Question 2 (31 Marks)

2(a) (4 marks)

$$\begin{aligned}\widehat{HOUSTNSA} = \underset{(4.196)}{92.871} &- \underset{(5.911)}{4.592}\ Jan - \underset{(5.911)}{1.935}\ Feb + \underset{(5.934)}{26.184}\ Mar \\ &+ \underset{(5.934)}{41.452}\ Apr + \underset{(5.934)}{46.786}\ May + \underset{(5.934)}{46.263}\ Jun + \underset{(5.934)}{40.937}\ Jul \\ &+ \underset{(5.934)}{38.714}\ Aug + \underset{(5.934)}{32.252}\ Sep + \underset{(5.934)}{36.170}\ Oct + \underset{(5.934)}{15.600}\ Nov\end{aligned} \tag{1}$$

The above linear regression estimates the US monthly 'housing starts' based on the month that is being modelled. The intercept $\beta_0 = 92.781$ is the estimated 'housing starts', on average, in US for the month of December. The variables in the linear regression are seasonal dummies which mean they only take a binary value (0 or 1). The β values for the seasonal dummies are the average change in the 'housing starts' relative to the month December. - January, on average, has 4.592 less 'housing starts' than the month of December, i.e. NUMBER of 'housing starts in January.

2(b) (4 marks)

Steps

In order to formulate the linear regression, first we need to determine the intercept: From equation 1 we can determine the values of each month because of the dummy variables. $92.871 - 4.592 = c \rightarrow c = 88.280$, where the LHS is the month of Jan from calculated from equation 1.

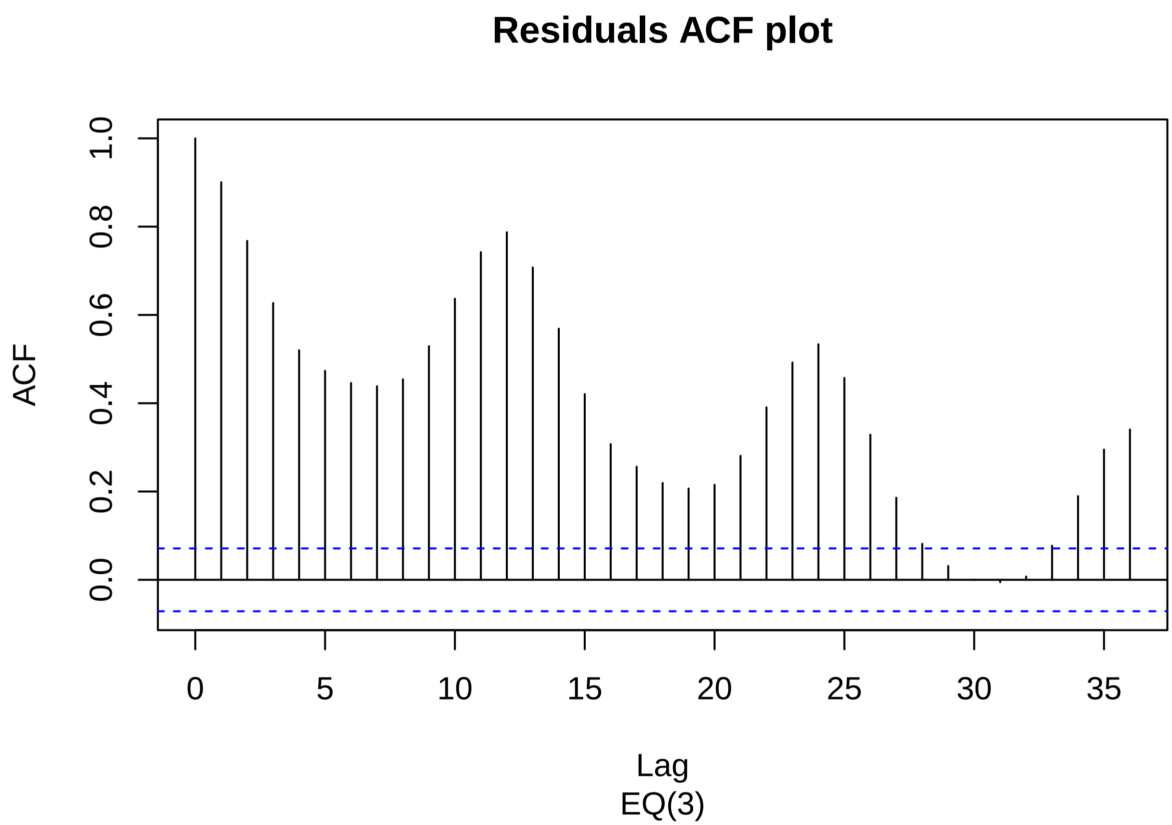
Next we need to determine the β values for Feb - Dec. Since we know the intercept for the new equation, we can substitute it in.

