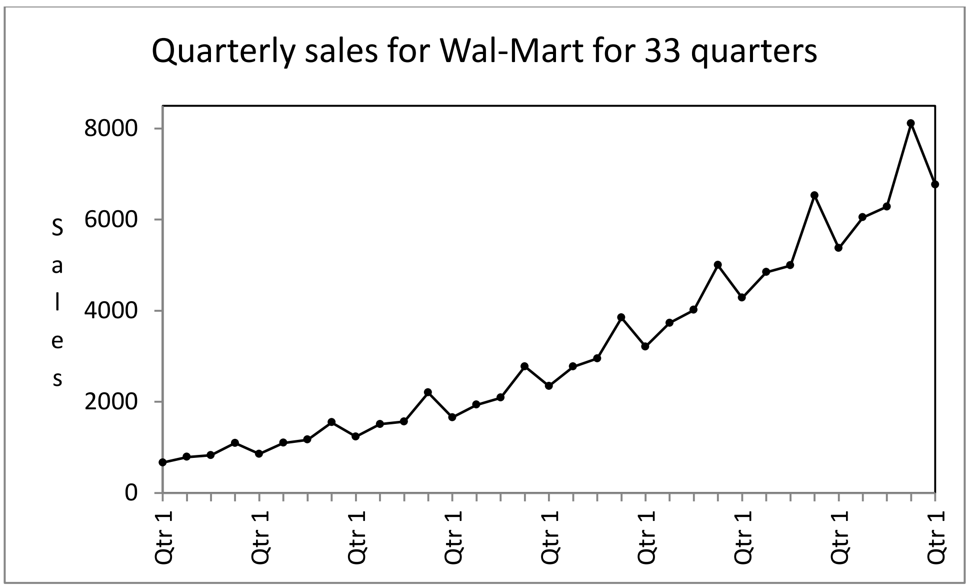
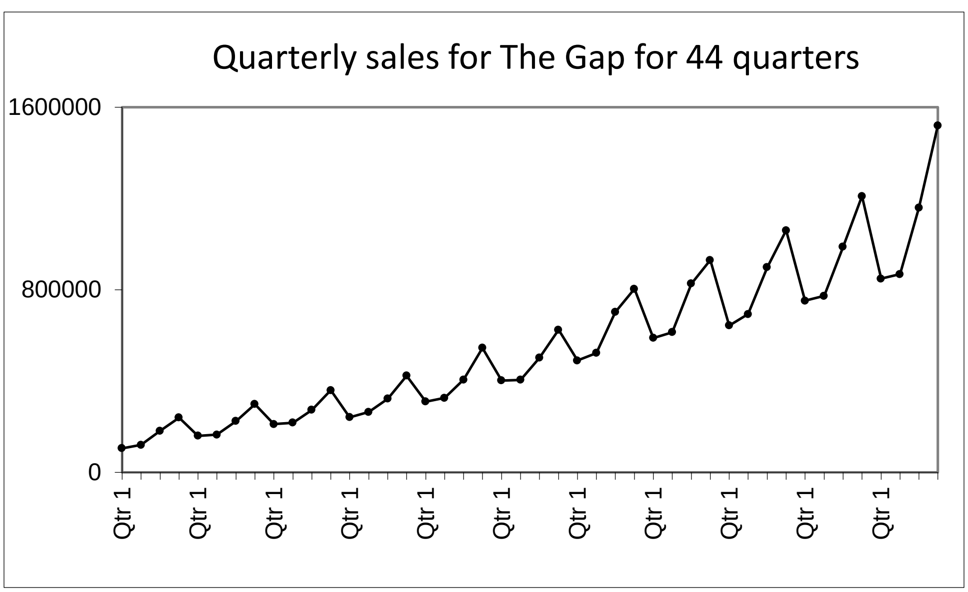
THE UNIVERSITY OF TEXAS AT AUSTIN
McCombs School of Business

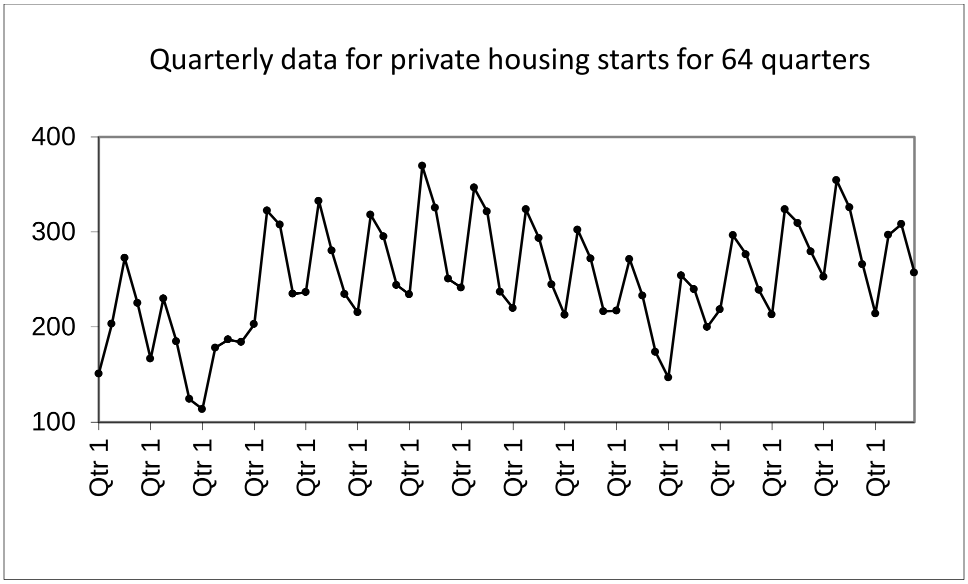
STA 372.5 **Tom Shively**

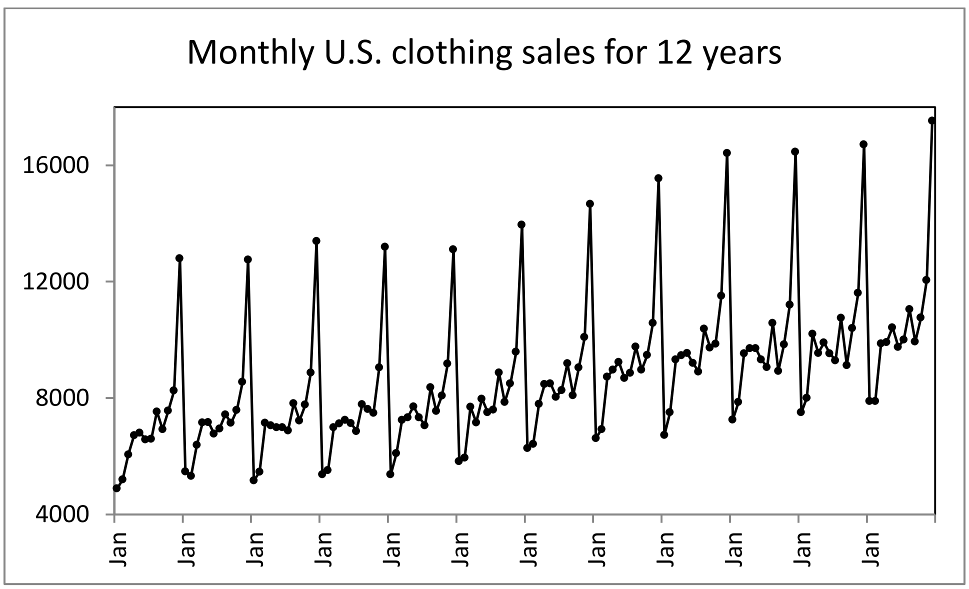
CLASSICAL SEASONAL DECOMPOSITION - MULTIPLICATIVE MODELS

Examples of Seasonality






Quarterly data for private housing starts for 64 quarters
400
300
200
100
Qtr 1


Monthly U.S. clothing sales for 12 years
16000
12000
8000
4000
Jan

A time series is typically decomposed into four components:

(1) Seasonal component (S_t);

(2) Trend component (T_t);

(3) Cyclical component (C_t);

(4) Irregular component (I_t).

The trend and cyclical components are often difficult to separate except in the long run. (See the figures on the following page.)

For this reason, the trend and cyclical components are often combined into one component - the Trend/Cycle component (TC_t).

Two types of models are used for seasonal decomposition:

(1) Additive model: $y_t = S_t + T_t + C_t + I_t$ or $y_t = S_t + TC_t + I_t$

(2) Multiplicative model: $y_t = S_t \times T_t \times C_t \times I_t$ or $y_t = S_t \times TC_t \times I_t$

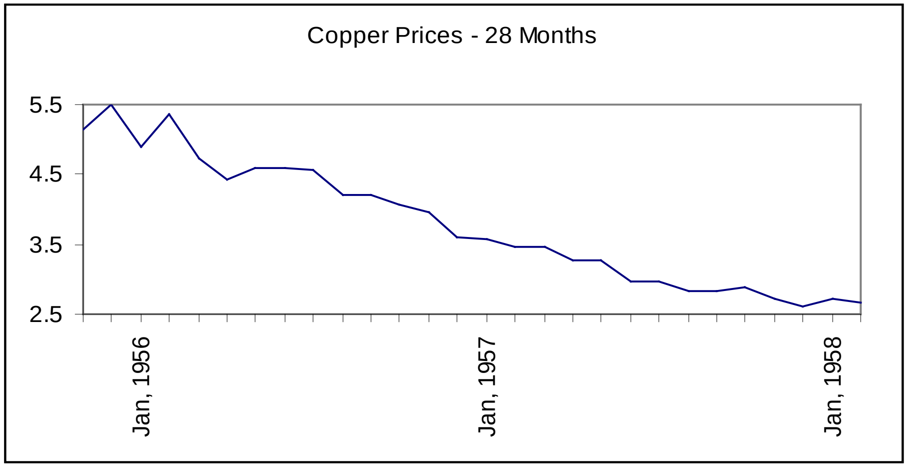
We will discuss multiplicative models since these are typically the models that are used in practice. Multiplicative models are used when there is increasing volatility such as in the sales data from The Gap (see page 1).

