A. Cable Wire Sales

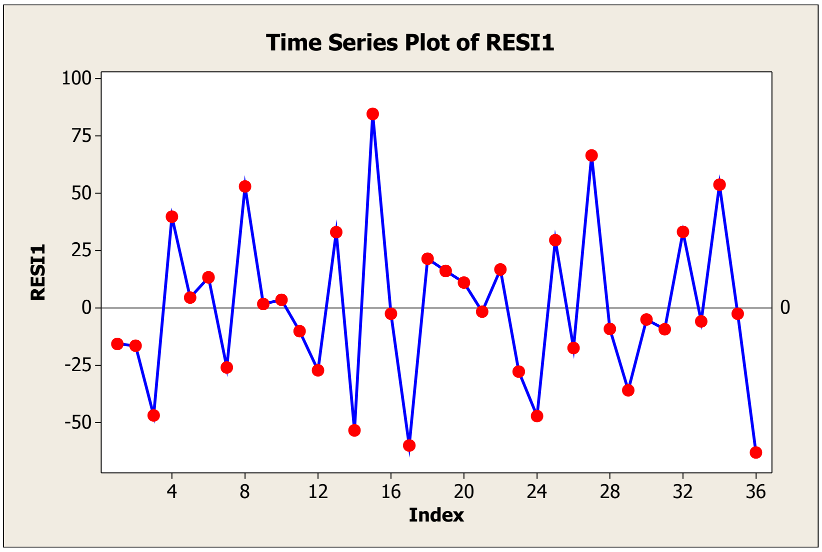
(a) Yes

(b)

- Observation
- 65.1%

(c) Cable sales increase by about 4350 feet each month.

(d)



- H_0: Residuals are random
  H_A: Residuals are autocorrelated
- P-value= 0.447
- The residuals are not autocorrelated. (They are random.)

(e) We are 95% certain that cable wire sales will be between 247,200 and 400,740 feet in March 2008, if present trends continue.

(f) No! Since regression residuals are random (not autocorrelated) we can use regression to predict sales in Observation 39.

B. Trade Employment

(a) Dec. 1974 = Month 60 $\Longrightarrow$ March 1976 = Month 75 $\Longrightarrow$ $\widehat{y} = \boxed{401{,}350}$

(b) Stop! Autocorrelation in the trend-only model renders prediction intervals unreliable. Do not use the interval (383.61, 419.10).

(c) 387,033

(d) Stop! Autocorrelation in the trend-and-seasonal model renders prediction intervals unreliable.

(e) $\widehat{\gamma}_1 = -19.030 \Longrightarrow \boxed{19{,}030}$

(f) 9,071

(g)

- From May to June each year there is <u>one additional month of trend</u> = <u>1.08819</u>
- We must also consider <u>seasonal effects</u>, after accounting for trend:

$$\text{June Effect} \; - \; \text{May Effect} \; = -9.071 - (-11.583) = \underline{2.512}$$

- Taken together, the average increase is

$$1.08819 + 2.512 = 3.60019 \approx 3.600 \Longrightarrow \boxed{\text{3600 employees}}$$

(h) $p = \boxed{\text{2 months}}$

(i) Dec. 1975 = 60 + 12 = 72

- Simple forecast:

$$F_t = \widehat{y}_t + \widehat{e}_t$$
$$F_{72} = \widehat{y}_{72} + \widehat{e}_{72} = 406.775 + 0.2034 = \boxed{406.9784}$$

- 95% interval forecast:

$$(406.775 - 9.6185\,,\; 406.775 + 10.0253) = (397.1565\,,\; 416.8003)$$

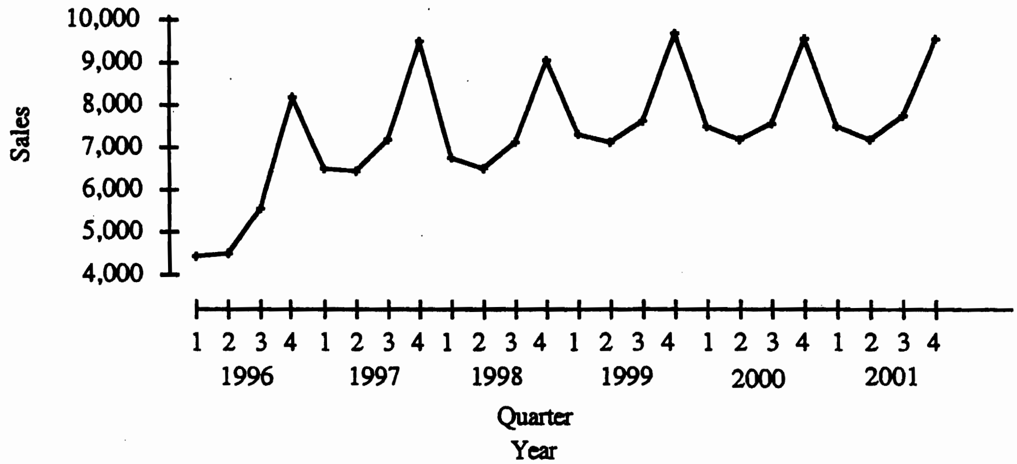
(j)