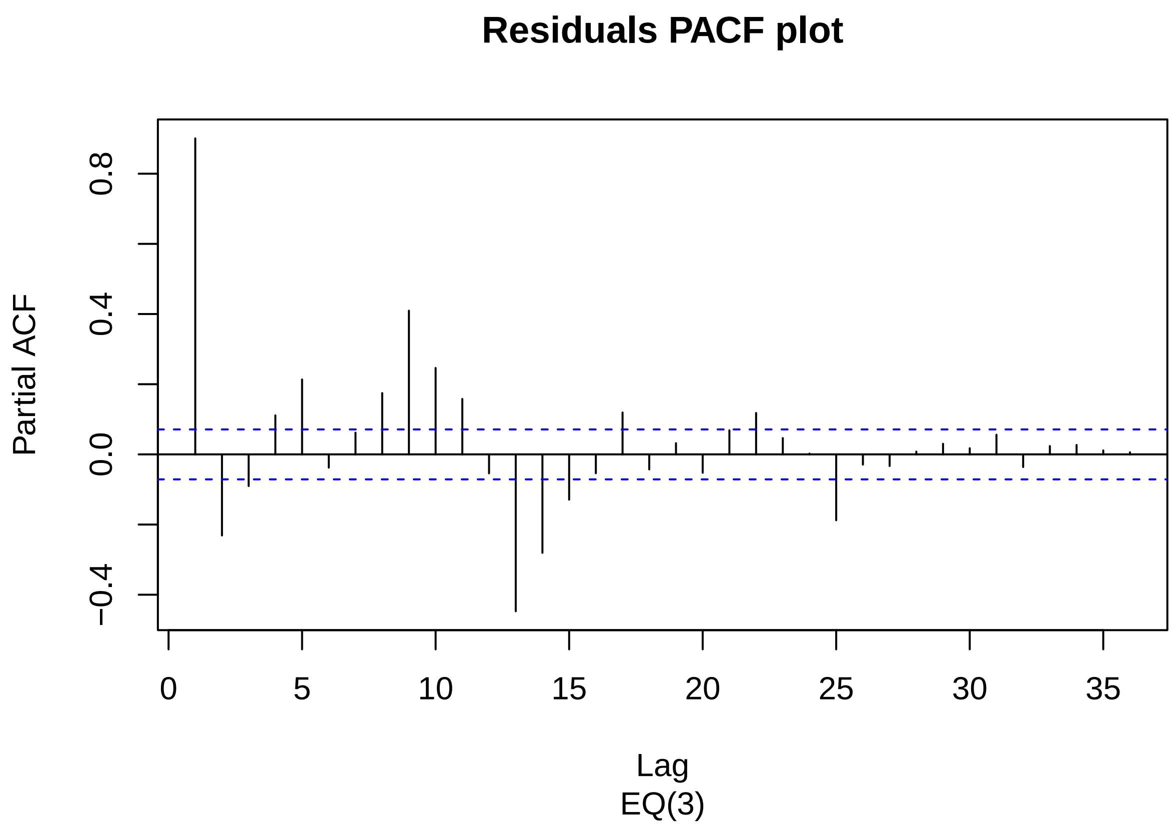
$$
\begin{aligned}
92.871 + 1.935 &= 88.280 + \beta_2 \ Feb \\
&\rightarrow \ \beta_2 = 2.656 \\[2ex]
92.871 + 26.184 &= 88.280 + \beta_3 \ Mar \\
&\rightarrow \ \beta_3 = 30.776 \\[2ex]
92.871 + 41.452 &= 88.280 + \beta_4 \ Apr \\
&\rightarrow \ \beta_4 = 46.044 \\[2ex]
92.871 + 46.786 &= 88.280 + \beta_5 \ May \\
&\rightarrow \ \beta_5 = 51.377 \\[2ex]
92.871 + 46.263 &= 88.280 + \beta_6 \ Jun \\
&\rightarrow \ \beta_6 = 50.855 \\[2ex]
92.871 + 40.937 &= 88.280 + \beta_7 \ Jul \\
&\rightarrow \ \beta_7 = 45.528 \\[2ex]
92.871 + 38.714 &= 88.280 + \beta_8 \ Aug \\
&\rightarrow \ \beta_8 = 43.306 \\[2ex]
92.871 + 32.252 &= 88.280 + \beta_9 \ Sep \\
&\rightarrow \ \beta_9 = 36.844 \\[2ex]
92.871 + 36.170 &= 88.280 + \beta_{10} \ Oct \\
&\rightarrow \ \beta_{10} = 40.762 \\[2ex]
92.871 + 15.600 &= 88.280 + \beta_{11} \ Nov \\
&\rightarrow \ \beta_{11} = 20.192 \\[2ex]
92.871 &= 88.280 + \beta_{12} \ Dec \\
&\rightarrow \ \beta_{12} = 4.592
\end{aligned}
$$