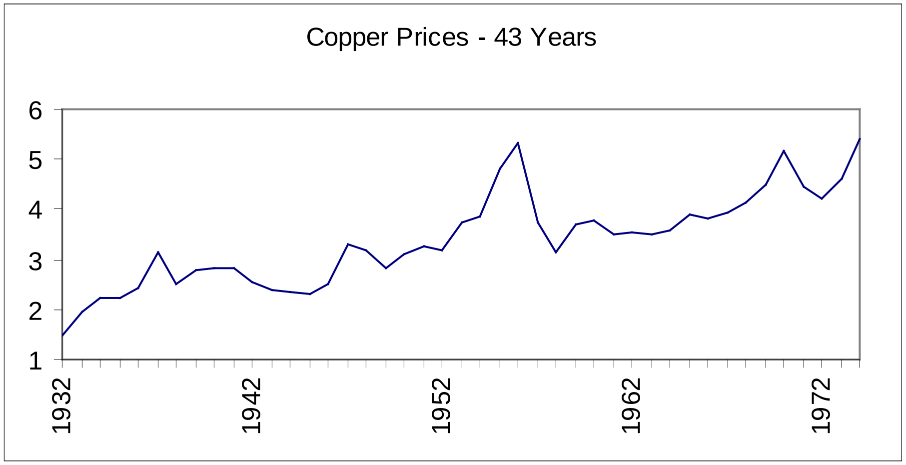
The reason for the increasing volatility in *Sales* for The Gap is that *Sales* grow about 30% on average between the third and fourth quarter every year.

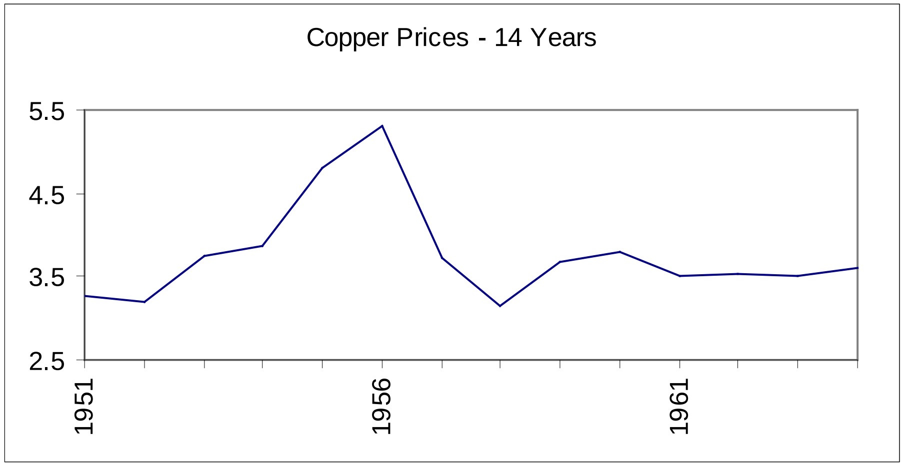
At the beginning of the series, where *Sales* are approximately $100 million, a 30% change in *Sales* from the third to fourth quarter gives a change of (0.30)($100m) = $30m.

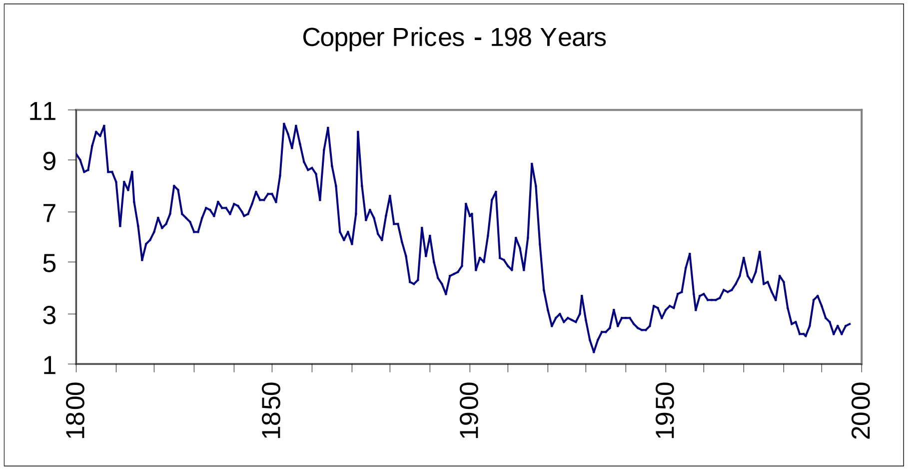
At the end of the series, where *Sales* are approximately $1,000 million (i.e. $1 billion), a 30% change in *Sales* gives a change of (0.30)($1,000m) = $300m.

Copper prices (in constant 1997 dollars) plotted for different length periods





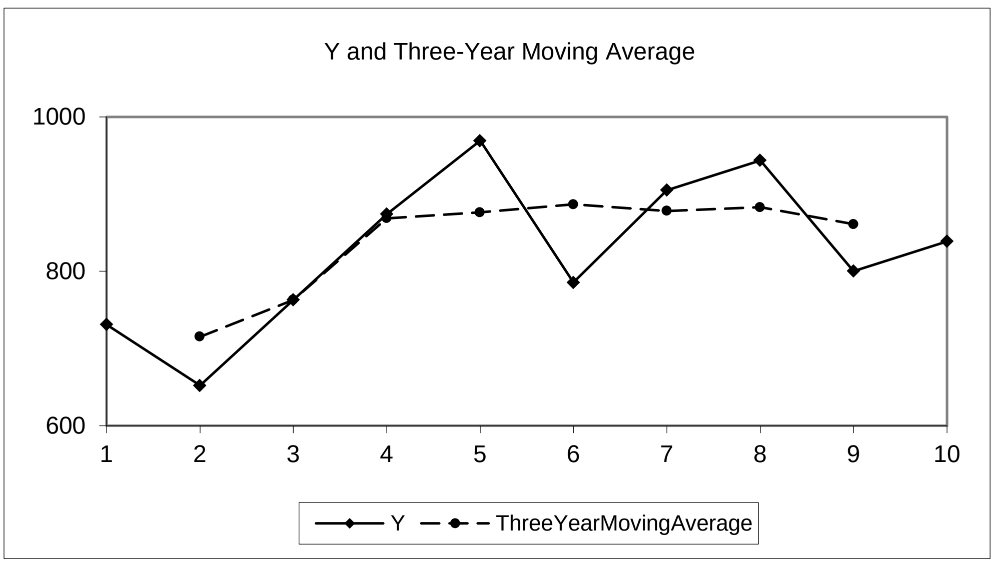




Calculation of Centered Moving Averages

Excel Spreadsheet for Computing Centered Moving Averages

Row	A	B	C	D
1	Y	ThreeYearMovingAverage $(y_{t-1} + y_t + y_{t+1})/3$	FiveYearMovingAverage $(y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2})/5$	SevenYearMovingAverage $(y_{t-3} + y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2} + y_{t+3})/7$
2	731.14			
3	652.10	715.40 **= AVERAGE(A2:A4)**		
4	762.95	763.06 **= AVERAGE(A3:A5)**	797.92 **= AVERAGE(A2:A6)**	
5	874.13	868.78	808.83 **= AVERAGE(A3:A7)**	795.88 **= AVERAGE(A2:A8)**
6	969.26	876.36	859.43	824.87 **= AVERAGE(A3:A9)**
7	785.69	886.69	895.59	873.48
8	905.11	878.18	880.83	879.72
9	943.75	883.07	854.77	
10	800.36	861.01		
11	838.92			



Classical Seasonal Decomposition for the Multiplicative Model: $y_t = S_t \times TC_t \times I_t$

From here on, the Trend/Cycle component will be denoted T_t (rather than TC_t).

Four Steps in a Classical Seasonal Decomposition Procedure for a Multiplicative Model

(1) Compute an initial estimate of the Trend/Cycle component T_t (the estimate is denoted $\hat{T}_t$) using the centered moving average

$$\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}.$$

(2) Compute $\frac{y_t}{\hat{T}_t}$. This is an estimate of the S_tI_t components.

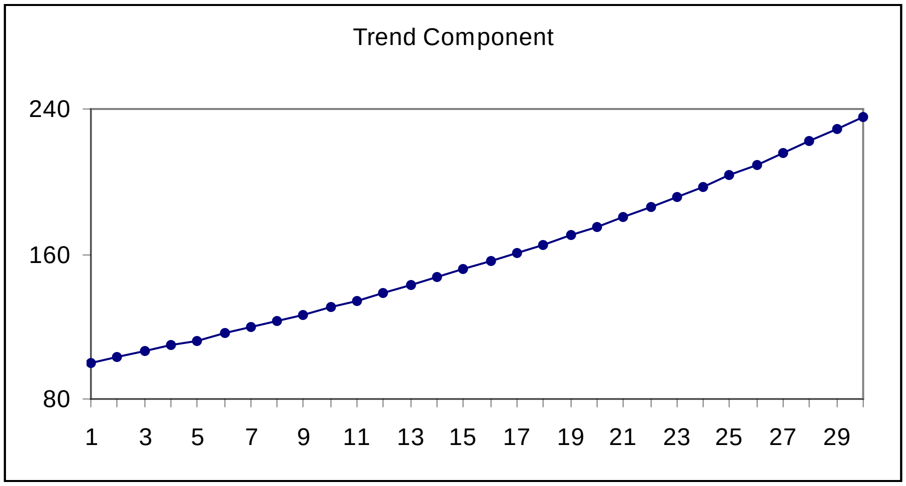
(3) Compute an estimate of S_t (denoted $\hat{S}_t$) by averaging the estimates of the S_tI_t components across corresponding quarters to average out the I_t components. For example, for quarterly data, the S_tI_t term is averaged across all the first quarters, all the second quarters, etc.