- Simple forecast = $\boxed{387.1576}$
- 95% forecast: (377.2126, 397.1026)

C. JCPenny Retail Sales

13.1 a) Each year, sales are lowest for the first two quarters, and then increase in the third and fourth quarters. Sales decrease from the fourth quarter of one year to the first quarter of the next. This pattern is repeated year after year.

b) A time plot of these data is given below.



c) There appears to be a positive trend, although the trend levels off after 1998.

d) There is an obvious repeating pattern. Each year, sales in the first two quarters are low, increase in the third quarter, and then increase by an even larger amount in the fourth quarter. Sales then drop substantially in the first quarter of the following year.

13.3 a) Using statistical software, we get the following.

Regression Analysis: Sales versus Order

```
The regression equation is
Sales = 5903 + 119 Order


Predictor      Coef  SE Coef       T      P
Constant     5903.2    492.9   11.98  0.000
Order        118.75    34.49    3.44  0.002


S = 1169.71    R-Sq = 35.0%    R-Sq(adj) = 32.1%
```

We see that the equation of the least-squares line is

$$\text{Sales} = 5903.20 + 118.75x$$

where sales are in millions of dollars and x takes on values 1, 2, 3,..., 24 as described in the statement of the problem.

b) The intercept corresponds to $x = 0$. $x = 1$ represents the first quarter of 1996, so $x = 0$ is the quarter preceding the first quarter of 1996. This means that $x = 0$ represents the fourth quarter of 1995.

c) The slope represents the increase in the response corresponding to a unit increase in the predictor variable x. In this case the response is sales, in millions of dollars, and a unit change in the predictor x corresponds to a change in time of one quarter. Thus, the slope is the increase in sales (in millions of dollars) that occurs from one quarter to the next. In particular, the least-squares regression line predicts a increase in sales of \$118.75 million each quarter, on average.

13.5 a) Using statistical software to estimate the trend-and-season model, we get the following results.

Regression Analysis: Sales versus Order, Q1, Q2, Q3

```
The regression equation is
Sales = 7859 + 99.5 Order - 2274 Q1 - 2565 Q2 - 2023 Q3


Predictor      Coef  SE Coef      T      P
Constant     7858.8    331.3  23.72  0.000
Order         99.54    16.93   5.88  0.000
Q1          -2274.2    331.1  -6.87  0.000
Q2          -2564.6    328.9  -7.80  0.000
Q3          -2022.8    327.6  -6.17  0.000


S = 566.720   R-Sq = 86.8%   R-Sq(adj) = 84.1%
```

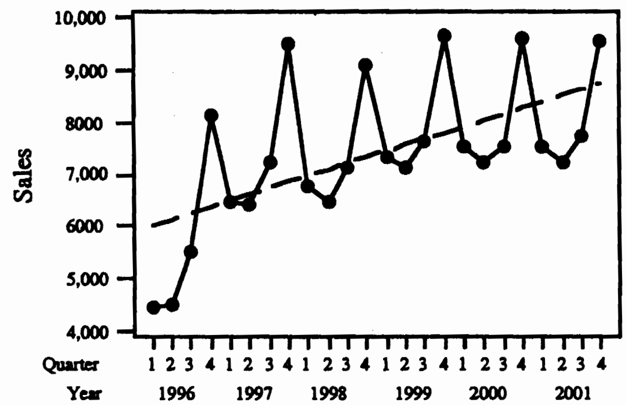
$$\text{Sales} = 7858.80 + 99.54x - 2274.2X1 - 2564.6X2 - 2022.8X3$$

b) If we know that $X1 = X2 = X3 = 0$, then we know that we are not in any of the first three quarters, and hence must be in the fourth quarter. Thus, another indicator variable, $X4$, to tell us if we are in the fourth quarter is not needed because we can tell whether we are in the fourth quarter from the values of $X1$, $X2$, and $X3$.

c) The intercept corresponds to $x = X1 = X2 = X3 = 0$. $x = 0$ is the quarter before the first quarter of 1996, namely the fourth quarter of 1995. $X1 = X2 = X3 = 0$ means we make no seasonal adjustments for being in the first, second, or third quarter. Thus, the intercept again represents the fourth quarter of 1995.

