## Tutorial letter 202/2/2018

Forecasting

## STA2604

Semester 2

## Department of Statistics

Solutions to Assignment 2

## STA2604 <br> Forecasting II <br> Solutions to Assignment 02

## QUESTION 1

The multiplicative decomposition model is:

$$
y_{t}=T R_{t} \times S N_{t} \times C L_{t} \times I R_{t}
$$

where $T R_{t}, S N_{t}, C L_{t}$ and $I R_{T}$ are the trend, seasonal, cyclical and irregular factors, respectively.
(1.1) The plot of sales versus time is given by the following graph:


This graph indicates that the trend for the data appears to be increasing.
(1.2) The four-point moving averages are given in the fifth column of Table 1.

Table 1: Analysis of sales data using multiplicative decomposition of a time series

| Yr | Quart | $t$ | $y_{t}$ | 4 MA | $C M A_{t}$ | $s n_{t} \times i r_{t}$ | $s n_{t}$ | $d_{t}$ | $t r_{t}$ | $\widehat{y}_{t}$ | $c l_{t} \times \mathrm{irt}_{t}$ | $c l_{t}$ | $i r_{t}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 10 |  |  |  | 0.47 | 21.28 | 23.13 | 10.87 | 0.92 |  |  |
| 1 | 2 | 2 | 29 |  |  |  | 1.21 | 23.97 | 23.91 | 28.93 | 1.00 | 0.98 | 1.03 |
| 1 | 3 | 3 | 40 | 24.75 | 25 | 1.6000 | 1.61 | 24.84 | 24.69 | 39.75 | 1.01 | 1.03 | 0.97 |
| 1 | 4 | 4 | 20 | 25.25 | 25.75 | 0.7767 | 0.72 | 27.78 | 25.47 | 18.34 | 1.09 | 1.02 | 1.07 |
| 2 | 1 | 5 | 12 | 26.25 | 26.875 | 0.4465 | 0.47 | 25.53 | 26.25 | 12.34 | 0.97 | 1.02 | 0.95 |
| 2 | 2 | 6 | 33 | 27.5 | 27.375 | 1.2055 | 1.21 | 27.27 | 27.03 | 32.70 | 1.01 | 1.00 | 1.01 |
| 2 | 3 | 7 | 45 | 27.25 | 27.625 | 1.6290 | 1.61 | 27.95 | 27.81 | 44.77 | 1.01 | 0.98 | 1.03 |
| 2 | 4 | 8 | 19 | 28 | 28.625 | 0.6638 | 0.72 | 26.39 | 28.59 | 20.58 | 0.92 | 1.01 | 0.92 |
| 3 | 1 | 9 | 15 | 29.25 | 30.125 | 0.4979 | 0.47 | 31.91 | 29.37 | 13.80 | 1.09 | 1.02 | 1.07 |
| 3 | 2 | 10 | 38 | 31 | 31.25 | 1.2160 | 1.21 | 31.40 | 30.15 | 36.48 | 1.04 | 1.06 | 0.98 |
| 3 | 3 | 11 | 52 | 31.5 |  |  | 1.61 | 32.30 | 30.93 | 49.79 | 1.04 | 1.00 | 1.04 |
| 3 | 4 | 12 | 21 |  |  |  | 0.72 | 29.17 | 31.71 | 22.83 | 0.92 |  |  |

(1.3) The centered moving averages, $C M A_{t}=t r_{t} \times c l_{t}$, are given in the sixth column of Table 1.
(1.4) The $s n_{t} \times i r_{t}=y_{t} / C M A_{t}$ values are given in the seventh column of Table 1.
(1.5) Steps for the calculation of seasonal indices are given in the following tables:

| Year | Quarter |  |  |  |
| :--- | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 |
| 1 | - | - | 1.600 | 0.7767 |
| 2 | 0.4465 | 1.2055 | 1.629 | 0.6638 |
| 3 | 0.4979 | 1.216 | - | - |
| Total | 0.9444 | 2.4215 | 3.229 | 1.4405 |
| Mean (unadjusted seasonal indices) | 0.4722 | 1.2108 | 1.6145 | 0.7203 |
| Then |  |  |  |  |