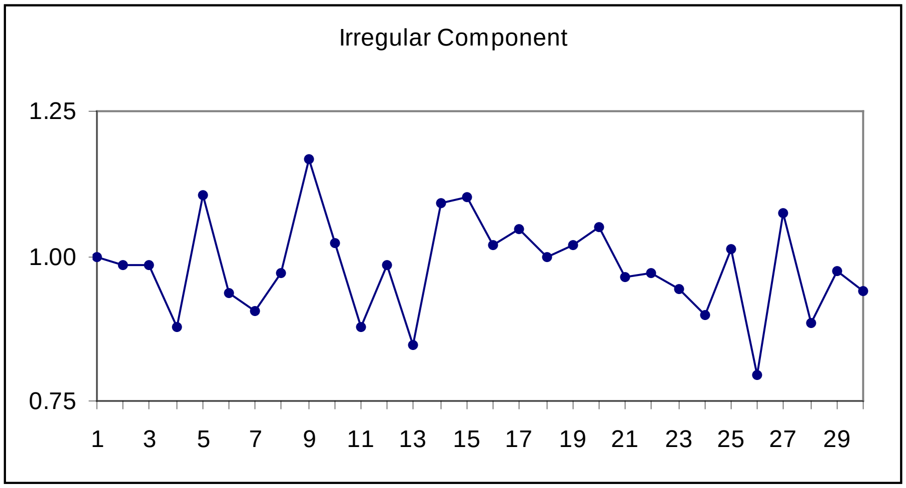
(4) Seasonally adjust the original data using $\frac{y_t}{\hat{S}_t}$.

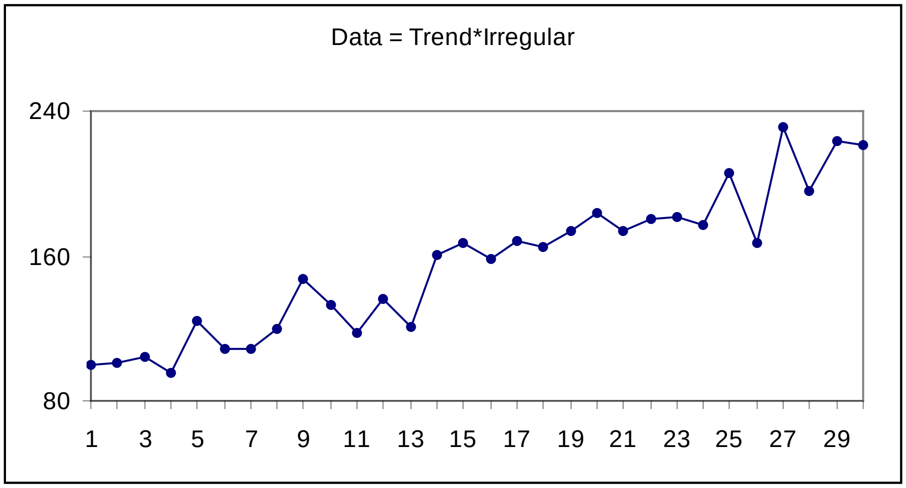
Important Question Regarding Step 1: How good an estimate is $\hat{T}_t$ of T_t?

To answer this question, first consider annual (simulated) data with no seasonality. Therefore, the model is $y_t = T_t \times I_t$.

Row	A	B	C	D
1	Year t	Trend T_t	Irregular I_t	Data=Trend*Irregular $y_t = T_t \times I_t$
2	1	100.000	0.997	99.667
3	2	103.000	0.984	101.327
4	3	106.090	0.984	104.417
5	4	109.273	0.877	95.780
6	5	112.551	1.105	124.321
7	6	115.927	0.935	108.401
8	7	119.405	0.906	108.197
9	8	122.987	0.970	119.267
10	9	126.677	1.166	147.718
11	10	130.477	1.022	133.412
12	11	134.392	0.878	118.024
13	12	138.423	0.983	136.056
14	13	142.576	0.845	120.534
15	14	146.853	1.092	160.304
16	15	151.259	1.103	166.884
17	16	155.797	1.018	158.640
18	17	160.471	1.045	167.730
19	18	165.285	1.000	165.214
20	19	170.243	1.020	173.636
21	20	175.351	1.050	184.046
22	21	180.611	0.962	173.813
23	22	186.029	0.970	180.536
24	23	191.610	0.945	181.030
25	24	197.359	0.897	177.050
26	25	203.279	1.012	205.629
27	26	209.378	0.796	166.698
28	27	215.659	1.074	231.577
29	28	222.129	0.883	196.135
30	29	228.793	0.976	223.208
31	30	235.657	0.939	221.227


Trend Component
240
160
80
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29


Irregular Component
1.25
1.00
0.75
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29


Data = Trend*Irregular
240
160
80
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29

To get an estimate of T_t we will compute a weighted moving average of the y_t's to average out the I_t components. In particular,

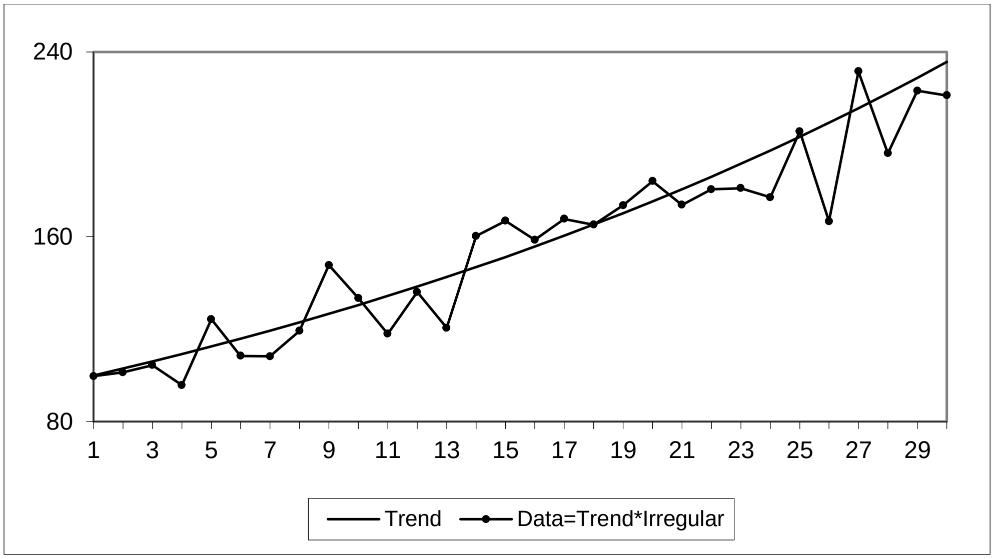
$$\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}.$$

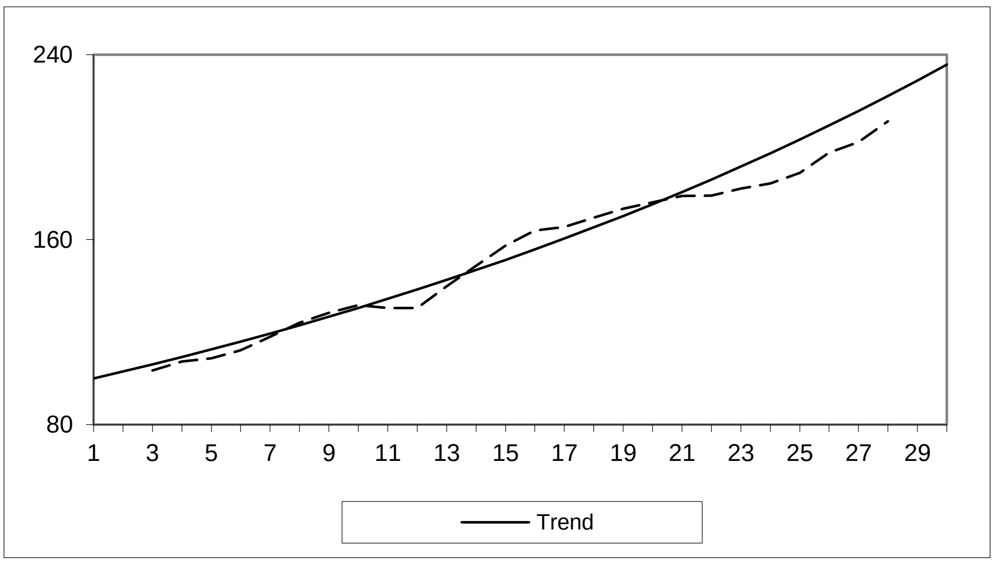
Suppose for the moment there is no irregular component (i.e. $I_t = 1$ in each period t). As shown below, even with no irregular component I_t we cannot recover T_t exactly, although we can come very close.

Row	A	B	C	D	E
1	Year t	Trend T_t	Irregular I_t	Data = Trend*Irregular $y_t = T_tI_t = T_t(1) = T_t$	Trend Estimate $\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}$
2	1	100.000	1	100.000	
3	2	103.000	1	103.000	
4	3	106.090	1	106.090	106.160 **= 0.125*D2 + 0.25*D3 + 0.25*D4 + 0.25*D5 + 0.125*D6**
5	4	109.273	1	109.273	109.344
6	5	112.551	1	112.551	112.625
7	6	115.927	1	115.927	116.003
8	7	119.405	1	119.405	119.483
9	8	122.987	1	122.987	123.068
10	9	126.677	1	126.677	126.760
11	10	130.477	1	130.477	130.563
12	11	134.392	1	134.392	134.480
13	12	138.423	1	138.423	138.514
14	13	142.576	1	142.576	142.670
15	14	146.853	1	146.853	146.950
16	15	151.259	1	151.259	151.358
17	16	155.797	1	155.797	155.899
18	17	160.471	1	160.471	160.576
19	18	165.285	1	165.285	165.393
---	---	---	---	---	---
25	24	197.359	1	197.359	197.488
26	25	203.279	1	203.279	203.413
27	26	209.378	1	209.378	209.515
28	27	215.659	1	215.659	215.800
29	28	222.129	1	222.129	222.274
30	29	228.793	1	228.793	
31	30	235.657	1	235.657	

Now consider the actual data $y_t = T_t \times I_t$. As shown below, we cannot recover T_t exactly because of the variability in the I_t component.