The values of fourth-quarter sales in 1996, 1997, 1998, 1999, 2000, and 2002 are all above 8000, and given the clear pattern of seasonal variation the estimate of the intercept in part (a) appears to be a better estimate than that of Exercise 13.3(b) for fourth-quarter sales in 1995.

13.15 a) The dashed line in the time plot on the next page corresponds to the least squares line. Using the trend-only model, the sales for the first three quarters tend to be overpredicted and the sales for the fourth quarter are underpredicted. Generally, the overpredictions in the first three quarters are of smaller magnitude than the underpredictions in the fourth quarter.



b) The equation of the least-squares line is

$$\text{Sales} = 5903.20 + 118.75x$$

where sales are in millions of dollars and x takes on values 1, 2, 3,...,24. The first quarter of 2002 corresponds to $x = 25$, and the fourth quarter of 2002 corresponds to $x = 28$. The predictions are

$x = 25$ Predicted Sales = 5903.20 + 118.75(25) = 8871.95 million dollars
$x = 28$ Predicted Sales = 5903.20 + 118.75(28) = 9228.20 million dollars

c) Based on previous history, we expect the first-quarter sales to be overpredicted and the fourth quarter to be underpredicted. Which forecast will be more accurate is difficult to guess, although in the past, predictions in the first three quarters tended to be slightly more accurate than in the fourth quarter.

13.16

(a) First-quarter sales for 2002 are predicted to be \$8073.10 (MINITAB's prediction using Options is \$8073) and fourth-quarter sales for 2002 are predicted to be \$10,645.92 (MINITAB's prediction is \$10,646.)

(b) The trend-seasonal model gives a lower forecast for the first quarter and a higher forecast for the fourth quarter than the trend-only model. Since the trend model tends to overestimate first quarters and underestimate fourth quarters, the trend-seasonal forecasts are likely to be more realistic and accurate.

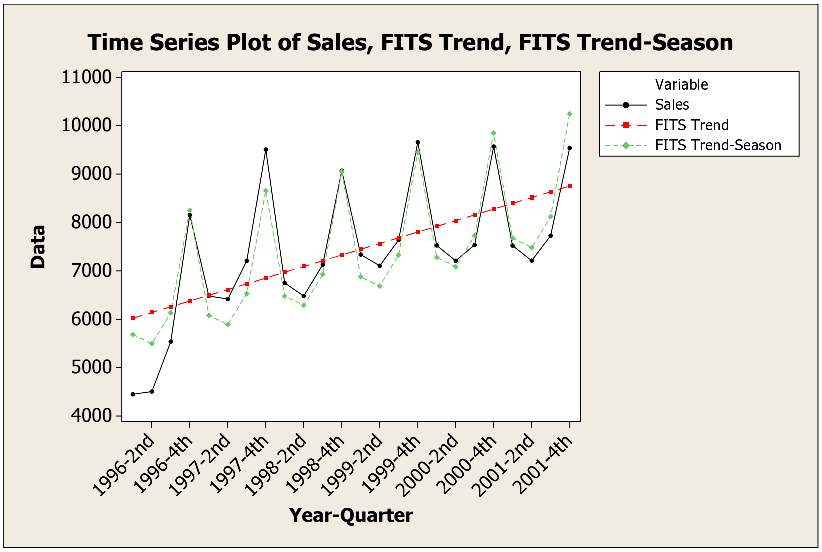
C. JCPenny Retail Sales, continued

- 13.19

(a) Trend-only: $R^2 = 35.0\%$. Trend-and-seasonal: $R^2 = 86.8\%$. R^2 is much higher for the trend-and-seasonal model.

(b) Trend-only: $s = 1169.71$. Trend-and-seasonal: $s = 566.720$. s is much smaller for the trend-and-seasonal model.