The total index is: $0.4722+1.2108+1.6145+0.7203=4.0178$.
Adjusted seasonal indices (must sum to 4, the number of seasons; quarters here):

| Quarter | Value |
| :---: | :---: |
| 1 | $\frac{0.4722 \times 4}{4.0178}=0.4701$ |
| 2 | $\frac{1.208 \times 4}{4.0177}=1.2054$ |
| 3 | $\frac{1.1519 \times 4}{4.0178}=1.6073$ |
| 4 | $\frac{0.7203 \times 4}{4.0178}=0.7171$ |

These adjusted seasonal indices are copied in the eighth column of Table 1
(1.6) The deseasonalised observations, $d_{t}=y_{t} / s n_{t}$, are given in the ninth column of Table 1.
(1.7) The plots of these deseasonalised observations versus time is given in the following figure. The figure indicates an overall increasing trend.

Plot of d_t versust

(1.8) This can easily be done using Excel, but must be done by hand in an exam.

| Year | Quarter | $t$ | $d_{t}$ | $t d_{t}$ | $t^{2}$ | $(t-\bar{t})^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 21 | 21 | 1 | 30.25 |
| 1 | 2 | 2 | 24 | 48 | 4 | 20.25 |
| 1 | 3 | 3 | 25 | 75 | 9 | 12.25 |
| 1 | 4 | 4 | 28 | 112 | 16 | 6.25 |
| 2 | 1 | 5 | 26 | 130 | 25 | 2.25 |
| 2 | 2 | 6 | 27 | 162 | 36 | 0.25 |
| 2 | 3 | 7 | 28 | 196 | 49 | 0.25 |
| 2 | 4 | 8 | 26 | 208 | 64 | 2.25 |
| 3 | 1 | 9 | 32 | 288 | 81 | 6.25 |
| 3 | 2 | 10 | 31 | 310 | 100 | 12.25 |
| 3 | 3 | 11 | 32 | 352 | 121 | 20.25 |
| 3 | 4 | 12 | 29 | 348 | 144 | 30.25 |
| Total |  | 78 | 329 | 2250 | 650 | 143 |
| Mean |  | 6.5 | 27.4167 |  |  |  |

$$
\widehat{\beta}_{1}=\frac{\sum_{t=1}^{12} t d_{t}-\frac{1}{12}\left(\sum_{t=1}^{12} t\right)\left(\sum_{t=1}^{12} d_{t}\right)}{\sum_{t=1}^{12} t^{2}-\frac{\left(\sum_{t=1}^{12} t\right)^{2}}{12}}=\frac{2250-\frac{78 \times 329}{12}}{650-\frac{78^{2}}{12}}=0.7797
$$

and

$$
\widehat{\beta}_{0}=\bar{d}-\widehat{\beta}_{1} \bar{t}=\frac{1}{12} \sum_{t=1}^{12} d_{t}-\frac{1}{12} \widehat{\beta}_{1} \sum_{t=1}^{12} t=\frac{1}{12} \times 329-\frac{1}{12} \times 0.7797 \times 78=22.3486
$$

Hence the estimate of the trend is: $t r_{t}=22.3486+0.7797 t$.
(1.9) The trend values $t r_{t}$ are given in the tenth column of Table 1, then the values $\widehat{y}_{t}=t r_{t} \times s n_{t}$ are given in the eleventh column of Table 1. Finally, the values $c l_{t} \times i r_{t}$ are given in the twelfth column of Table 1.
(1.10) The cyclical factors calculated as the three-point moving averages

