Problem 1

Part a

It's been a while since the HW2, so I want to use that question as an opportunity to repeat what we did in previous chapters. I started from first plots and applied them to slaughtered pigs in New South Wales data. I performed usual plots in first four figures, however, I was able to see trend and seasonality clearly in Fig-5.

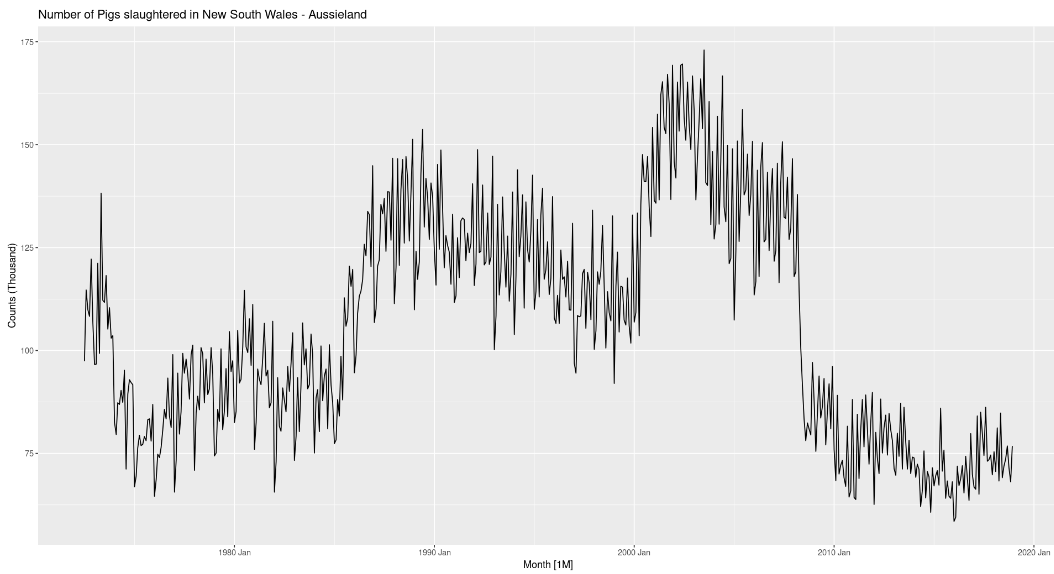


Figure 1: Number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. From the plot, it is hard to tell of existence of any patterns.

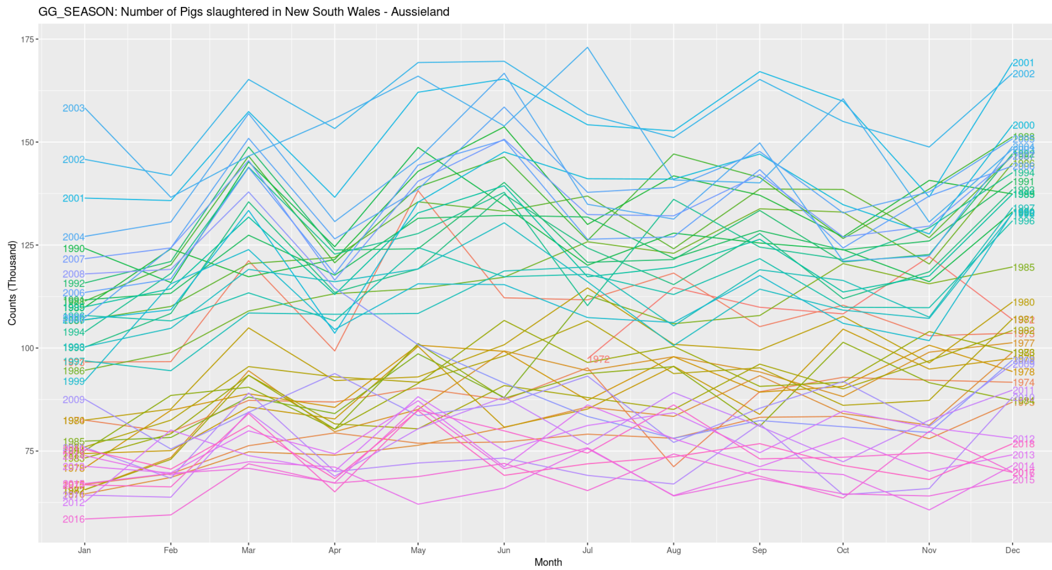


Figure 2: Seasonal plot for the same data. There seems like a seasonal pattern in the data, however, solely from this plot, we cannot be sure.