Row	A	B	C	D	E
1	Year t	Trend T_t	Irregular I_t	Data=Trend*Irregular $y_t = T_t \times I_t$	Trend Estimate $\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}$
2	1	100.000	0.997	99.667	
3	2	103.000	0.984	101.327	
4	3	106.090	0.984	104.417	103.380 **= 0.125*D2 + 0.25*D3 + 0.25*D4 + 0.25*D5 + 0.125*D6**
5	4	109.273	0.877	95.780	107.346
6	5	112.551	1.105	124.321	108.702
7	6	115.927	0.935	108.401	112.111
8	7	119.405	0.906	108.197	117.971
9	8	122.987	0.970	119.267	124.022
10	9	126.677	1.166	147.718	128.377
11	10	130.477	1.022	133.412	131.704
12	11	134.392	0.878	118.024	130.405
13	12	138.423	0.983	136.056	130.368
14	13	142.576	0.845	120.534	139.837
15	14	146.853	1.092	160.304	148.767
16	15	151.259	1.103	166.884	157.490
17	16	155.797	1.018	158.640	164.003
18	17	160.471	1.045	167.730	165.461
19	18	165.285	1.000	165.214	169.481
20	19	170.243	1.020	173.636	173.417
21	20	175.351	1.050	184.046	176.093
22	21	180.611	0.962	173.813	178.932
23	22	186.029	0.970	180.536	178.982
24	23	191.610	0.945	181.030	182.084
25	24	197.359	0.897	177.050	184.332
26	25	203.279	1.012	205.629	188.920
27	26	209.378	0.796	166.698	197.624
28	27	215.659	1.074	231.577	202.207
29	28	222.129	0.883	196.135	211.221
30	29	228.793	0.976	223.208	
31	30	235.657	0.939	221.227	





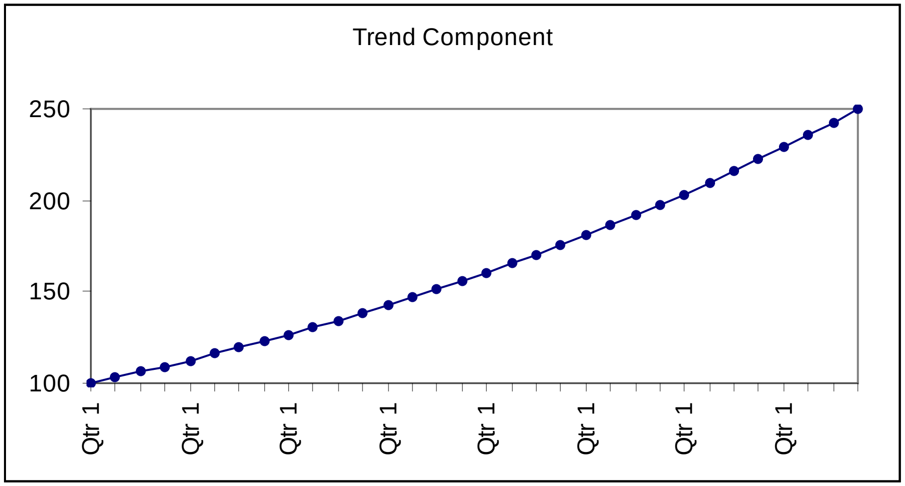
There are two sources of estimation error when estimating T_t:

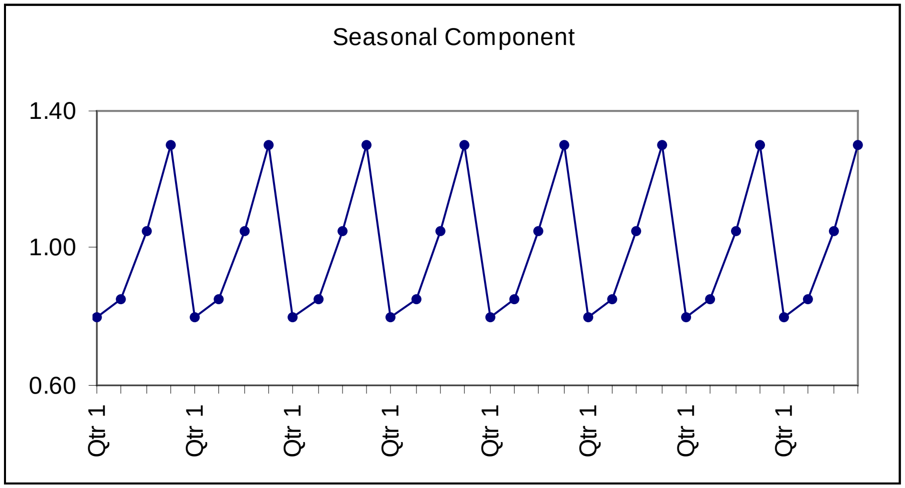
(1) Even with no irregular component the weighted moving average cannot recover T_t exactly (this is typically a very small source of error). The only time T_t can be recovered exactly is when $T_t = \alpha + \beta t$ (i.e. the trend is linear).

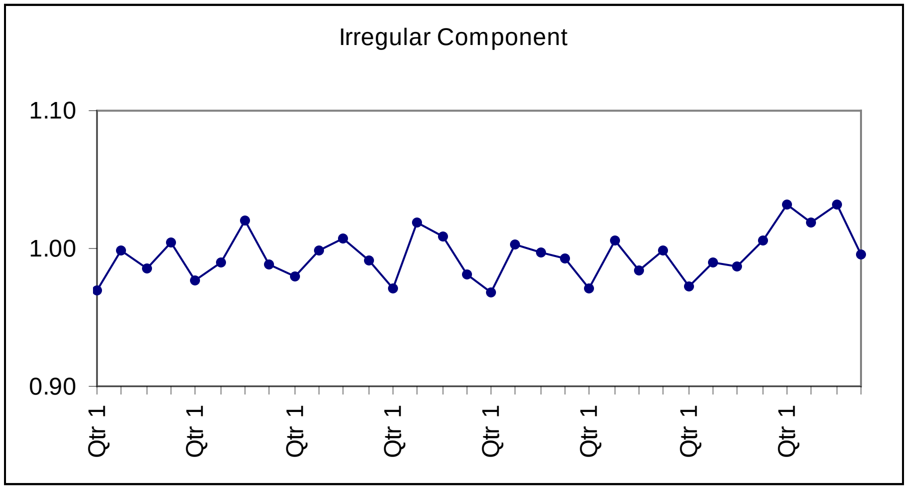
(2) The I_t component cannot be completely averaged out with a short weighted moving average. The size of the estimation error due to I_t depends on how much variability there is in I_t (i.e. it depends on what the variance of I_t is).

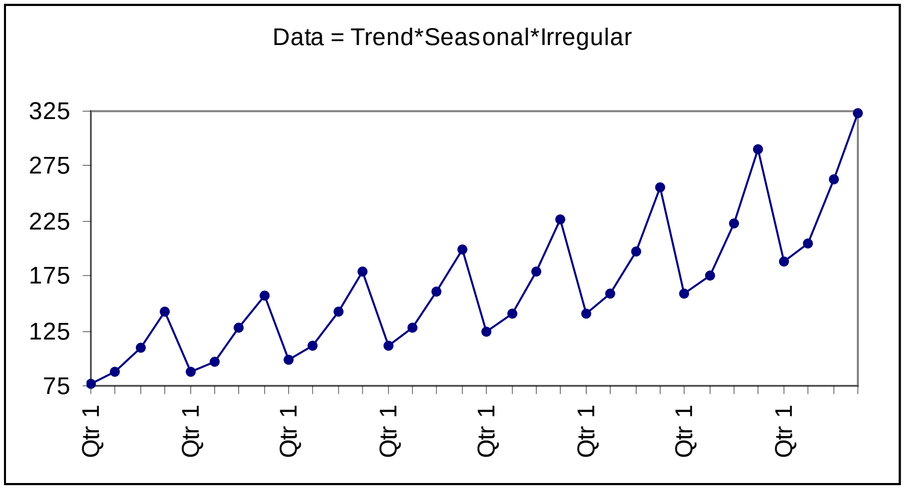
Now consider estimating T_t in a full seasonal model (using simulated quarterly data). The model is $y_t = T_t \times S_t \times I_t$.

Row	A	B	C	D	E	F
1	Time t	Quarter	Trend T_t	Seasonal S_t	Irregular I_t	Data=Trend*Seasonal*Irregular $y_t = T_t \times S_t \times I_t$
2	1	Qtr 1	100.000	0.80	0.969	77.525
3	2	Qtr 2	103.000	0.85	0.998	87.418
4	3	Qtr 3	106.090	1.05	0.985	109.733
5	4	Qtr 4	109.273	1.30	1.005	142.756
6	5	Qtr 1	112.551	0.80	0.977	87.994
7	6	Qtr 2	115.927	0.85	0.989	97.485
8	7	Qtr 3	119.405	1.05	1.020	127.935
9	8	Qtr 4	122.987	1.30	0.988	157.940
10	9	Qtr 1	126.677	0.80	0.980	99.278
11	10	Qtr 2	130.477	0.85	0.998	110.694
12	11	Qtr 3	134.392	1.05	1.008	142.203
13	12	Qtr 4	138.423	1.30	0.991	178.392
14	13	Qtr 1	142.576	0.80	0.971	110.782
15	14	Qtr 2	146.853	0.85	1.020	127.260
16	15	Qtr 3	151.259	1.05	1.009	160.270
17	16	Qtr 4	155.797	1.30	0.981	198.618
18	17	Qtr 1	160.471	0.80	0.968	124.294
19	18	Qtr 2	165.285	0.85	1.003	140.892
20	19	Qtr 3	170.243	1.05	0.997	178.207
21	20	Qtr 4	175.351	1.30	0.992	226.157
22	21	Qtr 1	180.611	0.80	0.972	140.394
23	22	Qtr 2	186.029	0.85	1.006	159.011
24	23	Qtr 3	191.610	1.05	0.984	197.988
25	24	Qtr 4	197.359	1.30	0.998	256.067
26	25	Qtr 1	203.279	0.80	0.973	158.242
27	26	Qtr 2	209.378	0.85	0.990	176.244
28	27	Qtr 3	215.659	1.05	0.987	223.446
29	28	Qtr 4	222.129	1.30	1.006	290.539
30	29	Qtr 1	228.793	0.80	1.032	188.926
31	30	Qtr 2	235.657	0.85	1.019	204.113
32	31	Qtr 3	242.726	1.05	1.032	263.121
33	32	Qtr 4	250.008	1.30	0.996	323.686


Trend Component
250
200
150
100
Qtr 1


Seasonal Component
1.40
1.00
0.60
Qtr 1


Irregular Component
1.10
1.00
0.90
Qtr 1


Data = Trend*Seasonal*Irregular
325
275
225
175
125
75
Qtr 1

To get an estimate of the T_t component we will compute a weighted moving average of y_t to average out the S_t and I_t components:

$$\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}.$$

Suppose for the moment there is no seasonal or irregular component (i.e. $S_t = 1$ and $I_t = 1$ in each period t). As shown below, even with no irregular component I_t we cannot recover T_t exactly, although we can come very close.