$$
c l_{t}=\frac{c l_{t-1} i r_{t-1}+c l_{t} i r_{t}+c l_{t+1} i r_{t+1}}{3}
$$

(see textbook, page 335) are given the thirteenth column of Table 1.
(1.11) The irregular values, $i r_{t}=\frac{c l_{t} \times i r_{t}}{c l_{t}}$, are given in the fourteenth column of Table 1. All the values are close to 1 , and thus this time series seems not having the irregular variation component.
(1.12) All values of $c l_{t}$ are close to 1 , thus we cannot determine any well-defined cycle for these data. Following the result in this part and part (1.11), this time series can only be characterised by trend and seasonal components.
(1.13) At year 4 , the values of $t$ are $13,14,15$ and 16 for the four quarters, respectively. The trend was found to be $t r_{t}=22.3486+0.7797 t$., and the four seasonal indices were found to be $0.4701,1.2054,1.6073$ and 0.7171 . Now using the prediction equation $\widehat{y}_{t}=t r_{t} \times s n_{t}$, we obtain:
$\widehat{y}_{13}=(22.3486+0.7797 \times 13) \times 0.4701=15.2711$
$\widehat{y}_{14}=(22.3486+0.7797 \times 14) \times 1.2054=40.0969$
$\widehat{y}_{15}=(22.3486+0.7797 \times 15) \times 1.6073=54.7191$
$\widehat{y}_{16}=(22.3486+0.7797 \times 16) \times 0.7171=24.9721$
(1.14) Here we first calculate prediction intervals of the trend at the four time points then use the formula $\left[\hat{y} \pm B_{t}(1-\alpha)\right]$ where $B_{t}(1-\alpha)$ is the error band in the prediction interval $\left[\operatorname{tr}_{t} \pm B_{t}(1-\alpha)\right]$ (see textbook on pages 335-336). This time, we fit the trend using Excel to obtain the standard error, that $s=1.8433$, or equivalently $s=\sqrt{M S E}=\sqrt{3.3978}=1.8433$.

| SUMMARY OUTPUT |  |
| :--- | :---: |
| Regression Statistics |  |
| Multiple R | 0.8479 |
| R Square | 0.7190 |
| Adjusted R Square | 0.6909 |
| Standard Error | 1.8433 |
| Observations | 12 |


| ANOVA |  |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | df | SS | MS | F | Significance F |  |
| Regression | 1 | 86.9388 | 86.9388 | 15.8302 | 0.0026 |  |
| Residual | 10 | 33.9779 | 3.3978 |  |  |  |
| Total | 11 | 120.9167 |  |  |  |  |


|  | Coefficients | Standard Error | t Stat | P-value | Lower 95\% | Upper 95\% |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept | 22.3485 | 1.1345 | 19.6994 | $<0.0001$ | 19.8207 | 24.8763 |
| X Variable 1 | 0.7797 | 0.1541 | 5.0583 | 0.0005 | 0.4363 | 1.1232 |

The mean value of $t$ is: $\bar{t}=6.5$.
For time 13, the error band of the trend is:

$$
\begin{aligned}
B_{13}(0.95) & =s \times t_{[0.025]}^{(12-2)} \sqrt{1+\frac{1}{12}+\frac{(13-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =1.84 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(13-6.5)^{2}}{143}} \\
& =4.8137 .
\end{aligned}
$$

Therefore, the prediction interval is: $[15.3 \pm 4.8]=[10.5 ; 20.1]$.
For time 14, the error band of the trend is:

$$
\begin{aligned}
B_{14}(0.95) & =s \times t_{[0.25]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(14-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =1.84 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(14-6.5)^{2}}{143}} \\
& =4.9817 .
\end{aligned}
$$

Therefore, the prediction interval is: $[40.1 \pm 5.0]=[35.1 ; 45.1]$. For time 15, the error band is:

$$
\begin{aligned}
B_{15}(0.95) & =s \times t_{[0.25]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(15-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =1.84 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(15-6.5)^{2}}{143}} \\
& =5.1670 .
\end{aligned}
$$

Therefore, the prediction interval is: $[54.7 \pm 5.2]=[49.5 ; 59.9]$.
For time 16, the error band is:

$$
\begin{aligned}
B_{16}(0.95) & =s \times t_{[0.25]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(16-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =1.84 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(16-6.5)^{2}}{143}} \\
& =5.3678 .
\end{aligned}
$$

Therefore, the prediction interval is: $[25.0 \pm 5.4]=[19.6 ; 29.4]$.