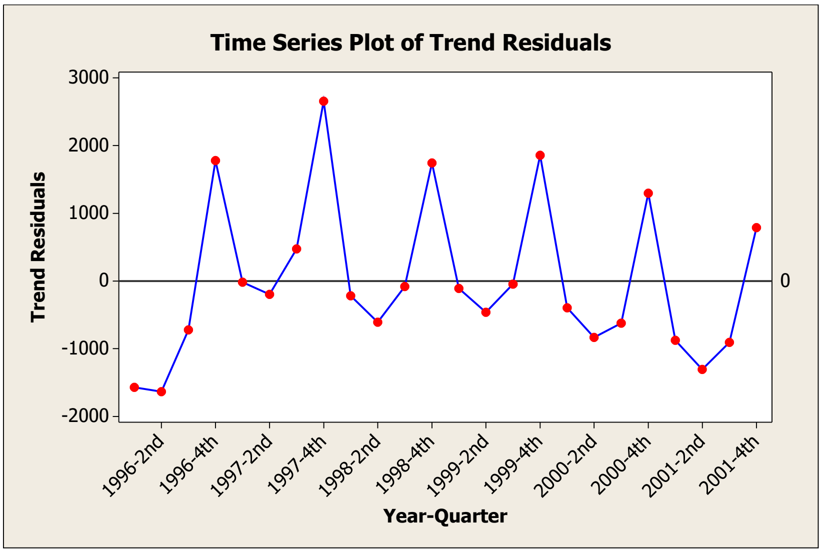
(c)



(d) The trend-and-seasonal model is a big improvement over the trend-only model.

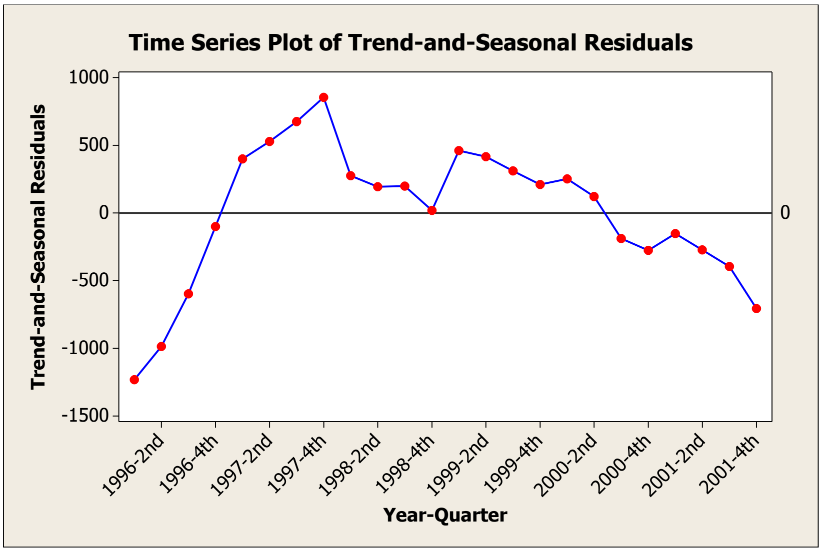
- 13.23

(a) The residuals plot from the trend-only model shows seasonality.



• 13.24

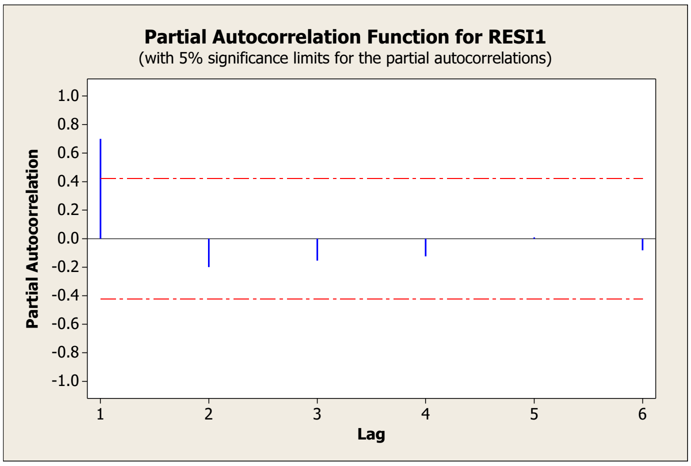
(a) The residuals plot from the trend-and-seasonal model shows autocorrelation.