Row	A	B	C	D	E	F	G
1	Quarter	Time t	Trend T_t	Seasonal S_t	Irregular I_t	Data = Trend*Seasonal*Irregular $y_t = T_t \times S_t \times I_t$	Trend Estimate- No Seasonal Or Irregular Comp $\hat{T}_t = 0.125y_{t-2}+0.25y_{t-1}+0.25y_t+0.25y_{t+1}+0.125y_{t+2}$
2	Qtr 1	1	100.000	1.00	1.000	100.000	
3	Qtr 2	2	103.000	1.00	1.000	103.000	
4	Qtr 3	3	106.090	1.00	1.000	106.090	106.160 **= 0.125*F2 + 0.25*F3 + 0.25*F4 + 0.25*F5 + 0.125*F6**
5	Qtr 4	4	109.273	1.00	1.000	109.273	109.344
6	Qtr 1	5	112.551	1.00	1.000	112.551	112.625
7	Qtr 2	6	115.927	1.00	1.000	115.927	116.003
8	Qtr 3	7	119.405	1.00	1.000	119.405	119.483
9	Qtr 4	8	122.987	1.00	1.000	122.987	123.068
10	Qtr 1	9	126.677	1.00	1.000	126.677	126.760
11	Qtr 2	10	130.477	1.00	1.000	130.477	130.563
12	Qtr 3	11	134.392	1.00	1.000	134.392	134.480
13	Qtr 4	12	138.423	1.00	1.000	138.423	138.514
14	Qtr 1	13	142.576	1.00	1.000	142.576	142.670
15	Qtr 2	14	146.853	1.00	1.000	146.853	146.950
16	Qtr 3	15	151.259	1.00	1.000	151.259	151.358
17	Qtr 4	16	155.797	1.00	1.000	155.797	155.899
18	Qtr 1	17	160.471	1.00	1.000	160.471	160.576
19	Qtr 2	18	165.285	1.00	1.000	165.285	165.393
20	Qtr 3	19	170.243	1.00	1.000	170.243	170.355
---	---	---	---	---	---	---	---
30	Qtr 1	29	228.793	1.00	1.000	228.793	228.943
31	Qtr 2	30	235.657	1.00	1.000	235.657	235.811
32	Qtr 3	31	242.726	1.00	1.000	242.726	
33	Qtr 4	32	250.008	1.00	1.000	250.008	

Now suppose there is a seasonal component but no irregular component (i.e. I_t = 1 in each period so $y_t = T_t \times S_t$). As shown below, we can obtain an accurate (but not perfect) estimate of T_t.

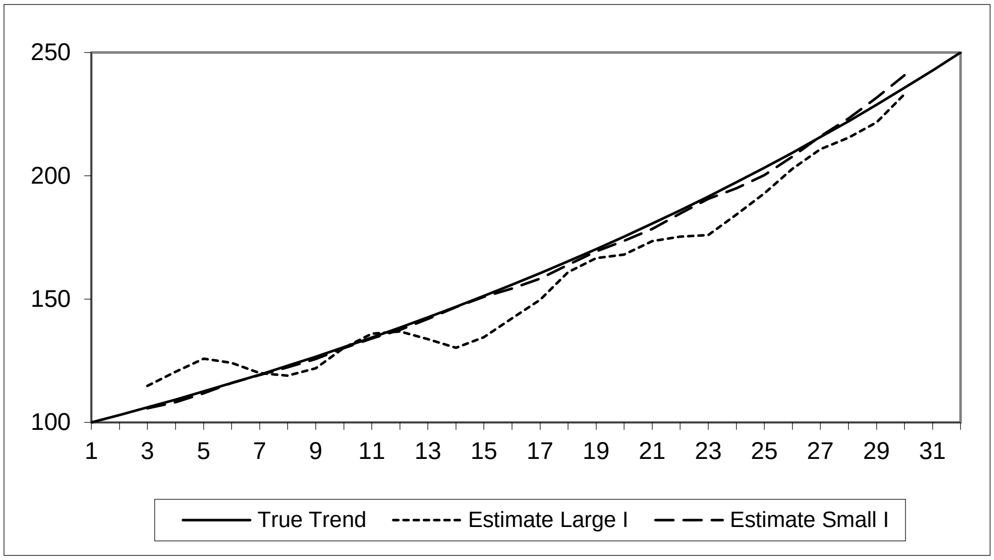
Row	A	B	C	D	E	F	G
1	Quarter	Time t	Trend T_t	Seasonal S_t	Irregular I_t	Data = Trend*Seasonal*Irregular $y_t = T_t \times S_t \times I_t$	Trend Estimate- No Irregular Component $\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}$
2	Qtr 1	1	100.000	0.80	1.000	80.000	
3	Qtr 2	2	103.000	0.85	1.000	87.550	
4	Qtr 3	3	106.090	1.05	1.000	111.395	106.505 **= 0.125*F2 + 0.25*F3 + 0.25*F4 + 0.25*F5 + 0.125*F6**
5	Qtr 4	4	109.273	1.30	1.000	142.055	109.133
6	Qtr 1	5	112.551	0.80	1.000	90.041	112.255
7	Qtr 2	6	115.927	0.85	1.000	98.538	116.231
8	Qtr 3	7	119.405	1.05	1.000	125.375	119.872
9	Qtr 4	8	122.987	1.30	1.000	159.884	122.831
10	Qtr 1	9	126.677	0.80	1.000	101.342	126.344
11	Qtr 2	10	130.477	0.85	1.000	110.906	130.819
12	Qtr 3	11	134.392	1.05	1.000	141.111	134.917
13	Qtr 4	12	138.423	1.30	1.000	179.950	138.247
14	Qtr 1	13	142.576	0.80	1.000	114.061	142.201
15	Qtr 2	14	146.853	0.85	1.000	124.825	147.238
16	Qtr 3	15	151.259	1.05	1.000	158.822	151.850
17	Qtr 4	16	155.797	1.30	1.000	202.536	155.598
18	Qtr 1	17	160.471	0.80	1.000	128.377	160.048
19	Qtr 2	18	165.285	0.85	1.000	140.492	165.717
20	Qtr 3	19	170.243	1.05	1.000	178.755	170.909
21	Qtr 4	20	175.351	1.30	1.000	227.956	175.127
22	Qtr 1	21	180.611	0.80	1.000	144.489	180.136
23	Qtr 2	22	186.029	0.85	1.000	158.125	186.516
24	Qtr 3	23	191.610	1.05	1.000	201.191	192.360
25	Qtr 4	24	197.359	1.30	1.000	256.566	197.107
26	Qtr 1	25	203.279	0.80	1.000	162.624	202.744
27	Qtr 2	26	209.378	0.85	1.000	177.971	209.926
28	Qtr 3	27	215.659	1.05	1.000	226.442	216.502
29	Qtr 4	28	222.129	1.30	1.000	288.768	221.846
30	Qtr 1	29	228.793	0.80	1.000	183.034	228.191
31	Qtr 2	30	235.657	0.85	1.000	200.308	236.273
32	Qtr 3	31	242.726	1.05	1.000	254.863	
33	Qtr 4	32	250.008	1.30	1.000	325.010	

Now consider the full model $y_t = T_t \times S_t \times I_t$. If the variability in I_t is not large we can obtain a fairly accurate estimate of T_t.

Row	A	B	C	D	E	F	G
1	Quarter	Time t	Trend T_t	Seasonal S_t	Irregular I_t	Data = Trend*Seasonal*Irregular $y_t = T_t \times S_t \times I_t$	Trend Estimate $\hat{T}_t = 0.125y_{t-2}+0.25y_{t-1}+0.25y_t+0.25y_{t+1}+0.125y_{t+2}$
2	Qtr 1	1	100.000	0.80	0.969	77.525	
3	Qtr 2	2	103.000	0.85	0.998	87.418	
4	Qtr 3	3	106.090	1.05	0.985	109.733	105.667 **= 0.125*F2 + 0.25*F3 + 0.25*F4 + 0.25*F5 + 0.125*F6**
5	Qtr 4	4	109.273	1.30	1.005	142.756	108.234
6	Qtr 1	5	112.551	0.80	0.977	87.994	111.767
7	Qtr 2	6	115.927	0.85	0.989	97.485	115.941
8	Qtr 3	7	119.405	1.05	1.020	127.935	119.249
9	Qtr 4	8	122.987	1.30	0.988	157.940	122.311
10	Qtr 1	9	126.677	0.80	0.980	99.278	125.746
11	Qtr 2	10	130.477	0.85	0.998	110.694	130.086
12	Qtr 3	11	134.392	1.05	1.008	142.203	134.080
13	Qtr 4	12	138.423	1.30	0.991	178.392	137.589
14	Qtr 1	13	142.576	0.80	0.971	110.782	141.918
15	Qtr 2	14	146.853	0.85	1.020	127.260	146.704
16	Qtr 3	15	151.259	1.05	1.009	160.270	150.922
17	Qtr 4	16	155.797	1.30	0.981	198.618	154.315
18	Qtr 1	17	160.471	0.80	0.968	124.294	158.261
19	Qtr 2	18	165.285	0.85	1.003	140.892	163.945
20	Qtr 3	19	170.243	1.05	0.997	178.207	169.400
21	Qtr 4	20	175.351	1.30	0.992	226.157	173.677
22	Qtr 1	21	180.611	0.80	0.972	140.394	178.415
23	Qtr 2	22	186.029	0.85	1.006	159.011	184.626
24	Qtr 3	23	191.610	1.05	0.984	197.988	190.596
25	Qtr 4	24	197.359	1.30	0.998	256.067	194.981
26	Qtr 1	25	203.279	0.80	0.973	158.242	200.317
27	Qtr 2	26	209.378	0.85	0.990	176.244	207.809
28	Qtr 3	27	215.659	1.05	0.987	223.446	215.953
29	Qtr 4	28	222.129	1.30	1.006	290.539	223.272
30	Qtr 1	29	228.793	0.80	1.032	188.926	231.715
31	Qtr 2	30	235.657	0.85	1.019	204.113	240.818
32	Qtr 3	31	242.726	1.05	1.032	263.121	
33	Qtr 4	32	250.008	1.30	0.996	323.686	

If the variability in I_t is large, this makes it more difficult to obtain an accurate estimate of T_t.