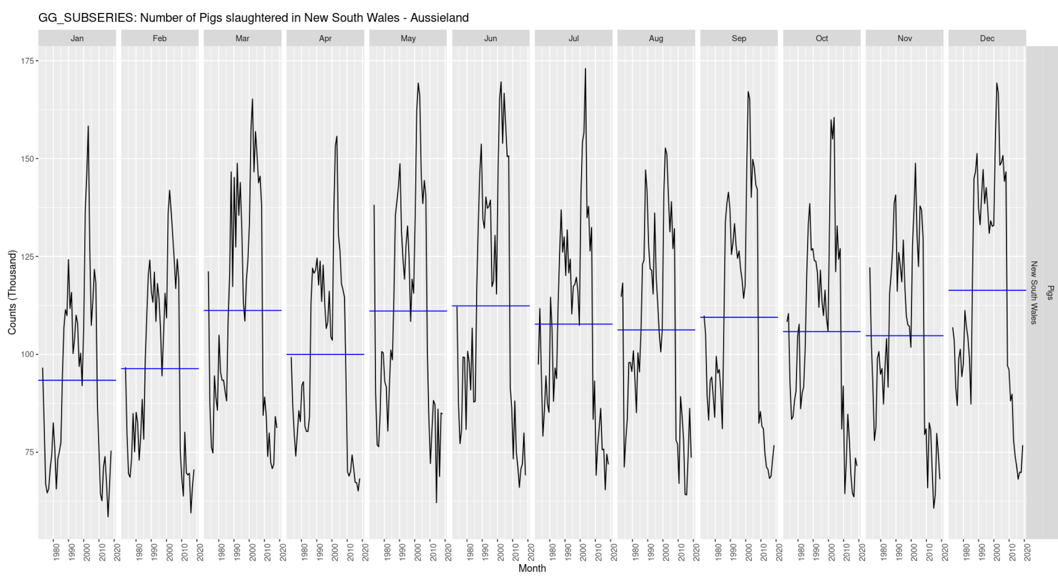


Figure 3: Subseries plot for number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. Monthly averages look more or less the same except 4 individual months (January, February, April and December).

