**Answer 1:** The line graph for the number of new cars registered shows that there have been seasonal changes in the data with overall declining trend. There has been a declining trend from 1990 to 1995 with touching the bottom in 1994 and then maintaining for a while. However, the trend picked up thereafter with touching its peak in about 2000.



**Answer 2:** The results for the linear trend shows that trend has significant impact on predicting the new registration of cars. However, trend explains only about 2% variation the new registration of cars. The coefficient of trend has been positive at 242.53 which means that with every month there has been an increase in the new registration of cars by 242.53 on an average.

The plot of the residuals shows that the mean is close to zero as the residuals hover around zero. However, there seems to be a seasonal component in the data. There might be a non-linear relationship present in the two variables. Furthermore, it also shows the presence of serial correlation in the model.

Residuals:

Min 1Q Median 3Q Max

-361825 -106717 1181 112010 354364

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 846135.24 16536.19 51.169 < 2e-16 \*\*\*

md$trend 242.53 85.56 2.835 0.00487 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 150800 on 332 degrees of freedom

Multiple R-squared: 0.02363, Adjusted R-squared: 0.02069

F-statistic: 8.035 on 1 and 332 DF, p-value: 0.004869



**Answer 3:** The results for the quadratic trend show that both the linear and quadratic terms of trend has significant impact on predicting the new registration of cars. However, the two terms of trend explains about 7% variation the new registration of cars. The coefficient of trend has been positive which means that with every month there has been an increase in the new registration of cars while the quadratic term has been negative, which suggests that after certain period of time the trend is negative that is with every month there has been an decrease in the new registration of cars.

The plot of the residuals shows that the mean is close to zero as the residuals hover around zero. However, there is presence of a seasonal component in the data. Furthermore, it also shows the presence of serial correlation in the model.

Call:

lm(formula = md$PureNumber ~ poly(md$trend, 2), data = md)

Residuals:

Min 1Q Median 3Q Max

-338377 -111325 227 103324 351627

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 886760 8064 109.969 < 2e-16 \*\*\*

poly(md$trend, 2)1 427365 147370 2.900 0.00398 \*\*

poly(md$trend, 2)2 -598223 147370 -4.059 6.15e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 147400 on 331 degrees of freedom

Multiple R-squared: 0.06993, Adjusted R-squared: 0.06431

F-statistic: 12.44 on 2 and 331 DF, p-value: 6.154e-06



The results for the polynomial trend show that both the linear and quadratic terms of trend has significant impact on predicting the new registration of cars while the cubic term is not significant. There is no further improvement in the model as the three terms of trend explains about 7% variation the new registration of cars and there has not been any change in the R-square. The coefficient of trend has been positive which means that with every month there has been an increase in the new registration of cars while the quadratic term has been negative, which suggests that after certain period of time the trend is negative that is with every month there has been an decrease in the new registration of cars.

Call:

lm(formula = md$PureNumber ~ poly(md$trend, 3), data = md)

Residuals:

Min 1Q Median 3Q Max

-338354 -111393 247 103397 351526

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 886760 8076 109.803 < 2e-16 \*\*\*

poly(md$trend, 3)1 427365 147593 2.896 0.00404 \*\*

poly(md$trend, 3)2 -598223 147593 -4.053 6.31e-05 \*\*\*

poly(md$trend, 3)3 1636 147593 0.011 0.99116

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 147600 on 330 degrees of freedom

Multiple R-squared: 0.06993, Adjusted R-squared: 0.06148

F-statistic: 8.271 on 3 and 330 DF, p-value: 2.546e-05

**Answer 4:** The results for the model with quadratic trend and seasonal dummies for the months of March and August show that both the linear and quadratic terms of trend has significant impact on predicting the new registration of cars. Moreover, the dummies for the months of March and August also have significant impact on the new registration of cars.

All the terms included in the model explains about 43% variation in the new registration of cars. The coefficient of trend has been positive which means that with every month there has been an increase in the new registration of cars while the quadratic term has been negative, which suggests that after certain period of time the trend is negative that is with every month there has been an decrease in the new registration of cars. The coefficient of march dummy has been positive which means that the new registration of cars is more on an average if the month is March in comparison to if the month is not March while the August dummy has a negative coefficient, which suggests that the new registration of cars is less on an average if the month is August in comparison to if the month is not August.

Both the monthly dummies included in the model capture the highest (March) and the lowest (August) new registered car in the year.

Call:

lm(formula = md$PureNumber ~ poly(md$trend, 2) + md$month[, 3] + md$month[, 8], data = md)

Residuals:

Min 1Q Median 3Q Max

-319277 -76685 1731 74234 349481

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 890324 6937 128.344 < 2e-16 \*\*\*

poly(md$trend, 2)1 409706 115670 3.542 0.000454 \*\*\*

poly(md$trend, 2)2 -597393 115665 -5.165 4.18e-07 \*\*\*

md$month[, 3] 201004 22933 8.765 < 2e-16 \*\*\*

md$month[, 8] -243525 22933 -10.619 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 115700 on 329 degrees of freedom

Multiple R-squared: 0.4305, Adjusted R-squared: 0.4236

F-statistic: 62.19 on 4 and 329 DF, p-value: < 2.2e-16

**Answer 5:** The plot of the residuals shows that the mean is close to zero as the residuals hover around zero with seasonal component partially captured in the model. Furthermore, it also shows the presence of serial correlation in the model.