$$\widehat{HOUSTNSA} = 88.280 \ + \ 2.656 \ Feb \ + \ 30.776 \ Mar \ + \ 46.044 \ Apr \\ + \ 51.377 \ May \ + \ 50.855 \ Jun \ + \ 45.528 \ Jul \ + \ 43.306 \ Aug \\ + \ 36.844 \ Sep \ + \ 40.762 \ Oct \ + \ 20.192 \ Nov \ + \ 4.592 \ Dec \tag{2}$$

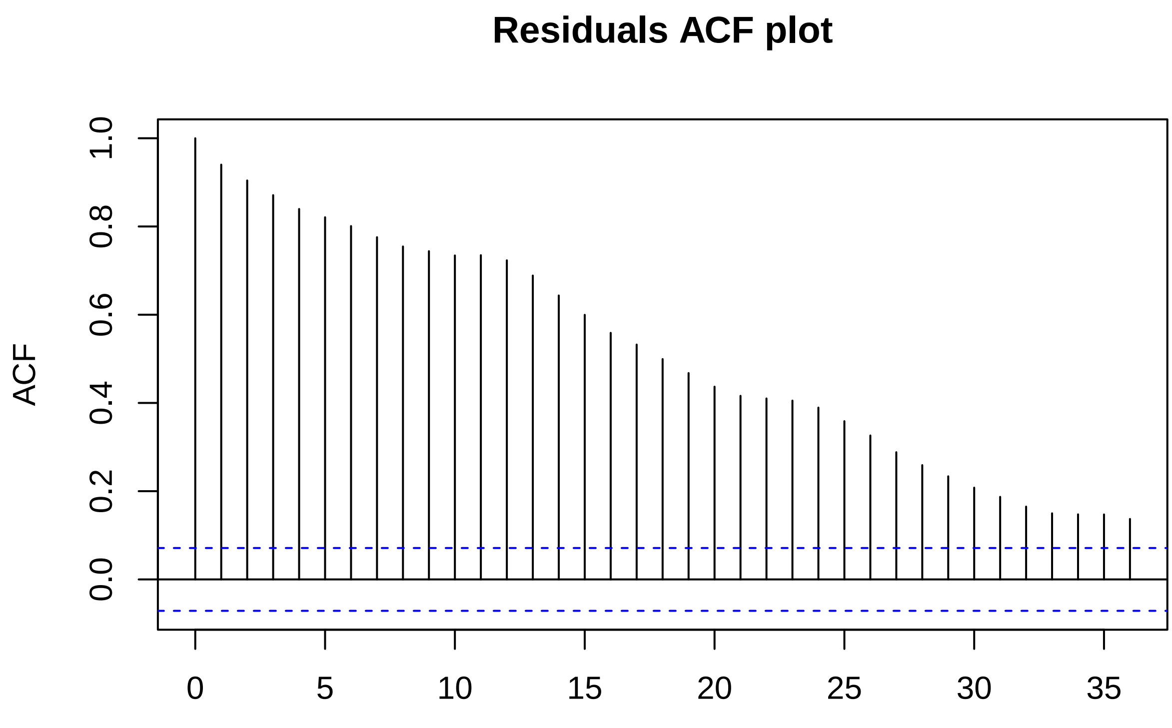
2(c) (6 marks)