## QUESTION 2

(2.1) The additive decomposition model is

$$
y_{t}=T R_{t}+S N_{t}+C L_{t}+I R_{t}
$$

where $T R_{t}, S N_{t}, C L_{t}$ and $I R_{T}$ are the trend, seasonal, cyclical and irregular terms, respectively. All calculations will be done as in Question 1 by replacing the sign $\times$ by the sign + , and the sign / by the sign -.

Table 2: Analysis of sales data using additive decomposition of a time series

| Yr | Quart | $t$ | $y_{t}$ | 4MA | $C M A_{t}$ | $s n_{t}+i r_{t}$ | $s n_{t}$ | $d_{t}$ | $t r_{t}$ | $\widehat{y}_{t}$ | $c l_{t}+i r t_{t}$ | $c l_{t}$ | $i r_{t}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 10 |  |  |  | -14.92 | 24.92 | 23.68 | 8.76 | 1.24 |  |  |
| 1 | 2 | 2 | 29 |  |  |  | 6.27 | 22.73 | 24.48 | 30.75 | -1.75 | -0.69 | -1.06 |
| 1 | 3 | 3 | 40 | 24.75 | 25 | 15 | 16.27 | 23.73 | 25.28 | 41.55 | -1.55 | -0.59 | -0.96 |
| 1 | 4 | 4 | 20 | 25.25 | 25.75 | -5.75 | -7.61 | 27.61 | 26.08 | 18.47 | 1.53 | 0.01 | 1.52 |
| 2 | 1 | 5 | 12 | 26.25 | 26.875 | -14.875 | -14.92 | 26.92 | 26.88 | 11.96 | 0.04 | 0.20 | -0.17 |
| 2 | 2 | 6 | 33 | 27.5 | 27.375 | 5.625 | 6.27 | 26.73 | 27.68 | 33.95 | -0.95 | -0.22 | -0.73 |
| 2 | 3 | 7 | 45 | 27.25 | 27.625 | 17.375 | 16.27 | 28.73 | 28.48 | 44.75 | 0.25 | -1.13 | 1.37 |
| 2 | 4 | 8 | 19 | 28 | 28.625 | -9.625 | -7.61 | 26.61 | 29.28 | 21.67 | -2.67 | -0.86 | -1.81 |
| 3 | 1 | 9 | 15 | 29.25 | 30.125 | -15.125 | -14.92 | 29.92 | 30.09 | 15.17 | -0.17 | -0.67 | 0.50 |
| 3 | 2 | 10 | 38 | 31 | 31.25 | 6.75 | 6.27 | 31.73 | 30.89 | 37.16 | 0.84 | 1.57 | -0.73 |
| 3 | 3 | 11 | 52 | 31.5 |  |  | 16.27 | 35.73 | 31.69 | 47.96 | 4.04 | 0.34 | 3.71 |
| 3 | 4 | 12 | 21 |  |  |  | -7.61 | 28.61 | 32.49 | 24.88 | -3.88 |  |  |

(2.1.1) The plot of sales versus time is the same as in Question 1, part (1.1).
(2.1.2) The four-point moving averages are given in the fifth column of Table 2, but same as those in Table 1.
(2.1.2) The centered moving averages are given in the sixth column of Table 2, but same as those in Table 1.
(2.1.4) The $s n_{t}+i r_{t}=y_{t}-C M A_{t}$ are given in the seventh column of Table 2.
(2.1.5) Steps for the calculation of seasonal indices are given in the following tables:

| Year | Quarter |  |  |  |
| :--- | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 |
| 1 | - | - | 15 | -5.75 |
| 2 | -14.875 | 5.625 | 17.375 | -9.625 |
| 3 | -15.125 | 6.75 | - | - |
| Total | -30 | 12.375 | 32.375 | -15.375 |
| Mean (unadjusted seasonal indices) | -15 | 6.1875 | 16.1875 | -7.6875 |