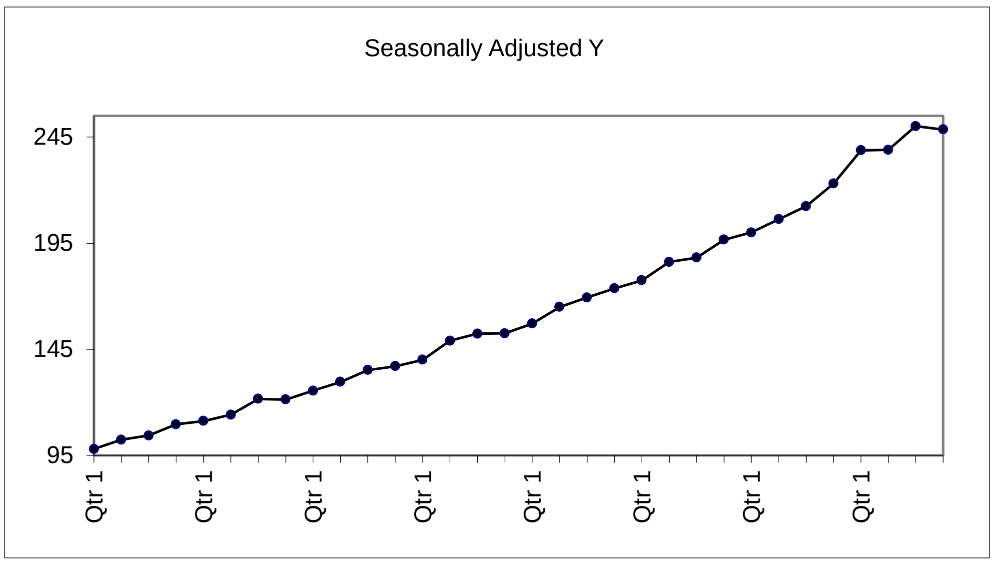
Row	A	B	C	D	E	F	G
1	Quarter	Time t	Trend T_t	Seasonal S_t	Irregular I_t	Data = Trend*Seasonal*Irregular $y_t = T_t \times S_t \times I_t$	Trend Estimate $\hat{T}_t = 0.125y_{t-2}+0.25y_{t-1}+0.25y_t+0.25y_{t+1}+0.125y_{t+2}$
2	Qtr 1	1	100.000	0.80	1.009	80.714	
3	Qtr 2	2	103.000	0.85	0.922	80.697	
4	Qtr 3	3	106.090	1.05	0.983	109.522	114.826 **= 0.125*F2 + 0.25*F3 + 0.25*F4 + 0.25*F5 + 0.125*F6**
5	Qtr 4	4	109.273	1.30	1.250	177.543	120.479
6	Qtr 1	5	112.551	0.80	1.137	102.367	125.832
7	Qtr 2	6	115.927	0.85	1.058	104.269	124.191
8	Qtr 3	7	119.405	1.05	1.027	128.774	119.996
9	Qtr 4	8	122.987	1.30	0.908	145.165	118.900
10	Qtr 1	9	126.677	0.80	0.998	101.186	121.969
---	---	---	---	---	---	---	---
31	Qtr 2	30	235.657	0.85	1.034	207.215	233.089
32	Qtr 3	31	242.726	1.05	0.975	248.552	
33	Qtr 4	32	250.008	1.30	1.043	338.963	



The full spreadsheet to implement classical decomposition in a multiplicative model is given on the next page.

Row	A	B	C	D	E	F	G	H
1	Quarter	Data y_t	T Estimate $\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}$	SxI Estimate $\frac{y_t}{\hat{T}_t}$	SeasonalSum **This column sums the S × I components for corresponding quarters**	SeasonalAvg **This column is an average of the S × I components for corresponding quarters**	S Estimate $\hat{S}_t$ **This column forces the S components to average to one**	SeasonalAdjY $\frac{y_t}{\hat{S}_t}$
2	Qtr 1	77.525					0.791 **Copied manually**	97.976 **= B2/G2**
3	Qtr 2	87.418					0.854 **Copied manually**	102.334 **= B3/G3**
4	Qtr 3	109.733	105.667 **= 0.125*B2 + 0.25*(B3+B4+B5)+ 0.125*B6**	1.038 **= B4/C4**	7.359 **= SUMIF(A$4:A$31, "Qtr 3",D$4:D$31)**	1.051 **= E4/7**	1.052 **= 4*F4/SUM(F$4:F$7)**	104.303 **= B4/G4**
5	Qtr 4	142.756	108.234 **= 0.125*B3 + 0.25*(B4+B5+B6)+ 0.125*B7**	1.319 **= B5/C5**	9.111 **= SUMIF(A$4:A$31, "Qtr 4",D$4:D$31)**	1.302 **= E5/7**	1.302 **= 4*F5/SUM(F$4:F$7)**	109.608 **= B5/G5**
6	Qtr 1	87.994	111.767 **= 0.125*B4 + 0.25*(B5+B6+B7)+ 0.125*B8**	0.787 **= B6/C6**	5.535 **= SUMIF(A$4:A$31, "Qtr 1",D$4:D$31)**	0.791 **= E6/7**	0.791 **= 4*F6/SUM(F$4:F$7)**	111.207 **= B6/G6**
7	Qtr 2	97.485	115.941 **= 0.125*B5 + 0.25*(B6+B7+B8)+ 0.125*B9**	0.841 **= B7/C7**	5.976 **= SUMIF(A$4:A$31, "Qtr 2",D$4:D$31)**	0.854 **= E7/7**	0.854 **= 4*F7/SUM(F$4:F$7)**	114.119 **= B7/G7**
8	Qtr 3	127.935	119.249	1.073			1.052 **Copied manually**	121.604 **= B8/G8**
9	Qtr 4	157.940	122.311	1.291			1.302 **Copied manually**	121.266 **= B9/G9**
10	Qtr 1	99.278	125.746	0.790			0.791	125.468
11	Qtr 2	110.694	130.086	0.851			0.854	129.582
12	Qtr 3	142.203	134.080	1.061			1.052	135.166
13	Qtr 4	178.392	137.589	1.297			1.302	136.969
14	Qtr 1	110.782	141.918	0.781			0.791	140.006
15	Qtr 2	127.260	146.704	0.867			0.854	148.974
16	Qtr 3	160.270	150.922	1.062			1.052	152.339
17	Qtr 4	198.618	154.315	1.287			1.302	152.498
18	Qtr 1	124.294	158.261	0.785			0.791	157.083

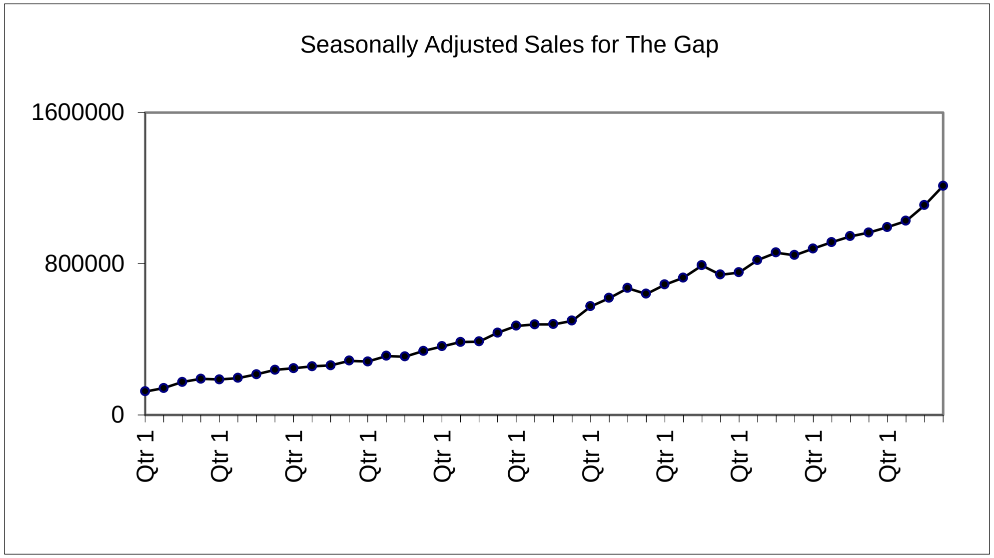
19	Qtr 2	140.892	163.945	0.859			0.854	164.931
20	Qtr 3	178.207	169.400	1.052			1.052	169.388
21	Qtr 4	226.157	173.677	1.302			1.302	173.643
22	Qtr 1	140.394	178.415	0.787			0.791	177.430
23	Qtr 2	159.011	184.626	0.861			0.854	186.142
24	Qtr 3	197.988	190.596	1.039			1.052	188.190
25	Qtr 4	256.067	194.981	1.313			1.302	196.607
26	Qtr 1	158.242	200.317	0.790			0.791	199.986
27	Qtr 2	176.244	207.809	0.848			0.854	206.316
28	Qtr 3	223.446	215.953	1.035			1.052	212.388
29	Qtr 4	290.539	223.272	1.301			1.302	223.075
30	Qtr 1	188.926	231.715	0.815			0.791	238.765
31	Qtr 2	204.113	240.818	0.848			0.854	238.939
32	Qtr 3	263.121					1.052	250.099
33	Qtr 4	323.686					1.302	248.525