$$\widehat{HOUSTNSA} = \underset{(1.377)}{119.3} \tag{3}$$

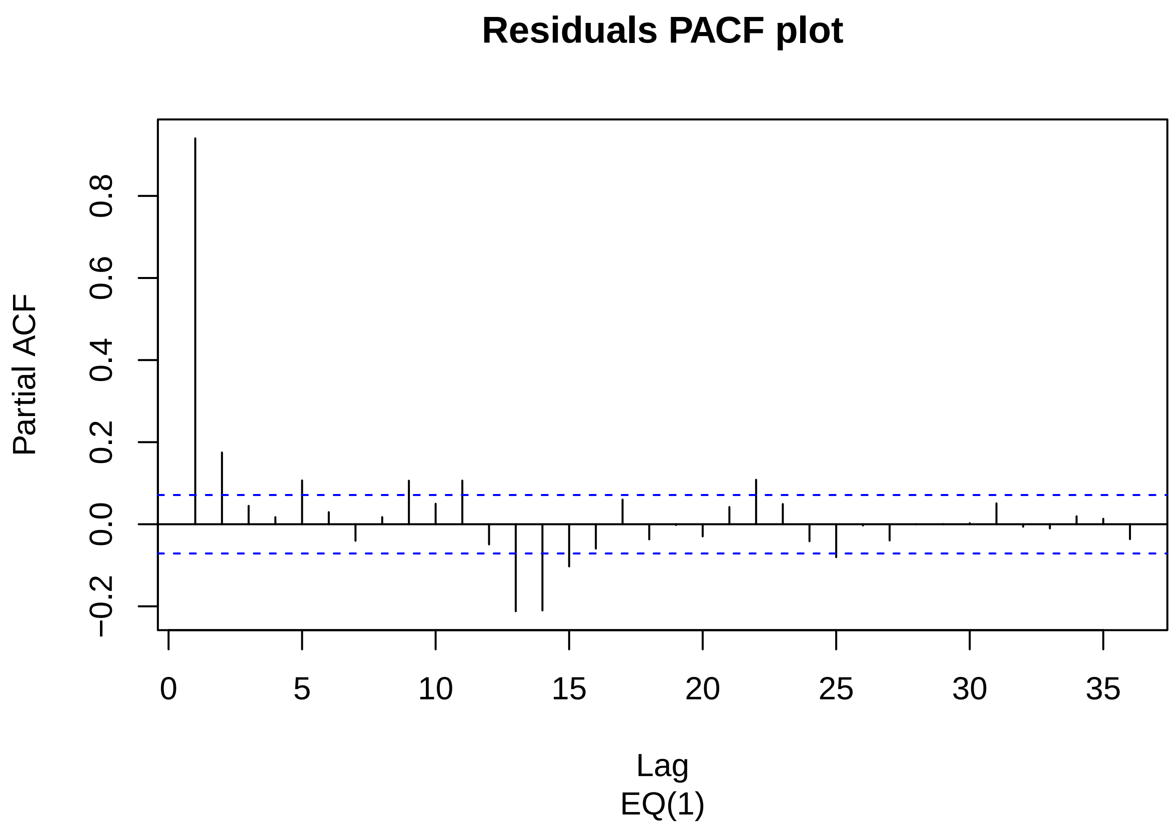




The residual ACF plot for the linear model EQ(3) shows a few things. Firstly, each non-seasonal lag (lag $\neq \{12, 24, 36\}$) have a large positive spike which indicates the existence of a trend. Secondly, the seasonal lags have large positive spikes which indicates that there is also seasonality in the data. So we can attribute the gradual decrease in the lag values is because of trend while the scallop pattern is due to the seasonality. As a result of the trend and seasonality it can safely be said that there is no stationarity in the data and thus implying that the error term is not white noise.