The total index is: $-15+6.1875+16.1875-7.6875=-0.3125$.
Adjusted seasonal or normalised indices (sum must be equal to zero):

| Quarter | Value |
| :---: | :---: |
| 1 | $-15-\frac{-0.3125}{4}=-14.9219$ |
| 2 | $6.1875-\frac{-0.3125}{4}=6.2656$ |
| 3 | $16.1875-\frac{-0.3125}{4}=16.2656$ |
| 4 | $-7.6875-\frac{-0.3125}{4}=-7.6094$ |

These adjusted seasonal indices are copied in the eighth column of Table 2
(2.1.6) The deseasonalised observations $d_{t}=y_{t}-s n_{t}$ are given in the ninth column of Table 2.
(2.1.7) The plots of these observations versus time is given in the following figure. The figure indicates an overall slightly increasing trend.

Plot of d_t versus t

(2.1.8) This can easily be done using Excel, but must be done by hand in an exam. We need to use a different table here.

| Year | Quarter | $t$ | $d_{t}$ | $t d_{t}$ | $t^{2}$ | $(t-\bar{t})^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 25 | 25 | 1 | 30.25 |
| 1 | 2 | 2 | 23 | 46 | 4 | 20.25 |
| 1 | 3 | 3 | 24 | 72 | 9 | 12.25 |
| 1 | 4 | 4 | 28 | 112 | 16 | 6.25 |
| 2 | 1 | 5 | 27 | 135 | 25 | 2.25 |
| 2 | 2 | 6 | 27 | 162 | 36 | 0.25 |
| 2 | 3 | 7 | 29 | 203 | 49 | 0.25 |
| 2 | 4 | 8 | 27 | 216 | 64 | 2.25 |
| 3 | 1 | 9 | 30 | 270 | 81 | 6.25 |
| 3 | 2 | 10 | 32 | 320 | 100 | 12.25 |
| 3 | 3 | 11 | 36 | 396 | 121 | 20.25 |
| 3 | 4 | 12 | 29 | 348 | 144 | 30.25 |
| Total |  | 78 | 337 | 2305 | 650 | 143 |
| Mean |  | 6.5 | 28.0833 |  |  |  |
| $\widehat{\beta}_{1}=\frac{\sum_{t=1}^{12} t d_{t}-\frac{1}{12}\left(\sum_{t=1}^{12} t\right)\left(\sum_{t=1}^{12} d_{t}\right)}{\sum_{t=1}^{12} t^{2}-\frac{\left(\sum_{t=1}^{12} t\right)^{2}}{12}}=\frac{2305-\frac{78 \times 337}{12}}{650-\frac{77^{2}}{12}}=0.8007$ |  |  |  |  |  |  |

and

$$
\widehat{\beta}_{0}=\bar{d}-\widehat{\beta}_{1} \bar{t}=\frac{1}{12} \sum_{t=1}^{12} d_{t}-\frac{1}{12} \widehat{\beta}_{1} \sum_{t=1}^{12} t=\frac{1}{12} \times 337-\frac{1}{12} \times 0.8007 \times 78=22.8788
$$

Hence the estimate of the trend is: $t r_{t}=22.8788+0.8007 t$.
This equation is slightly different from the one found in Question 1, part (1.8) since the deseasonalised data are based on a different model. The trend values $t r_{t}$ are given in the tenth column of Table 2.
(2.1.9) The point forecast values $\widehat{y}_{t}=t r_{t}+s n_{t}$ are given in the eleventh column of Table 2. Finally, the values $c l_{t}+i r_{t}$ are given in the twelfth column of Table 2.
(2.1.10) The cyclical terms calculated as the three-point moving averages

$$
c l_{t}=\frac{\left(c l_{t-1}+i r_{t-1}\right)+\left(c l_{t}+i r_{t}\right)+\left(c l_{t+1}+i r_{t+1}\right)}{3}
$$