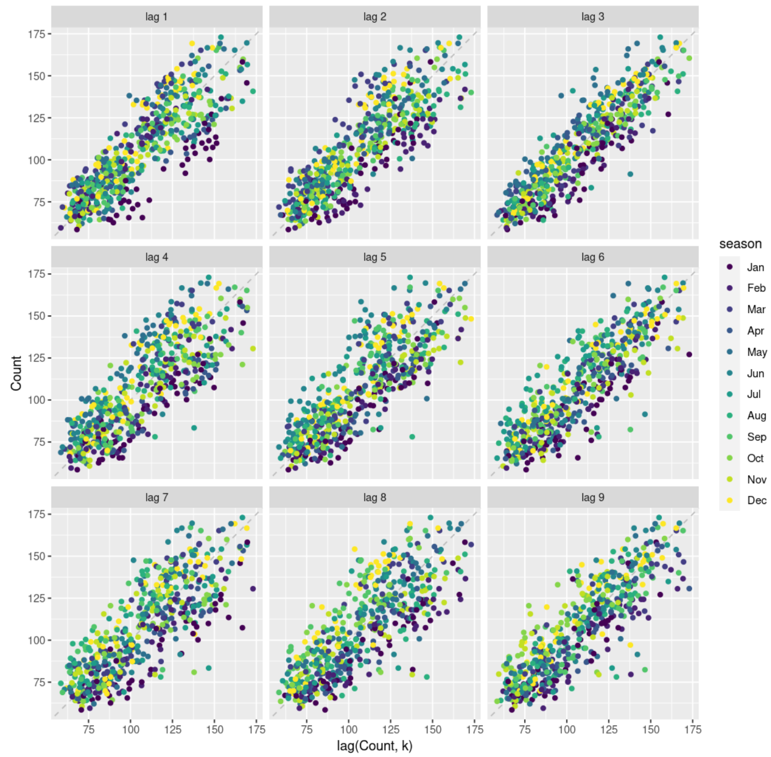


Figure 4: Lags respect to months for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. All the lags show a linear relation in the data.

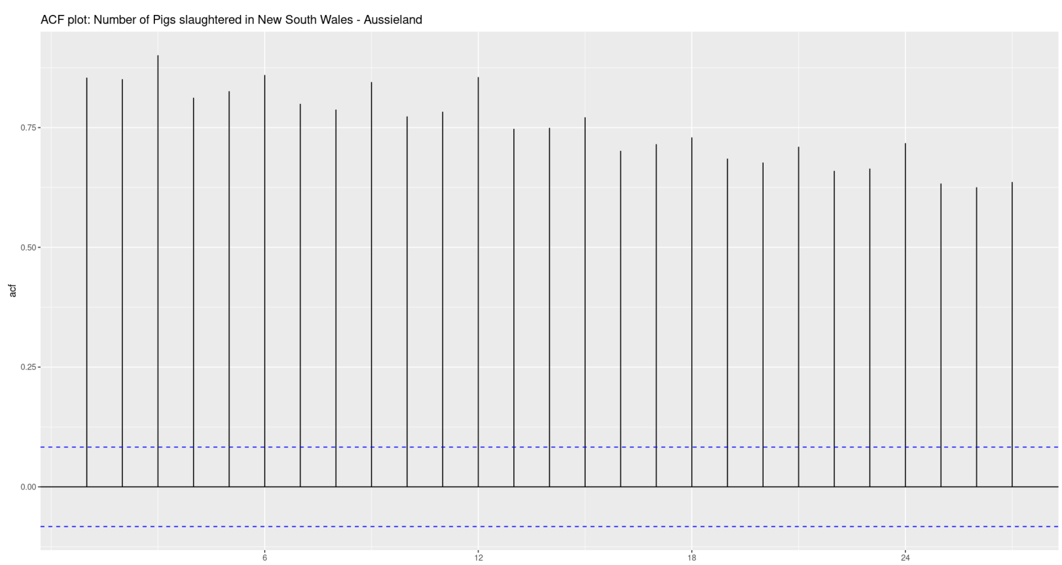


Figure 5: Auto correlation function for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. Plot shows that the data set is meaningful in terms of patterns and there is a strong visual evidence for trend and seasonality.