EQ(1)



The residual ACF plot for the linear model EQ(1) shows a significant difference compared to the residual ACF plot for the linear model EQ(3). The linear model EQ(1) has no seasonality in the residual ACF plot since the linear model has already captured it. What is remaining is the trend as depicted by decreasing values as the lags increase, showing that the seasonal dummy variables have improved the model. Despite the improvement, there is still a trend which means that there is no stationarity in the data and thus implying the error term is not white noise.

2(d) (9 marks)

$$\begin{aligned}\widehat{HOUSTNSA} = & \underset{(2.165)}{-14.042} - \underset{(2.003)}{11.014}\ Jan - \underset{(2.091)}{20.084}\ Feb + \underset{(2.179)}{47.498}\ Mar \\ & + \underset{(2.652)}{40.081}\ Apr + \underset{(2.360)}{28.510}\ May + \underset{(2.198)}{21.153}\ Jun + \underset{(2.123)}{15.297}\ Jul \\ & + \underset{(2.067)}{17.297}\ Aug + \underset{(2.095)}{13.490}\ Sep + \underset{(2.049)}{22.810}\ Oct + \underset{(2.168)}{0.347}\ Nov \\ & + \underset{(0.036)}{0.775}\ HOUSTNSA_{t-1} + \underset{(0.036)}{0.177}\ HOUSTNSA_{t-2}\end{aligned} \tag{4}$$

$$
\begin{aligned}
H_0 &: \underset{i\in\{1,2,3,4,5,6,7,8,9,10,11\}}{\forall} \beta_i = 0 \\
H_1 &: \underset{i\in\{1,2,3,4,5\}}{\exists} \beta_i \neq 0 \text{ at least one regressor coef is zero} \\
\text{Significance Level} &: \alpha = 0.05 \\
\text{Unresticted Model} &: \widehat{HOUSTNSA} = \underset{(2.165)}{-14.042} \; - \; \underset{(2.003)}{11.014}\; Jan \; - \; \underset{(2.091)}{20.084}\; Feb \\
&+ \; \underset{(2.179)}{47.498}\; Mar \; + \; \underset{(2.652)}{40.081}\; Apr \; + \; \underset{(2.360)}{28.510}\; May \\
&+ \; \underset{(2.198)}{21.153}\; Jun \; + \; \underset{(2.123)}{15.297}\; Jul \; + \; \underset{(2.067)}{17.297}\; Aug \\
&+ \; \underset{(2.095)}{13.490}\; Sep \; + \; \underset{(2.049)}{22.810}\; Oct \; + \; \underset{(2.168)}{0.347}\; Nov \\
&+ \; \underset{(0.036)}{0.775}\; HOUSTNSA_{t-1} \; + \; \underset{(0.036)}{0.177}\; HOUSTNSA_{t-2} \\
\text{Resticted Model} &: \widehat{HOUSTNSA} = \underset{(1.970)}{14.598} \; + \; \underset{(0.035)}{1.110}\; HOUSTNSA_{t-1} \; - \; \underset{(0.035)}{0.232}\; HOUSTNSA_{t-2} \\
\text{Test stat and null dist} &: \frac{(SSR_R - SSR_{UR})}{SSR_{UR}} \frac{(n-k-1)}{q} \sim F_{(q,n-k-1)} = F_{11,742} \\
F_{calc} &= 74.7468276 \\
F_{crit} &= 2.0101347 \\
\text{Decision rule} &: \text{reject } H_0 \text{ if } F_{calc} > F_{crit} \\
\text{Decision} &: \text{Since } 74.7468276 > 2.0101347\text{, reject } H_0
\end{aligned}
$$

In conclusion, at 0.05 level of significance, we reject the null hypothesis that the seasonal dummies (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct) are jointly insignificant in effecting the 'Housing Starts' in favour of the alternative hypothesis that at least one of these variables are significant.