(see textbook, page 339) are given the thirteenth column of Table 2.
(2.1.11) The irregular values $i r_{t}=\left(c l_{t}+i r_{t}\right)-c l_{t}$ are given in the fourteenth column of Table 2. Notice that some, but a few, irregular values are not close to zero.
(2.1.12) There is no much evidence to state that the data contain a well defined cycle since the $c l_{t}$ values do not largely vary.
(2.1.13) At year 4, the values of $t$ are $13,14,15$ and 16 for the four quarters, respectively. The trend was found to be $t r_{t}=22.8788+0.8007 t$, and the four seasonal indices were found to be $-14.9219,6.2656,16.2656$ and -7.6094 . Now using the prediction equation $\widehat{y}_{t}=$ $t r_{t}+s n_{t}$, we obtain:
$\widehat{y}_{13}=(22.8788+0.8007 \times 13)-14.9219=18.366$
$\widehat{y}_{14}=(22.8788+0.8007 \times 14)+6.2656=40.3542$
$\widehat{y}_{15}=(22.8788+0.8007 \times 15)+16.2656=51.1549$
$\widehat{y}_{16}=(22.8788+0.8007 \times 16)-7.6094=28.0806$
(2.1.14) The prediction intervals are given by the formula $\left[\widehat{y} \pm B_{t}(1-\alpha)\right]$ where $B_{t}(1-\alpha)$ is the error band in the prediction interval $\left[\operatorname{tr}_{t} \pm B_{t}(1-\alpha)\right]$ (see textbook on pages 335-336). This time, we fit the trend using Excel to obtain the standard error. We only report the relevant output here.

| SUMMARY OUTPUT |  |
| :--- | :---: |
| Regression Statistics |  |
| Multiple R | 0.8124 |
| R Square | 0.6599 |
| Adjusted R Square | 0.6260 |
| Standard Error | 2.1734 |
| Observations | 12 |

The mean value of $t$ is: $\bar{t}=6.5$.
For time 13, the error band is:

$$
\begin{aligned}
B_{13}(0.95) & =s \times t_{[0.025]}^{(12-2)} \sqrt{1+\frac{1}{12}+\frac{(13-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =2.2 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(13-6.5)^{2}}{143}} \\
& =5.7555 .
\end{aligned}
$$

Therefore, the prediction interval is: $[18.4 \pm 5.8]=[12.6 ; 24.2]$.

For time 14, the error band of the trend is:

$$
\begin{aligned}
B_{14}(0.95) & =s \times t_{[0.025]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(14-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =2.2 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(14-6.5)^{2}}{143}} \\
& =5.9564 .
\end{aligned}
$$

Therefore, the prediction interval is: $[40.4 \pm 6.0]=[34.4 ; 46.4]$.
For time 15, the error band is:

$$
\begin{aligned}
B_{15}(0.95) & =s \times t_{[0.025]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(15-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =2.2 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(15-6.5)^{2}}{143}} \\
& =6.1779 .
\end{aligned}
$$

Therefore, the prediction interval is: $[51.2 \pm 6.2]=[45.0 ; 57.4]$.
For time 16, the error band is:

$$
\begin{aligned}
B_{16}(0.95) & =s \times t_{[0.025]}^{(10)} \sqrt{1+\frac{1}{12}+\frac{(16-\bar{t})^{2}}{\sum_{t=1}^{12}(t-\bar{t})^{2}}} \\
& =2.2 \times 2.228 \sqrt{1+\frac{1}{12}+\frac{(16-6.5)^{2}}{143}} \\
& =6.4180 .
\end{aligned}
$$