(continued)

- **Extra Exercise**

(a) From MINITAB Partial Autocorrelation:



Try $p = 1$ with MINITAB ARIMA:

```
Final Estimates of Parameters

Type       Coef  SE Coef      T      P
AR    1  0.9965   0.0997  10.00  0.000
```

Try $p = 2$:

```
Final Estimates of Parameters

Type        Coef  SE Coef      T      P
AR    1   1.3964   0.2002   6.98  0.000
AR    2  -0.5109   0.1876  -2.72  0.012
```

Try $p = 3$:

```
Final Estimates of Parameters

Type        Coef  SE Coef      T      P
AR    1   1.2813   0.2235   5.73  0.000
AR    2  -0.1784   0.3362  -0.53  0.601
AR    3  -0.3180   0.2070  -1.54  0.139
```

So $p = \boxed{\text{2 quarters}}$ is the correct value of p since the AR(3) model has two non-significant time periods (P-value $= 0.601$ and P-value $= 0.139$.)

(b) The residuals model is

$$\widehat{e}_t = (1.3964)\,\widehat{e}_{t-1} - (0.5109)\,\widehat{e}_{t-2}$$

(see remaining answers next page)

(c) First quarter of 2002:

The simple forecast is $F_t = \widehat{y}_t + \widehat{e}_t$

$$F_{25} = \widehat{y}_{25} + \widehat{e}_{25} = 8073 - 782.97 = \boxed{7290.03 \text{ million dollars}}$$

The 95% interval forecast is

$$(8073 - 1240.58,\ 8073 - 325.35) = \boxed{(6832.42,\ 7747.65) \text{ million dollars}}$$

(d) 10,150.29 million dollars

(8965.64, 11,334.73) million dollars

(e) 8737.62 million dollars

(7369.07, 10,106.18) million dollars

(f) The residuals plot from (Trend + Seasonal) regression shows negative residuals for the last 6 quarters of data. So it makes sense that residuals forecast into the future (beyond the 4th quarter of 2001) are also negative.

(g)

It is harder, as evidenced by a larger width in the 95% forecast interval:

- width for 4th quarter, 2002 = 2369.09
- width for 3rd quarter, 2003 = 2737.11

(continued)

C. JCPenny Retail Sales

13.7 a) In Exercise 13.3, we computed the trend to be

$$\text{Sales} = 5903.2 + 118.75x$$

For each quarter, we compute the value of this trend and then divide the actual sales by the trend. The results are summarized below. (Note that we have rounded off the estimates of the slope and intercept in our trend. If you did not round off, your results may differ slightly from ours.)

x	Sales	Trend	Sales/Trend
1	4452	6021.9700	0.739
2	4507	6140.7200	0.734
3	5537	6259.4700	0.885
4	8157	6378.2200	1.279
5	6481	6496.9700	0.998
6	6420	6615.7200	0.970
7	7208	6734.4700	1.070
8	9509	6853.2200	1.388
9	6755	6971.9700	0.969
10	6483	7090.7200	0.914
11	7129	7209.4700	0.989
12	9072	7328.2200	1.238
13	7339	7446.9700	0.986
14	7104	7565.7200	0.939
15	7639	7684.4700	0.994
16	9661	7803.2200	1.238
17	7528	7921.9700	0.950
18	7207	8040.7200	0.896
19	7538	8159.4700	0.924
20	9573	8278.2200	1.156
21	7522	8396.9700	0.896
22	7211	8515.7200	0.847
23	7729	8634.4700	0.895
24	9542	8753.2200	1.090

1st quarter ratios

To get the seasonality factors, we average together the values of the last column corresponding to each of the four quarters. For example, the seasonality factor for the first quarter is the average of the entries in the last column corresponding to x = 1, 5, 9, 13, 17, and 21, i.e., the average of 0.739, 0.998, 0.969, 0.986, 0.950, and 0.896. We summarize the results below.

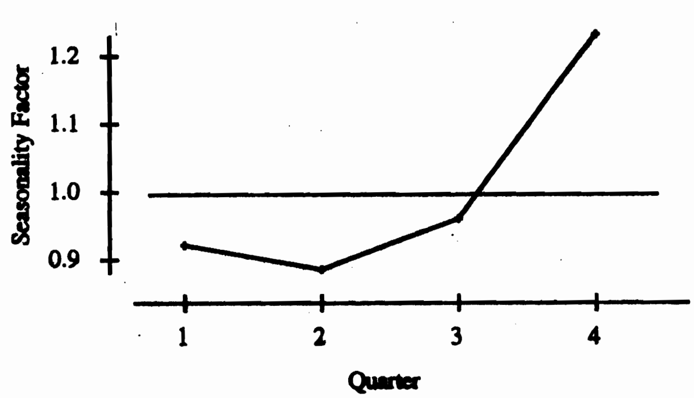
Quarter	Seasonality Factor
1	0.923
2	0.885
3	0.960
4	1.231

