället. Vilken indexserie ska man då jämföra med? Och hur gör man en sådan jämförelse? Vad kallas en sådan jämförande indexserie?

3

Variabeln som ska analyseras kallas Utsläpp till luft NOX i ton
Graf över data:



Nedan ses utskrifter från fyra olika regressionsmodeller:

Förklarande variabler:

Tid: en numrering av tidpunkterna 1,2,...,36

Kvartal: Numrering av kvartal 1,2,3,4,1,2,3,...,3,4

Kvartal_i är indikator för kvartal i. *Kvartal_i* är 1 om kvartal i och 0 annars.

Indikator är en indikator för obs nr 27 så indikator är 1 om obs nr 27 och 0 annars. I grafen ovan ses ett extremt värde för obs nr 27 inringad med rött

Modell 1

Regression Analysis: Utsläpp till luft NOX versus tid; kvartal

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	827345680	413672840	23,01	0,000
Error	33	593253204	17977370		
Total	35	1420598884			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	47531	2048	23,21	0,000	
tid	-464,2	68,4	-6,78	0,000	1,01
kvartal	489	636	0,77	0,447	1,01

Modell 2

Regression Analysis: Utsläpp till luft NOX versus tid; kvartal_1; kvartal_2; kvartal_3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	853047136	213261784	11,65	0,000
Error	31	567551748	18308121		
Total	35	1420598884			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	48571	1985	24,47	0,000	
tid	-464,2	69,0	-6,72	0,000	1,01
kvartal_1	-1068	2028	-0,53	0,602	1,52
kvartal_2	55	2022	0,03	0,978	1,51
kvartal_3	1744	2018	0,86	0,394	1,50

Modell 3

Regression Analysis: Utsläpp till luf versus tid; kvartal_1; kvartal_2; kvartal_3; indikator

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	942048820	188409764	11,81	0,000
Error	30	478550064	15951669		
Total	35	1420598884			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	48992	1862	26,32	0,000	
tid	-485,2	65,1	-7,46	0,000	1,03
kvartal_1	-1131	1893	-0,60	0,555	1,52
kvartal_2	13	1887	0,01	0,994	1,51
kvartal_3	601	1945	0,31	0,760	1,60
indikator	10101	4277	2,36	0,025	1,11

Modell 4

Regression Analysis: Utsläpp till luft NOX versus tid; indikator

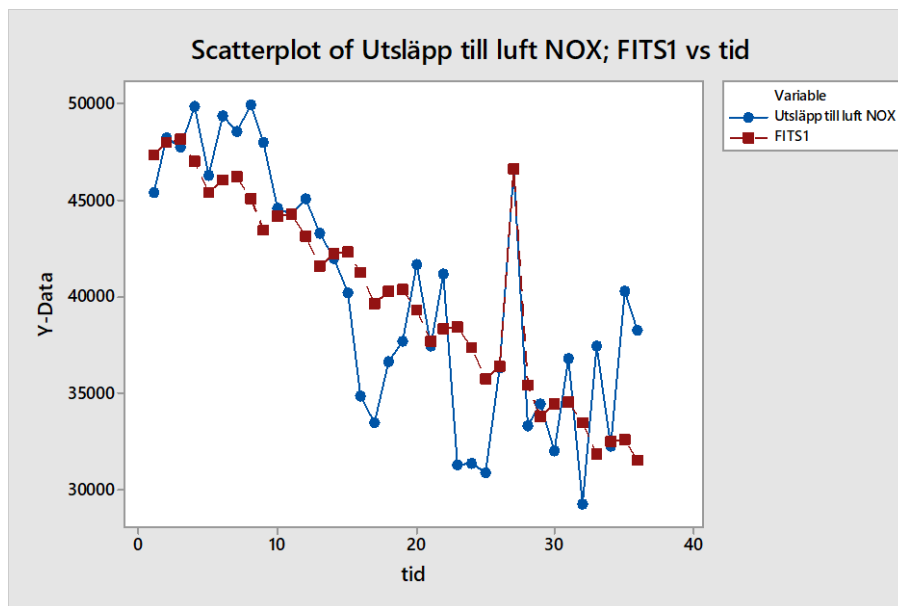
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	928462466	464231233	31,13	0,000
Error	33	492136418	14913225		
Total	35	1420598884			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	48787	1316	37,09	0,000	
tid	-482,2	62,6	-7,71	0,000	1,02
indikator	10826	3955	2,74	0,010	1,02

- Modell 1 har ett allvarligt fel. Förklara vad som är fel och varför det är fel
- Vilken av modellerna 2, 3 eller 4 anser du vara bäst. Motivera noga.
- Tolka regressionskoefficienten för kvartal_1 i modell 3
- Hur många ton NOX minskar utsläpp till luft varje år enligt modell 4?
- Hur många fler ton utsläpp av NOX var det vid tidpunkt 27 jämfört med trenden enligt modell 3? Grafen nedan är kanske till hjälp.
- Beräkna en prognos för tidpunkt 37 med hjälp av modell 2
- Vilket värde skulle VIF ha om alla förklarande variabler var okorrelerade?



Denna graf visar anpassade värden (rött) för modell 3.

4

Ur statistisk årsbok har följande uppgifter hämtats:

År	1998	1999	2000	2001	2002	2003
Mjölprodukter Värde, milj. kr	108.6	113.7	112.4	117.4	119.5	121.6
Grupp kött och fläsk Värde, milj. kr	335.1	344.0	341.8	338.5	371.2	385.4
Rågbröd 400g, medelpris i kr (Ur grupp mjölprod)	12.80	13.30	13.40	13.90	15.10	15.10
Rökt skinka kg, medelpris i kr (Ur grupp kött och fläsk)	121.70	126.70	130.60	142.40	142.10	136.40
KPI	257.0	258.1	260.7	267.1	272.8	278.1

- a) Bilda två enkla prisindexserier för Rågbröd och Rökt skinka med basår 1998.
- b) Bilda på lämpligt sätt ett sammansatt kedjeindex för de båda grupperna Mjölprodukter och Kött och fläsk. Rågbröd och Rökt skinka ska vara gruppernas respektive representantvaror. Använd Laspeyrevikter. Basår 1998.
- c) Hur ser prisutvecklingen ut på mjölprodukter och kött jämfört med den allmänna prisutvecklingen på konsumtion mellan åren 1998 och 2003?