Therefore, the prediction interval is: $[28.1 \pm 6.4]=[21.7 ; 28.1]$.
(2.2) Overall both models perform similarly in term of the detection of the trend and seasonality, and also perform similarly in terms of predicting future values. Therefore, the additive decomposition model does not seem to be more appropriate than the multiplicative decomposition model. Rather, we have noticed that with the multiplicative model there is no evidence of irregular variations while some irregular variations are evident when the additive is used. This may probably be due to lack-of-fit. Moreover, the following table indicates that the ranges of prediction intervals in year 4 are larger (i.e. wider prediction intervals), except the last one for the additive model than for the multiplicative model.

|  | Multiplicative Model |  | Additive Model |  |
| :---: | :---: | :---: | :---: | :---: |
| Time | $95 \%$ Prediction interval | Range | 95\% Prediction interval | Range |
| 13 | $[10.5 ; 20.1]$ | 9.5 | $[12.6 ; 24.2]$ | 11.6 |
| 14 | $[35.1 ; 45.1]$ | 10.0 | $[34.4 ; 46.4]$ | 12.0 |
| 15 | $[49.5 ; 59.9]$ | 10.4 | $[45.0 ; 57.4]$ | 12.4 |
| 16 | $[19.6 ; 29.4]$ | 9.8 | $[21.7 ; 28.1]$ | 7.4 |

## QUESTION 3

(3.1) Here, we don't use deseasonalised data, but the observed data, because we are only asked to fit a simple linear regression model to the data.

| Year | Quarter | $t$ | $y_{t}$ | $t y_{t}$ | $t^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 10 | 10 | 1 |
| 1 | 2 | 2 | 29 | 58 | 4 |
| 1 | 3 | 3 | 40 | 120 | 9 |
| 1 | 4 | 4 | 20 | 80 | 16 |
| 2 | 1 | 5 | 12 | 60 | 25 |
| 2 | 2 | 6 | 33 | 198 | 36 |
| 2 | 3 | 7 | 45 | 315 | 49 |
| 2 | 4 | 8 | 19 | 152 | 64 |
| 3 | 1 | 9 | 15 | 135 | 81 |
| 3 | 2 | 10 | 38 | 380 | 100 |
| 3 | 3 | 11 | 52 | 572 | 121 |
| 3 | 4 | 12 | 21 | 252 | 144 |
| Total |  | 78 | 334 | 2332 | 650 |

$$
\begin{gathered}
\widehat{\beta}_{1}=\frac{\sum_{t=1}^{12} t y_{t}-\frac{1}{12}\left(\sum_{t=1}^{12} t\right)\left(\sum_{t=1}^{12} y_{t}\right)}{\sum_{t=1}^{12} t^{2}-\frac{\left(\sum_{t=1}^{12} t\right)^{2}}{12}}=\frac{2332-\frac{78 \times 334}{12}}{650-\frac{78^{2}}{12}}=1.1259 . \\
\widehat{\beta}_{0}=\bar{y}-\widehat{\beta}_{1} \bar{t}=\frac{1}{12} \sum_{t=1}^{12} y_{t}-\frac{1}{12} \widehat{\beta}_{1} \sum_{t=1}^{12} t=\frac{1}{12} \times 334-\frac{1}{12} \times 1.1259 \times 78=20.5150 .
\end{gathered}
$$

Hence the estimate of the trend is: $t r_{t}=20.5150+1.1259 t$.
Excel would give the following output.

|  | Coefficients | Standard Error | t Stat | P-value | Lower 95\% | Upper 95\% |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept | 20.5152 | 8.4755 | 2.4205 | 0.0360 | 1.6307 | 39.3996 |
| X Variable 1 | 1.1259 | 1.1516 | 0.9777 | 0.3513 | -1.4400 | 3.6918 |