The deviation in the model around 1992/93 could be because of the after impact of the Iraq war. The deviation in the data around 2008/2009 could be because of the after impact of the financial crisis going on in the world.



**Answer 6:** The p-value of F-stat of all the three models is significant suggesting that atleast one of the variables included in the respective models have significant impact on the new registration of cars. The p-value of all the independent variables is significant suggesting that all the independent variables have significant impact on the new registration of the cars. The R-square of the third model has been highest suggesting it to be the most-good-fit model. Furthermore, the Aic and **B**ic values of all the three models are quite large but they are lowest for Model 3, again suggesting it to be best model amongst the three models.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 |
| p-value of F-stat | 0 | 0 | 0 |
| Independent variables | All significant | All significant | All significant |
| R-square | 2% | 7% | 43% |
| Aic | 8916.731 | 8902.504 | 8742.649 |
| Bic | 8928.165 | 8917.749 | 8765.516 |

**Answer 7**: The ACF graph for the model 1 shows that there has been a big spike at the lag 0 which decays gradually but then picks up after lag 8. And the PACF graph for model 1 shows that there has been a significant spike at lag 1. Both the ACF and PACF graphs for model 2 in ACF have similar spikes with a big spike at the lag 0 which decays gradually but then picks up after lag 8. The spikes in the PACF are also similar with a significant spike at lag 1 but there are some bigger spikes at lag 8 etc. The ACF and PACF graphs for model 3 are somewhat different with negative spikes in ACF turning into positive spikes and picking only at lag 9 and after. The large spikes in PACF graph have also become smaller suggesting it to be a better model. However, there seems to be an autocorrelation in the models present which can be corrected by modeling AR(1) model.

Furthermore, the durbin-watson values for all the three models have been smaller than 2 with significant p-values concluding that there has been a serial correlation present in the model which needs to be corrected.

 



 

Durbin Watson Results:

model 1: DW = 1.0942, p-value < 2.2e-16

model 2: DW = 1.1488, p-value = 1.257e-15

model 3: DW = 0.77055, p-value < 2.2e-16

**Answer 8:** The graph clearly shows that the model 3 with seasonal dummies for March and August captures the seasonality component and seems to be the good model. But the standard errors for the prediction for this model has also been large.

> fc1

$fit

1 2 3 4 5 6 7 8 9

846377.8 846620.3 846862.8 847105.4 847347.9 847590.4 847833.0 848075.5 848318.0

10 11 12 13 14 15 16 17 18

848560.6 848803.1 849045.6 849288.2 849530.7 849773.2 850015.8 850258.3 850500.8

19 20 21 22 23 24

850743.4 850985.9 851228.4 851471.0 851713.5 851956.0

$se.fit

[1] 16462.092 16388.106 16314.233 16240.474 16166.832 16093.308 16019.903 15946.620

[9] 15873.459 15800.423 15727.514 15654.732 15582.080 15509.560 15437.174 15364.923

[17] 15292.809 15220.834 15149.001 15077.311 15005.767 14934.370 14863.122 14792.027

> fc2

$fit

1 2 3 4 5 6 7 8 9

773838.4 775387.9 776929.6 778463.4 779989.3 781507.4 783017.6 784519.9 786014.3

10 11 12 13 14 15 16 17 18

787500.9 788979.6 790450.4 791913.3 793368.4 794815.6 796254.9 797686.3 799109.9

19 20 21 22 23 24

800525.6 801933.4 803333.4 804725.5 806109.7 807486.0

$se.fit

[1] 24046.98 23759.87 23475.39 23193.55 22914.36 22637.86 22364.04 22092.94 21824.57

[10] 21558.95 21296.09 21036.03 20778.77 20524.33 20272.75 20024.03 19778.21 19535.30

[19] 19295.32 19058.31 18824.27 18593.24 18365.23 18140.28

> fc3

$fit

1 2 3 4 5 6 7 8

779172.3 780710.0 538714.8 783761.9 785276.0 786782.2 788280.6 990775.6

9 10 11 12 13 14 15 16

791253.9 792728.7 794195.6 795654.7 797105.9 798549.3 556459.8 801412.4

17 18 19 20 21 22 23 24

802832.2 804244.1 805648.2 1008048.8 808432.7 809813.1 811185.7 812550.5

$se.fit

[1] 19098.311 18875.431 27776.348 18435.972 18219.419 18005.004 17792.738 27201.076