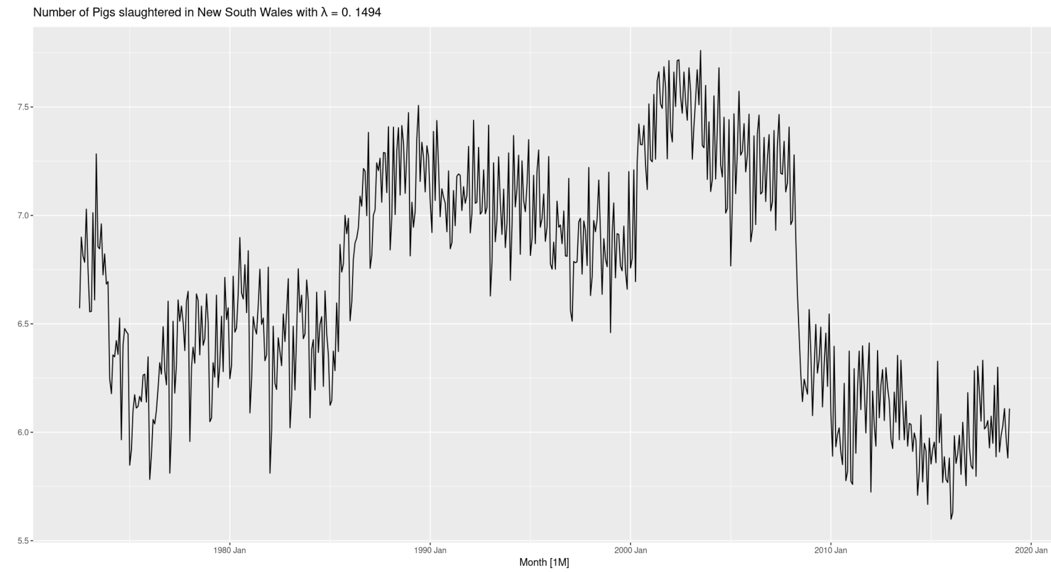


Figure 6: Box-Cox transformation ($\lambda = 0.1494$) for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. There is no significant changes other than the slight change in amplitude of fluctuations.

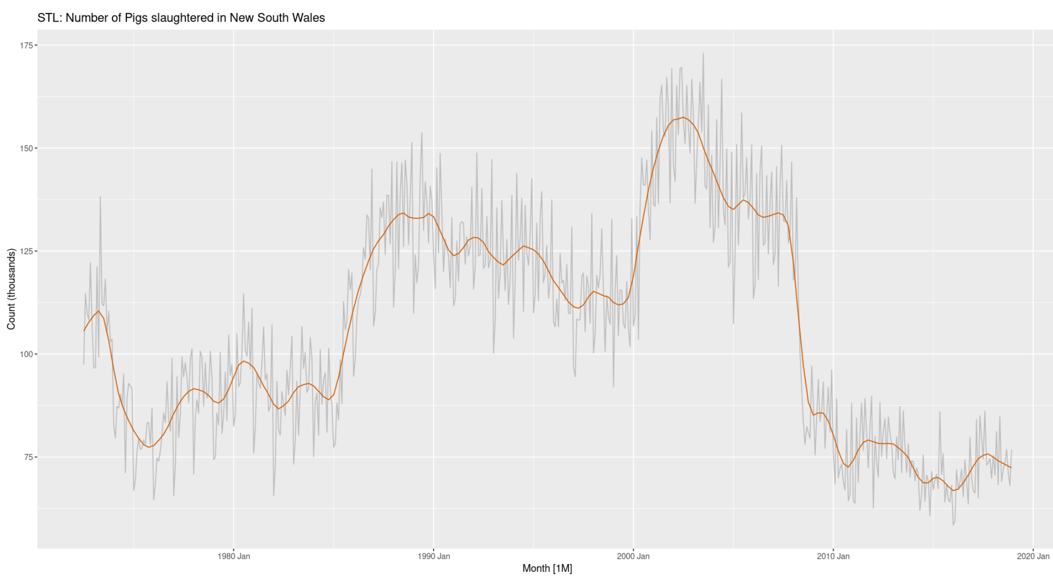


Figure 7: STL for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. Orange line represents trend in the data.

In addition to Fig-5, more detailed analysis is done in Fig-8. Seasonal and trend patterns are convenient. Through Fig-11-14, we see different methods applied to data, and different between methods mostly based on irregular components. In Fig-13 and 14, comparing to other methods, irregular components are depicted smaller but seasonality components do not look any better.