(3.2) There is no well defined way of choosing the initial level value $l_{0}$ and $b_{0}$. For a large dataset, a linear model is fit on part of the data, say half, so that $l_{0}=\widehat{\beta}_{0}$ and $b_{0}=\widehat{\beta}_{1}$. But, in this case we have a small dataset, and thus we use the entire dataset [No penalty will be applied if you use half of the data, or data for one year]. In this situation, $l_{0}=\widehat{\beta}_{0}=20.5150$ and $b_{0}=\widehat{\beta}_{1}=$ 1.1259. To determine the smoothed levels we start with $\widehat{y}_{1}(0)=l_{0}+b_{0}=20.5150+1.1259=$ 21.6276. Hence, using $l_{T}=\alpha y_{T}+(1-\alpha)\left[l_{T-1}+b_{T-1}\right]$ and $b_{T}=\gamma\left[l_{T}-l_{T-1}\right]+(1-\gamma) b_{T-1}$ we obtain:
$l_{1}=\alpha y_{1}+(1-\alpha)\left(l_{0}+b_{0}\right)=0.15(10)+0.85(21.6276)=19.8835 ;$
$b_{1}=\gamma\left(l_{1}-l_{0}\right)+(1-\gamma) b_{0}=0.1(19.8835-20.5150)+0.9(1.1259)=0.9502$.
$l_{2}=\alpha y_{2}+(1-\alpha)\left(l_{1}+b_{1}\right)=0.15(29)+0.85(19.8835+0.9502)=22.0586 ;$
$b_{2}=\gamma\left(l_{2}-l_{1}\right)+(1-\gamma) b_{1}=0.1(22.0586-19.8835)+0.9(0.9502)=1.0727$.
This process continues until calculating $l_{12}$ and $b_{12}$. Table 3 , columns 3 and 4 contains all the smoothed levels and growth rates manually calculated using Excel spreadsheet.

Table 3: Holt's trend corrected exponential smoothing for the sales data

| Time $(t)$ | Sales $\left(y_{t}\right)$ | Level $\left(l_{t}\right)$ | Growth $\left(b_{t}\right)$ | FMLP | $e_{t}$ | $e_{t}^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 |  | 20.515 | 1.1259 |  |  |  |
| 1 | 10 | 19.8835 | 0.9502 | 21.6409 | -11.6409 | 135.5106 |
| 2 | 29 | 22.0586 | 1.0727 | 20.8337 | 8.1663 | 66.6891 |
| 3 | 40 | 25.6616 | 1.3257 | 23.1313 | 16.8687 | 284.5542 |
| 4 | 20 | 25.9392 | 1.2209 | 26.9873 | -6.9873 | 48.8218 |
| 5 | 12 | 24.8860 | 0.9935 | 27.1601 | -15.1601 | 229.8271 |
| 6 | 33 | 26.9476 | 1.1003 | 25.8795 | 7.1205 | 50.7012 |
| 7 | 45 | 30.5907 | 1.3546 | 28.0479 | 16.9521 | 287.3745 |
| 8 | 19 | 30.0035 | 1.1604 | 31.9453 | -12.9453 | 167.5797 |
| 9 | 15 | 28.7393 | 0.9179 | 31.1639 | -16.1639 | 261.2703 |
| 10 | 38 | 30.9086 | 1.0431 | 29.6572 | 8.3428 | 69.6022 |
| 11 | 52 | 34.9589 | 1.3438 | 31.9517 | 20.0483 | 401.9345 |
| 12 | 21 | 34.0073 | 1.1143 | 36.3027 | -15.3027 | 234.1738 |
| Total |  |  |  |  |  | 2238.0390 |

(3.3) The growth rates $b_{T}=\gamma\left[l_{T}-l_{T-1}\right]+(1-\gamma) b_{T-1}$ where $l_{T}=\alpha y_{T}+(1-\alpha)\left[l_{T-1}+b_{T-1}\right]$ are given in the fourth column of Table 3.
(3.4) The forecast made last time ( $F M L P=F_{t}=l_{t-1}+b_{t-1}$ ) are given in column 5 of Table 3.
(3.5) The forecast errors $e_{t}=y_{t}-F_{t}$ are given in the sixth column of Table 3.
(3.6) The squared forecast errors are given in the seventh column of Table 3.
(3.7) The mean squared forecast error is MSFE $=\frac{\sum_{i=1}^{12} e_{t}^{2}}{12}=\frac{2238.0390}{12}=186.5033$.