b) The average of the four seasonality factors is

$$(0.923 + 0.885 + 0.960 + 1.231)/4 = 3.999/4 = 0.999$$

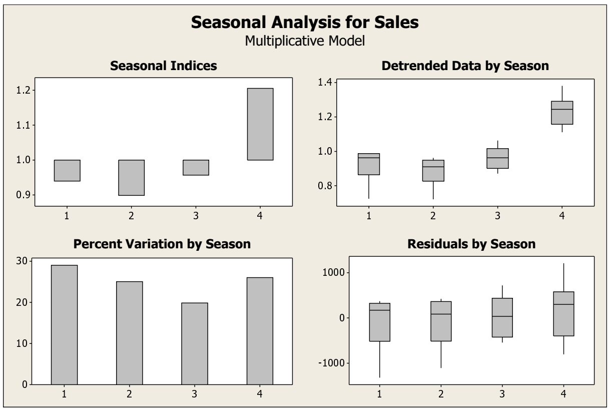
and this is close to 1. The fourth-quarter seasonality factor of 1.231 tells us that fourth-quarter sales are typically 23.1% above the average for all four quarters.

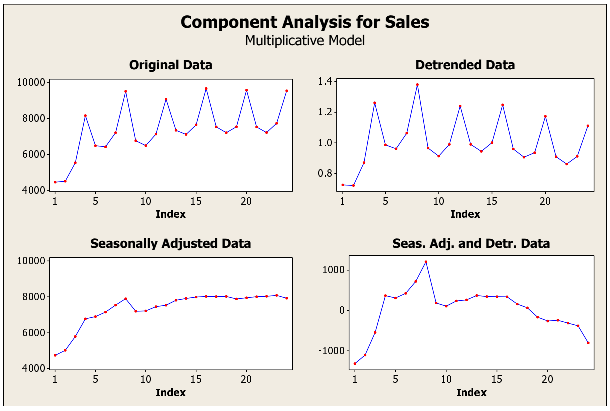
c) The plot is given below. Notice that it mimics the pattern of the seasonal variation seen in Exercise 13.1.

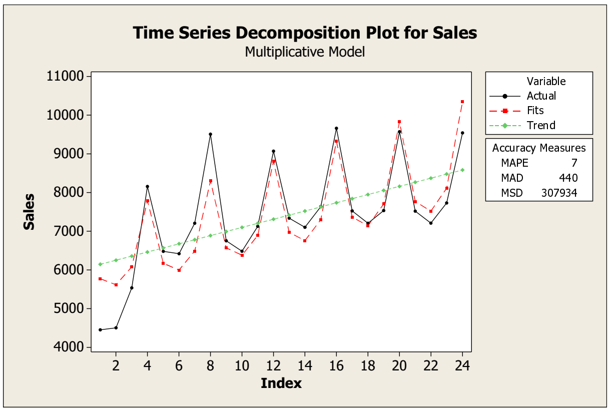


- 13.7

  (d)







Below are MINITAB's seasonal factors from the Decomposition procedure. They differ from the seasonal factors which you calculated by hand in part (a).

```
SeasonalIndices

Period    Index
     1  0.93963
     2  0.89853
     3  0.95668
     4  1.20516
```

MINITAB's seasonally-adjusted time series is one of the plots above. Seasonally-adjusted retail sales in the second quarter of 1998 are $6483/0.89853 = \boxed{\$7215.12 \text{ million}}$. This adjustment represents an <u>increase</u> from actual sales of $6483 million.

13.17 a) For the linear-only model and the seasonality factors, the predictions are

$$\text{Predicted Sales} = (5903.2 + 118.75x) \times \text{SF}$$

where the seasonality factors were calculated in Exercise 13.7 as

Quarter	Seasonality Factor
1	0.923
2	0.885
3	0.960
4	1.231

The first quarter of 2002 corresponds to $x = 25$ and the fourth quarter of 2002 corresponds to $x = 28$. The predictions are

$x = 25$ Predicted Sales = $(5903.2 + 118.75[25]) \times 0.923 =$ 8188.81 million dollars
$x = 28$ Predicted Sales = $(5903.2 + 118.75[28]) \times 1.231 =$ 11359.91 million dollars

b) The first-quarter forecast has been multiplied by 0.923, as the trend-only model typically overpredicts the first quarter while the fourth-quarter forecast has been multiplied by 1.231 to account for the fact that the trend-only model typically underpredicts the fourth quarter. The seasonality factors are simple adjustments to the trend-only model predictions that are designed to account for the seasonality in the time series.