Excel Spreadsheet for Classical Seasonal Decomposition of Sales for The Gap

Row	A	B	C	D	E	F	G	H
1	Quarter	Sales y_t	T Estimate $\hat{T}_t = 0.125y_{t-2} + 0.25y_{t-1} + 0.25y_t + 0.25y_{t+1} + 0.125y_{t+2}$	SxI Estimate $\frac{y_t}{\hat{T}_t}$	SeasonalSum **This column sums the S × I components for corresponding quarters**	SeasonalAvg **This column is an average of the S × I components for corresponding quarters**	S Estimate $\hat{S}_t$ **This column forces the S components to average to one**	SeasonalAdjSales $\frac{y_t}{\hat{S}_t}$
2	Qtr 1	105715					0.854 **Copied manually**	123784.73 **= B2/G2**
3	Qtr 2	120136					0.846 **Copied manually**	141974.74 **= B3/G3**
4	Qtr 3	181669	168616.375 **= 0.125*B2 + 0.25*(B3+B4+B5)+ 0.125*B6**	1.077 **= B4/C4**	10.457 **= SUMIF(A$4:A$43, "Qtr 3",D$4:D$43)**	1.046 **= E4/10**	1.046 **= 4*F4/SUM(F$4:F$7)**	173688.51 **= B4/G4**
5	Qtr 4	239813	180977.500 **= 0.125*B3 + 0.25*(B4+B5+B6)+ 0.125*B7**	1.325 **= B5/C5**	12.536 **= SUMIF(A$4:A$43, "Qtr 4",D$4:D$43)**	1.254 **= E5/10**	1.254 **= 4*F5/SUM(F$4:F$7**	191261.13 **= B5/G5**
6	Qtr 1	159980	191946.875 **= 0.125*B4 + 0.25*(B5+B6+B7)+ 0.125*B8**	0.833 **= B6/C6**	8.538 **= SUMIF(A$4:A$43, "Qtr 1",D$4:D$43)**	0.854 **= E6/10**	0.854 **= 4*F6/SUM(F$4:F$7**	187325.18 **= B6/G6**
7	Qtr 2	164760	204670.250 **= 0.125*B5 + 0.25*(B6+B7+B8)+ 0.125*B9**	0.805 **= B7/C7**	8.460 **= SUMIF(A$4:A$43, "Qtr 2",D$4:D$43)**	0.846 **= E7/10**	0.846 **= 4*F7/SUM(F$4:F$7**	194710.64 **= B7/G7**
8	Qtr 3	224800	218387.250	1.029			1.046 **Copied manually**	214924.82 **= B8/G8**
9	Qtr 4	298469	231396.375	1.290			1.254 **Copied manually**	238041.80 **= B9/G9**
10	Qtr 1	211060	244122.500	0.865			0.854	247136.22
11	Qtr 2	217753	257864.875	0.844			0.846	257336.89
12	Qtr 3	273616	269291.250	1.016			1.046	261596.40
13	Qtr 4	359592	278899.125	1.289			1.254	286790.00
14	Qtr 1	241348	290863.000	0.830			0.854	282601.31
15	Qtr 2	264328	305014.625	0.867			0.846	312378.46
16	Qtr 3	322752	321596.375	1.004			1.046	308573.92
17	Qtr 4	423669	337869.875	1.254			1.254	337894.15

18	Qtr 1	309925	355927.375	0.871			0.854	362900.09
19	Qtr 2	325939	381466.250	0.854			0.846	385189.32
20	Qtr 3	405601	408204.375	0.994			1.046	387783.47
21	Qtr 4	545131	429641.875	1.269			1.254	434765.29
---	---	---	---	---	---	---	---	---
38	Qtr 1	751670	882078.125	0.852			0.854	880152.00
39	Qtr 2	773131	912036.750	0.848			0.846	913673.43
40	Qtr 3	988346	942914.000	1.048			1.046	944929.22
41	Qtr 4	1.21E+06	966964.125	1.251			1.254	965026.77
42	Qtr 1	848688	1000343.750	0.848			0.854	993753.17
43	Qtr 2	868514	1060550.500	0.819			0.846	1026395.48
44	Qtr 3	1.16E+06					1.046	1109042.68
45	Qtr 4	1.52E+06					1.254	1212265.03



Algebra used to find the normalizing constant for the seasonal factors in Column G

We want to find a normalizing constant, denoted *c*, so that when we divide each of the seasonal factors by *c* they average to one, i.e. we want to find *c* so that

$$\frac{\frac{S_1}{c}+\frac{S_2}{c}+\frac{S_3}{c}+\frac{S_4}{c}}{4}=1.$$

Using some algebra, this gives

$$\frac{\frac{S_1+S_2+S_3+S_4}{c}}{4}=1$$

$$\frac{\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)}{c}=1$$

$$\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)=c$$

Therefore, the normalized seasonal factors are

$$\frac{S_1}{\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)},\ \frac{S_2}{\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)},\ \frac{S_3}{\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)}\ \text{and}\ \frac{S_4}{\frac{1}{4}\left(S_1+S_2+S_3+S_4\right)},$$

or equivalently,

$$\frac{4S_1}{S_1+S_2+S_3+S_4},\ \frac{4S_2}{S_1+S_2+S_3+S_4},\ \frac{4S_3}{S_1+S_2+S_3+S_4}\ \text{and}\ \frac{4S_4}{S_1+S_2+S_3+S_4}.$$

This is the formula used in Column G in the above spreadsheets.

Seasonally adjust sales for The Gap using StatTools

StatTools instructions to seasonally adjust data

To use StatTools in Excel, you must first open it outside Excel by clicking on its icon. Then inside Excel, click on *StatTools* in the menu at the top of the Excel screen. (Please note that "click" will always refer to a left click; if a right click is needed, I'll write "right click").

To run an analysis using StatTools, you must first create a StatTools data set containing the variable(s) you want to analyze. To do this, click on *Data Set Manager* in the top left hand corner of the StatTools screen. In the *Data Set Manager* dialog box, click on *New*, click on the *Select the range* icon immediately to the right of the *Excel Range* box, highlight the column in the Excel worksheet containing quarterly sales for The Gap, click *OK*, and then click *OK* again.

To seasonally adjust sales, click on *Time Series and Forecasting* at the top of the StatTools screen, and then click on *Forecast*. In the *StatTools-Forecast* dialog box, click the box next to *Sales*, click on *Time Scale* (at this point in the semester don't worry about any of the other options in the *Forecast* dialog box), click on *Quarterly* in the new dialog box that comes up, click on *Deseasonalize*, and then click *OK*.

The seasonally adjusted sales data will be put into a new worksheet labelled *Forecast* (the seasonally adjusted data will be in a column labelled *Deseason Sales* part way down the worksheet). We will discuss the other columns in this worksheet as the semester goes along.

The columns in the worksheet labelled *Forecast* related to seasonally adjusted sales are:

	Season	Deseason
Sales	**Index**	**Sales**
105715.00	0.85	123784.74
120136.00	0.85	141974.74
181669.00	1.05	173688.52
239813.00	1.25	191261.14
159980.00	0.85	187325.19
164760.00	0.85	194710.65
224800.00	1.05	214924.83
298469.00	1.25	238041.81
⋮	⋮	⋮
751670.00	0.85	880152.05
773131.00	0.85	913673.47
988346.00	1.05	944929.27
1210000.00	1.25	965026.82
848688.00	0.85	993753.22
868514.00	0.85	1026395.53
1160000.00	1.05	1109042.73
1520000.00	1.25	1212265.09

Seasonally adjust sales for The Gap using R

IMPORTANT:

You are not responsible for running R scripts in this class.
I have only included the R script for this analysis in case you are interested.

R script to decompose sales for The Gap into trend, seasonal and irregular (random) components

```
#####################################################################################################################
#
#   You must set the working directory properly
#   If the R package ggplot2 is not installed then it must be installed by typing the command (at the "> prompt"):
#      install.packages("ggplot2")
#   You must type (at the "> prompt"): library (ggplot2)
#   To run, type (at the "> prompt"): source("ClassicalSeasonalDecomposition_Sales_TheGap.R")
#     (where ClassicalSeasonalDecomposition_Sales_TheGap.R is the name of file containing the R script given below)
#
#####################################################################################################################

#
#   Open file for output
#
sink ("C:/Users/shivelyt/Box Sync/Courses/372_Spring2015/R_Scripts/Classical_Seasonal_Decomposition/00Practice.txt",
append=FALSE, split=TRUE)
#
#   Read data
#
file <- "Sales_TheGap_1985-95.dat"
Sales_table <- read.table(file, header = FALSE, sep = "")
colnames(Sales_table) <- c("Time", "Sales")
n_obs = nrow(Sales_table)
cat ("Number of observations is:", n_obs, "\n", "\n")
print(head(Sales_table))
#
#   Plot Sales vs. Time
#
g = ggplot()
```