[9] 17374.715 17168.987 16965.467 16764.172 16565.117 16368.317 26190.476 15981.549

[17] 15791.615 15604.001 15418.727 25751.979 15055.265 14877.112 14701.368 14528.052

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Model 1 | | Model 2 | | Model 3 | |
|  | Forecast | SE | Forecast | SE | Forecast | SE |
| 1 | 846377.8 | 16462.09 | 773838.4 | 24046.98 | 779172.3 | 19098.31 |
| 2 | 846620.3 | 16388.11 | 775387.9 | 23759.87 | 780710 | 18875.43 |
| 3 | 846862.8 | 16314.23 | 776929.6 | 23475.39 | 538714.8 | 27776.35 |
| 4 | 847105.4 | 16240.47 | 778463.4 | 23193.55 | 783761.9 | 18435.97 |
| 5 | 847347.9 | 16166.83 | 779989.3 | 22914.36 | 785276 | 18219.42 |
| 6 | 847590.4 | 16093.31 | 781507.4 | 22637.86 | 786782.2 | 18005 |
| 7 | 847833 | 16019.9 | 783017.6 | 22364.04 | 788280.6 | 17792.74 |
| 8 | 848075.5 | 15946.62 | 784519.9 | 22092.94 | 990775.6 | 27201.08 |
| 9 | 848318 | 15873.46 | 786014.3 | 21824.57 | 791253.9 | 17374.72 |
| 10 | 848560.6 | 15800.42 | 787500.9 | 21558.95 | 792728.7 | 17168.99 |
| 11 | 848803.1 | 15727.51 | 788979.6 | 21296.09 | 794195.6 | 16965.47 |
| 12 | 849045.6 | 15654.73 | 790450.4 | 21036.03 | 795654.7 | 16764.17 |
| 13 | 849288.2 | 15582.08 | 791913.3 | 20778.77 | 797105.9 | 16565.12 |
| 14 | 849530.7 | 15509.56 | 793368.4 | 20524.33 | 798549.3 | 16368.32 |
| 15 | 849773.2 | 15437.17 | 794815.6 | 20272.75 | 556459.8 | 26190.48 |
| 16 | 850015.8 | 15364.92 | 796254.9 | 20024.03 | 801412.4 | 15981.55 |
| 17 | 850258.3 | 15292.81 | 797686.3 | 19778.21 | 802832.2 | 15791.62 |
| 18 | 850500.8 | 15220.83 | 799109.9 | 19535.3 | 804244.1 | 15604 |
| 19 | 850743.4 | 15149 | 800525.6 | 19295.32 | 805648.2 | 15418.73 |
| 20 | 850985.9 | 15077.31 | 801933.4 | 19058.31 | 1008049 | 25751.98 |
| 21 | 851228.4 | 15005.77 | 803333.4 | 18824.27 | 808432.7 | 15055.27 |
| 22 | 851471 | 14934.37 | 804725.5 | 18593.24 | 809813.1 | 14877.11 |
| 23 | 851713.5 | 14863.12 | 806109.7 | 18365.23 | 811185.7 | 14701.37 |
| 24 | 851956 | 14792.03 | 807486 | 18140.28 | 812550.5 | 14528.05 |

**Answer 9**:

The results for the log linear trend show that trend has significant impact on predicting the new registration of cars. However, trend explains only about 2% variation in the new registration of cars. The coefficient of trend has been close to zero which means that with every month there has been an increase in the new registration of cars by 0.0027% on an average.

The plot of the residuals shows that the mean is close to zero as the residuals hover around zero. There seems to be a seasonal component in the data. Furthermore, it also shows the presence of serial correlation in the model. The Aic and **B**ic values are much smaller than the earlier models suggesting it to be better than the earlier models and the best model could be a mix of this model and model 3.

Call:

lm(formula = md$Number ~ md$trend, data = md)

Residuals:

Min 1Q Median 3Q Max

-0.49569 -0.11498 0.01627 0.13469 0.34730

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.363e+01 1.924e-02 708.535 < 2e-16 \*\*\*

md$trend 2.739e-04 9.957e-05 2.751 0.00627 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1754 on 332 degrees of freedom

Multiple R-squared: 0.02229, Adjusted R-squared: 0.01934

F-statistic: 7.568 on 1 and 332 DF, p-value: 0.006267

> AIC(model4)

[1] -210.7718

> BIC(model4)

[1] -199.3383



**Answer 10**:

The forecast using the model 4 with log of registrations is made as follows:

$forecast

1 2 3 4 5 6 7 8 9

834386.0 834614.5 834843.2 835071.9 835300.6 835529.5 835758.3 835987.3 836216.3

10 11 12 13 14 15 16 17 18

836445.4 836674.5 836903.7 837133.0 837362.3 837591.7 837821.1 838050.7 838280.2

19 20 21 22 23 24

838509.9 838739.6 838969.3 839199.2 839429.1 839659.0

The graph comparing the forecast of the four models is also presented in the figure below.