c) The estimated trend-and-season model from Exercise 13.5 is

$$\text{Sales} = 7858.8 + 99.54x - 2274.2\,X1 - 2564.6\,X2 - 2022.8\,X3$$

For the first quarter, we have $X1 = 1$ and $X2 = X3 = 0$, while for the fourth quarter we have $X1 = X2 = X3 = 0$. Thus, the first quarter of 2002 corresponds to $x = 25$, $X1 = 1$, and $X2 = X3 = 0$, and the fourth quarter of 2002 corresponds to $x = 28$ and $X1 = X2 = X3 = 0$. The predictions are

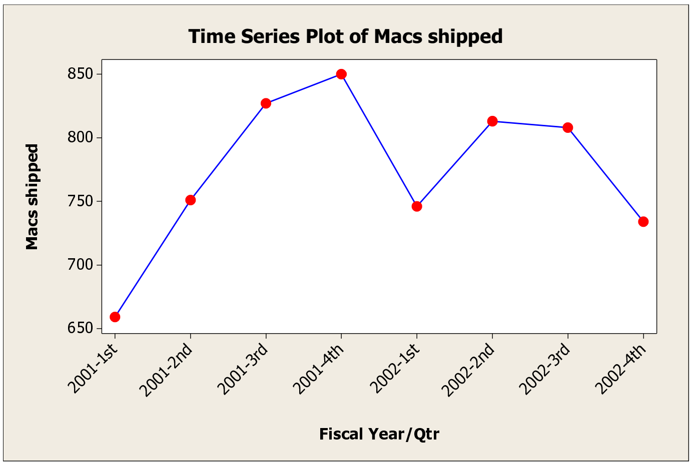
First quarter, 2002 Predicted Sales = $(7858.8 + 99.54[25]) - 2274.2 =$ 8073.1 million dollars
Fourth quarter, 2002 Predicted Sales = $(7858.8 + 99.54[28]) =$ 10645.92 million dollars

The trend-and-season model and the trend-only model with seasonality factors give similar predictions as both are adjusting the trend model for the seasonality effects.

D. Number of Macs Shipped

- 13.2

(a)



(b) The simple regression formulas

$$b_1 = r\,\frac{s_y}{s_x} \qquad\qquad b_0 = \bar{y} - b_1\bar{x}$$

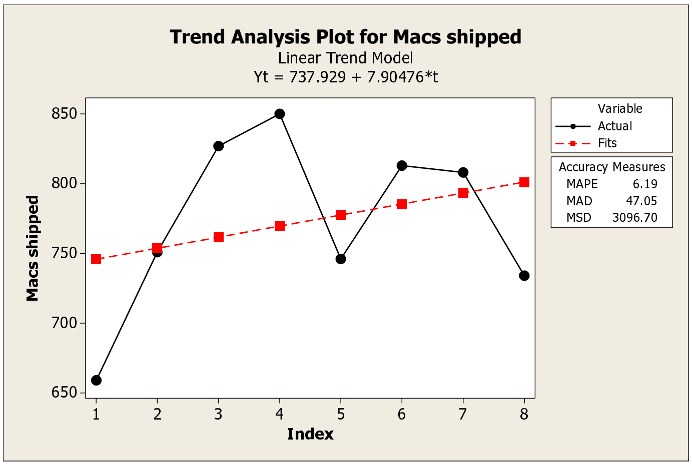
imply that $b_1 = \boxed{7.905}$ and $b_0 = \boxed{737.93}$ Therefore the regression line is

$$\widehat{y} = 737.93 + 7.905\,x$$

or in common English:

$$\text{Mac Sales (thousands)} = 737.93 + 7.905 \times \text{Quarter}$$

(c)



The linear model (regression line) does not seem to fit the data well.

(d) The trend line from the MINITAB Trend Analysis and the regression line which you calculated in part (b) are identical (except for roundoff error.) This shows that MINITAB is using simple linear regression to calculate the trend line in its Trend Analysis procedure. (If you wish, apply MINITAB Regression to the data for further confirmation.)