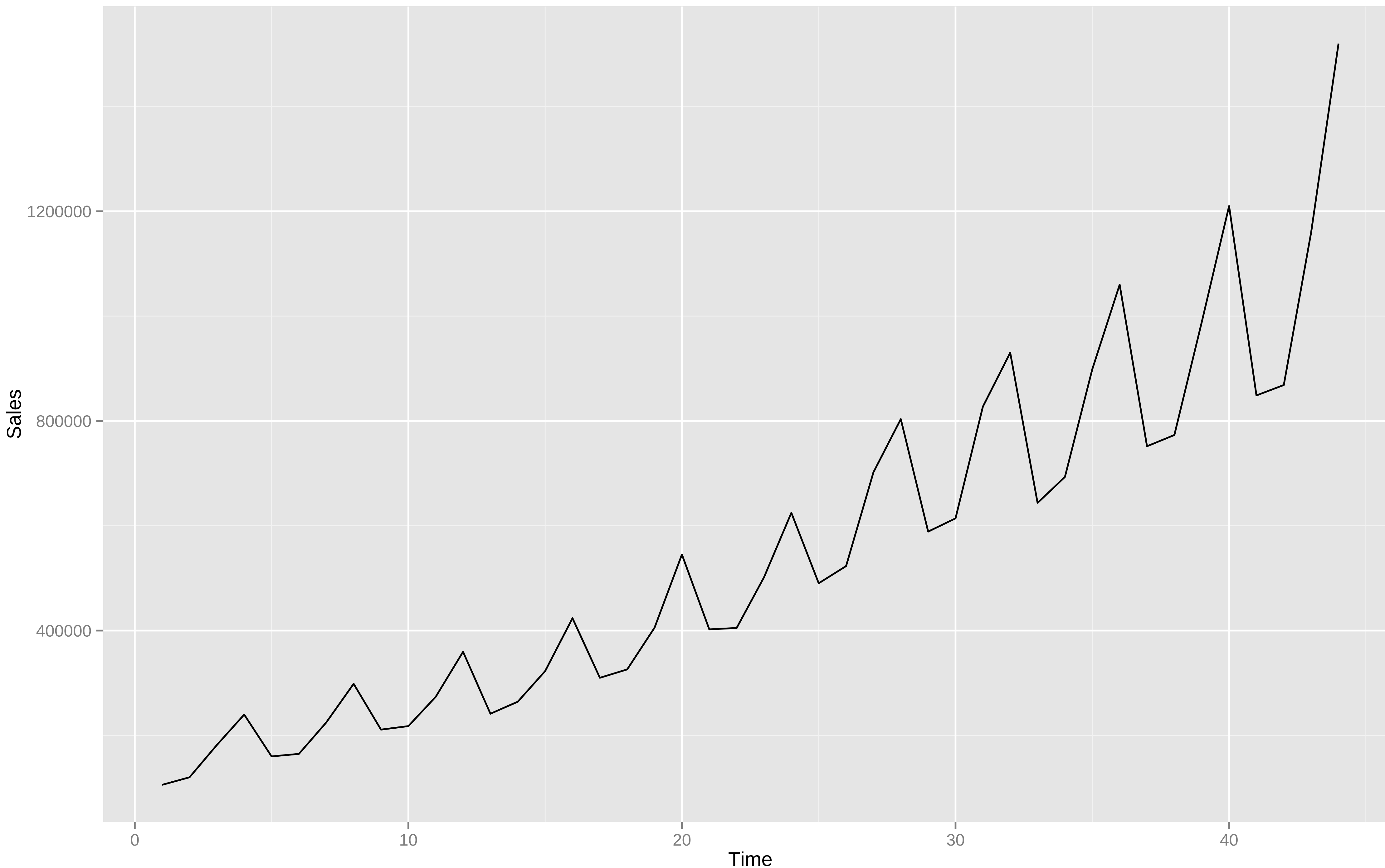
```
g <- g + geom_line(data=Sales_table, aes(Time,Sales), color="Black", lty=1)
ggsave('plot_Sales.pdf', g)
shell.exec(file.path(getwd(), "plot_Sales.pdf"))
#
#   Save data as a time series object
#
Sales_time_series <- ts(Sales_table[2], frequency=4)
colnames(Sales_time_series) <- "Sales"
#
#   Decompose Sales into trend, seasonal and irregular components, and plot and print components
#
Sales_time_series_components <- decompose(Sales_time_series, type="multiplicative")
pdf("plot_seasonal_components.pdf")
plot(Sales_time_series_components)
shell.exec(file.path(getwd(), "plot_seasonal_components.pdf"))
dev.off ()
print (Sales_time_series_components)
#
#   Close file for output
#
closeAllConnections()
```

Output from R

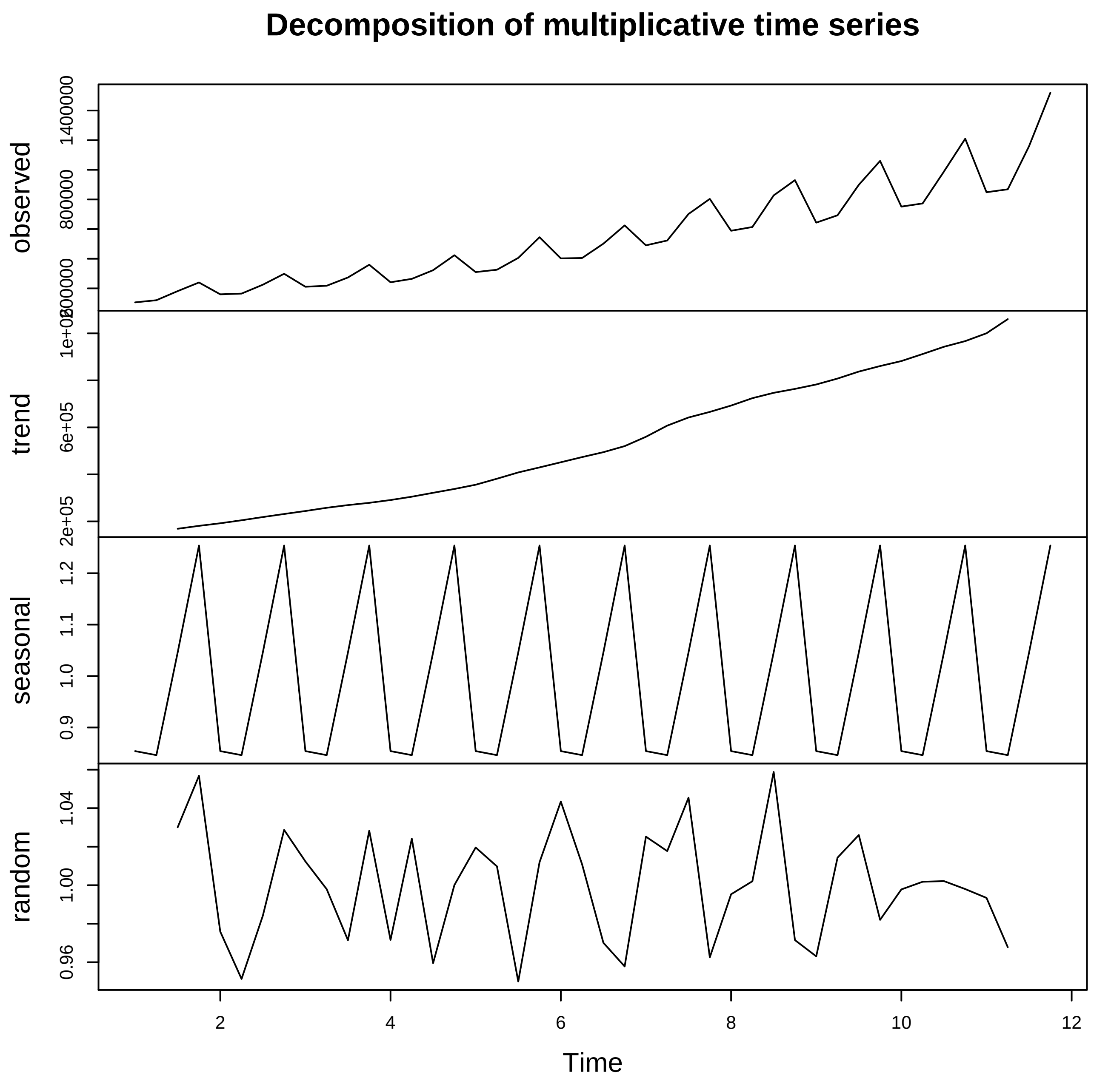
```
Number of observations is: 44

  Time  Sales
1    1 105715
2    2 120136
3    3 181669
4    4 239813
5    5 159980
6    6 164760

$x
      Qtr1    Qtr2    Qtr3    Qtr4
1   105715  120136  181669  239813
2   159980  164760  224800  298469
3   211060  217753  273616  359592
4   241348  264328  322752  423669
```

```
5   309925  325939  405601  545131
6   402368  404996  501690  624726
7   490300  523056  702052  803485
8   588864  614114  827222  930209
9   643580  693192  898677 1060000
10  751670  773131  988346 1210000
11  848688  868514 1160000 1520000

$seasonal
        Qtr1      Qtr2      Qtr3      Qtr4
1  0.8540229 0.8461787 1.0459471 1.2538512
2  0.8540229 0.8461787 1.0459471 1.2538512
3  0.8540229 0.8461787 1.0459471 1.2538512
4  0.8540229 0.8461787 1.0459471 1.2538512
5  0.8540229 0.8461787 1.0459471 1.2538512
6  0.8540229 0.8461787 1.0459471 1.2538512
7  0.8540229 0.8461787 1.0459471 1.2538512
8  0.8540229 0.8461787 1.0459471 1.2538512
9  0.8540229 0.8461787 1.0459471 1.2538512
10 0.8540229 0.8461787 1.0459471 1.2538512
11 0.8540229 0.8461787 1.0459471 1.2538512

$trend
        Qtr1      Qtr2      Qtr3      Qtr4
1         NA        NA  168616.4  180977.5
2   191946.9  204670.2  218387.2  231396.4
3   244122.5  257864.9  269291.2  278899.1
4   290863.0  305014.6  321596.4  337869.9
5   355927.4  381466.2  408204.4  429641.9
6   451535.1  473495.6  494436.5  520185.5
7   559988.2  607378.4  642043.8  665746.5
8   692775.0  724261.8  746941.8  763666.0
9   782482.6  807638.4  837373.5  860877.1
10  882078.1  912036.8  942914.0  966964.1
11 1000343.8 1060550.5        NA        NA
```

```
$random
        Qtr1      Qtr2      Qtr3      Qtr4
1         NA        NA 1.0300809 1.0568227
2  0.9759220 0.9513383 0.9841455 1.0287188
3  1.0123451 0.9979525 0.9714255 1.0282930
4  0.9715959 1.0241426 0.9595068 1.0000719
5  1.0195903 1.0097599 0.9499738 1.0119249
6  1.0434277 1.0108173 0.9700971 0.9578231
7  1.0252114 1.0177164 1.0454299 0.9625491
8  0.9952983 1.0020545 1.0588285 0.9714738
9  0.9630710 1.0143189 1.0260646 0.9820163
10 0.9978164 1.0017945 1.0021372 0.9979965
11 0.9934117 0.9677950        NA        NA

$figure
[1] 0.8540229 0.8461787 1.0459471 1.2538512

$type
[1] "multiplicative"

attr(,"class")
[1] "decomposed.ts"
```



1200000
800000
400000
Sales
0
10
20
30
40
Time



Decomposition of multiplicative time series
observed
200000
800000
1400000
trend
2e+05
6e+05
1e+06
seasonal
0.9
1.0
1.1
1.2
random
0.96
1.00
1.04
2
4
6
8
10
12
Time