2(e) (8 marks)

$$
\begin{aligned}
\widehat{HOUSTNSA} = \underset{(2.264)}{-1.503} \; + \; \underset{(1.590)}{17.299}\; Q1 \; + \; \underset{(1.772)}{16.273}\; Q2 \; + \; \underset{(1.534)}{6.685}\; Q3 \\
+ \; \underset{(0.040)}{1.007}\; HOUSTNSA_{t-1} \; - \; \underset{(0.039)}{0.079}\; HOUSTNSA_{t-2}
\end{aligned} \tag{5}
$$

Monthly Model (Unresticted):

$$
\begin{aligned}
LHS = \beta_0 + \beta_1\, Jan + \beta_2\, Feb + \beta_3\, Mar + \beta_4\, Apr + \beta_5\, May + \beta_6\, Jun + \beta_7\, Jul + \beta_8\, Aug \\
+ \; \beta_9\, Sep \; + \; \beta_{10}\, Oct \; + \; \beta_{11}\, Nov \; + \; \beta_{13}\, HOUSTNSA_{t-1} \; + \; \beta_{14}\, HOUSTNSA_{t-2}
\end{aligned} \tag{6}
$$

Quarterly Model:

$$RHS = \alpha_0 \ + \ \alpha_1 \, Q1 \ + \ \alpha_2 \, Q2 \ + \ \alpha_3 \, Q3 \ + \ \alpha_4 \, HOUSTNSA_{t-1} \ + \ \alpha_5 \, HOUSTNSA_{t-2} \quad (7)$$

LHS == RHS

$$\begin{aligned}
\widehat{HOUSTNSA} &= \beta_0 \ + \ \beta_1 \ Jan \ + \ \beta_2 \ Feb \ + \ \beta_3 \ Mar \ + \ \beta_4 \ Apr \ + \ \beta_5 \ May \ + \ \beta_6 \ Jun \\
&\quad + \ \beta_7 \ Jul \ + \ \beta_8 \ Aug \ + \ \beta_9 \ Sep \ + \ \beta_{10} \ Oct \ + \ \beta_{11} \ Nov \\
&\quad + \ \beta_{13} \ HOUSTNSA_{t-1} \ + \ \beta_{14} \ HOUSTNSA_{t-2} \\
\widehat{HOUSTNSA} &= \alpha_0 \ + \ \alpha_1 \ Q1 \ + \ \alpha_2 \ Q2 \ + \ \alpha_3 \ Q3 \ + \ \alpha_4 \ HOUSTNSA_{t-1} \ + \ \alpha_5 \ HOUSTNSA_{t-2}
\end{aligned}$$

$$\begin{aligned}
\alpha_0 &= \beta_0 + \beta_{10} + \beta_{11} \\
\alpha_1 &= \beta_1 + \beta_2 + \beta_3 \\
\alpha_2 &= \beta_4 + \beta_5 + \beta_6 \\
\alpha_3 &= \beta_7 + \beta_8 + \beta_9 \\
\alpha_4 &= \beta_{13} \\
\alpha_5 &= \beta_{14}
\end{aligned}$$

Quarterly Model (Restricted):

$$\begin{aligned}
RHS = \beta_0 + \beta_{10} + \beta_{11} \ &+ \ (\beta_1 + \beta_2 + \beta_3) \ Q1 \ + \ (\beta_4 + \beta_5 + \beta_6) \ Q2 \\
&+ \ (\beta_7 + \beta_8 + \beta_9) \ Q3 \ + \ \beta_{13} \ HOUSTNSA_{t-1} \ + \ \beta_{14} \ HOUSTNSA_{t-2}
\end{aligned} \tag{8}$$

$$\begin{aligned}
H_0 &: \underset{i \in \{0,1,2,3\}}{\forall} \alpha_i = 0 \\
H_1 &: \underset{i \in \{0,1,2,3\}}{\exists} \alpha_i \neq 0 \text{ at least one regressor coef is zero} \\
\text{Test stat and null dist} &: \frac{(SSR_R - SSR_{UR})}{SSR_{UR}} \frac{(n-k-1)}{q} \sim F_{(q,n-k-1)} = F_{3,752} \\
F_{calc} &= 189.6544468 \\
F_{crit} &= 3.1334909 \\
\text{Decision rule} &: \text{reject } H_0 \text{ if } F_{calc} > F_{crit} \\
\text{Decision} &: \text{Since } 189.6544468 > 3.1334909, \text{ reject } H_0
\end{aligned}$$

In conclusion, at 0.05 level of significance, we reject the null hypothesis that there is only quarterly seasonality in the AR(2) model in favour of the alternative hypothesis that there may also be monthly seasonality in the AR(2) model.