- 13.4

  (a)

$$\widehat{y} = 786.75 + (0.87)\text{Quarter} - (86.88)X_1 - (8.25)X_2 + (26.38)X_3$$

  (b) The fourth quarter is represented by

$$X_1 = X_2 = X_3 = 0$$

  (c) The F test has P-value $= 0.564$. The test indicates that this model is not a good predictor of Macs shipped.

- 13.6

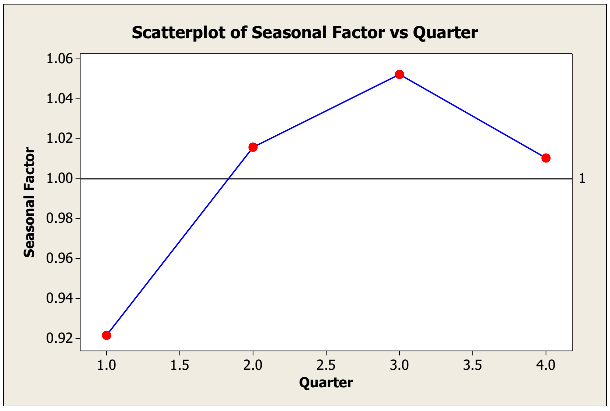
  (a)

Quarter	Seasonal Factor
1	0.9216
2	1.0158
3	1.0522
4	1.0104

  (b) The average of the four factors is $4.0/1.0 = 1.0$.

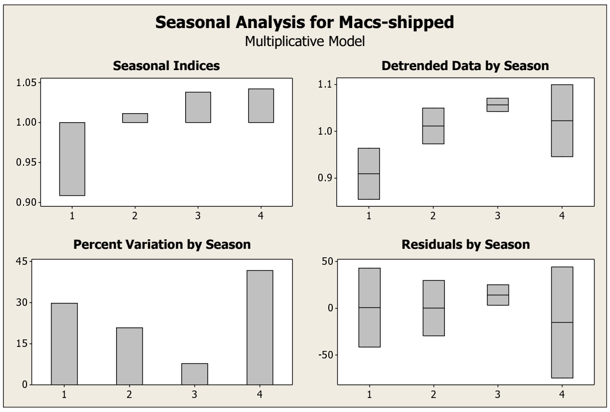
  The factor for the first quarter is 0.9216, which means that the seasonal forecast for number of Macs shipped during first quarters is approximately 8% below the amount forecast by the trend model alone.

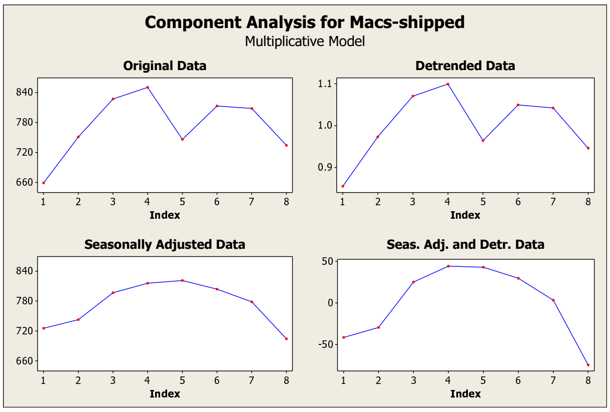
  (c)

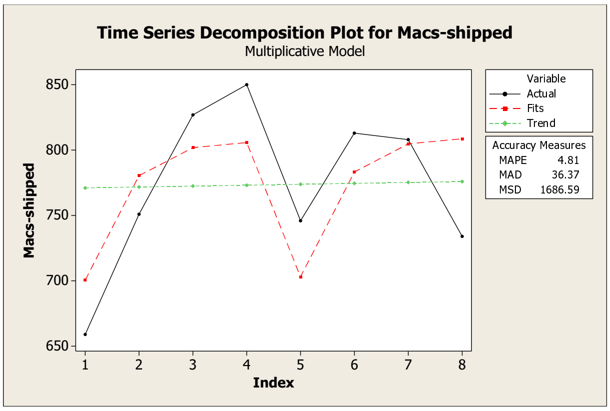


• 13.6

(d)







Below are MINITAB's seasonal factors from the Decomposition procedure. They differ from the seasonal factors which you calculated by hand in part (a).

```
SeasonalIndices

Period    Index
     1  0.90853
     2  1.01128
     3  1.03806
     4  1.04213
```

MINITAB's seasonally-adjusted time series is one of the plots above. Seasonally-adjusted Mac sales in the first quarter of 2002 are 821,107 units . This adjustment represents an <u>increase</u> from actual sales of 746,000 units.