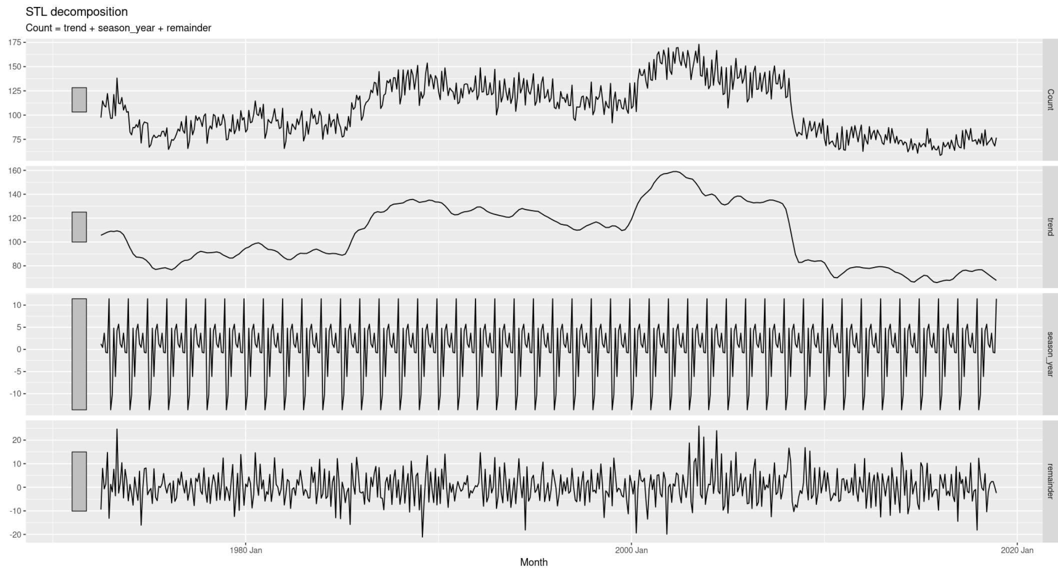


Figure 8: STL decomposition for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. Trend window (window=13) was selected from seasonal pattern (window= "periodic"). A clear seasonal pattern can be seen easily. Furthermore, the trend is in agreement what we have found in Fig-7.

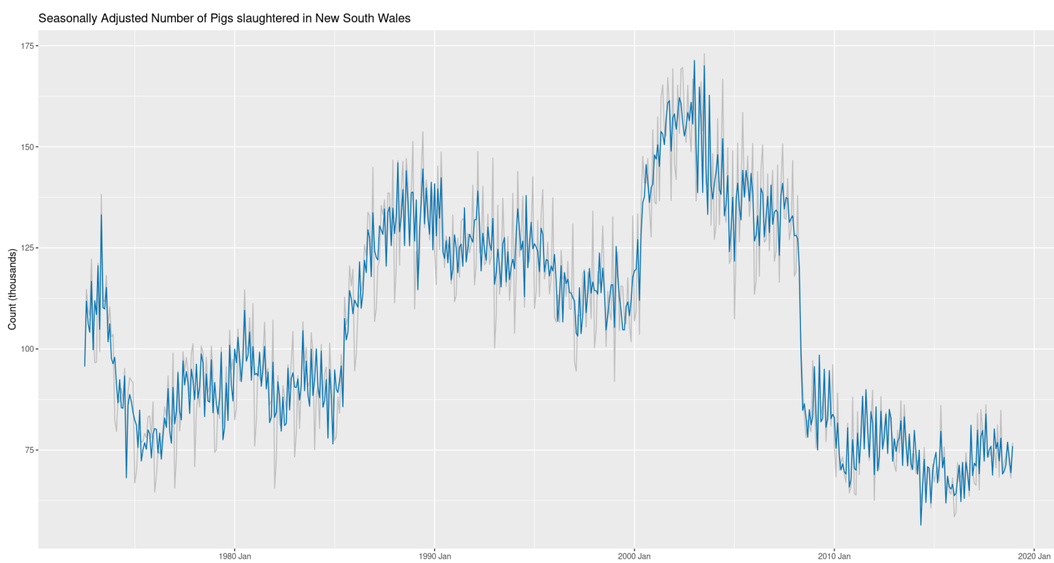


Figure 9: Seasonally adjusted data for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

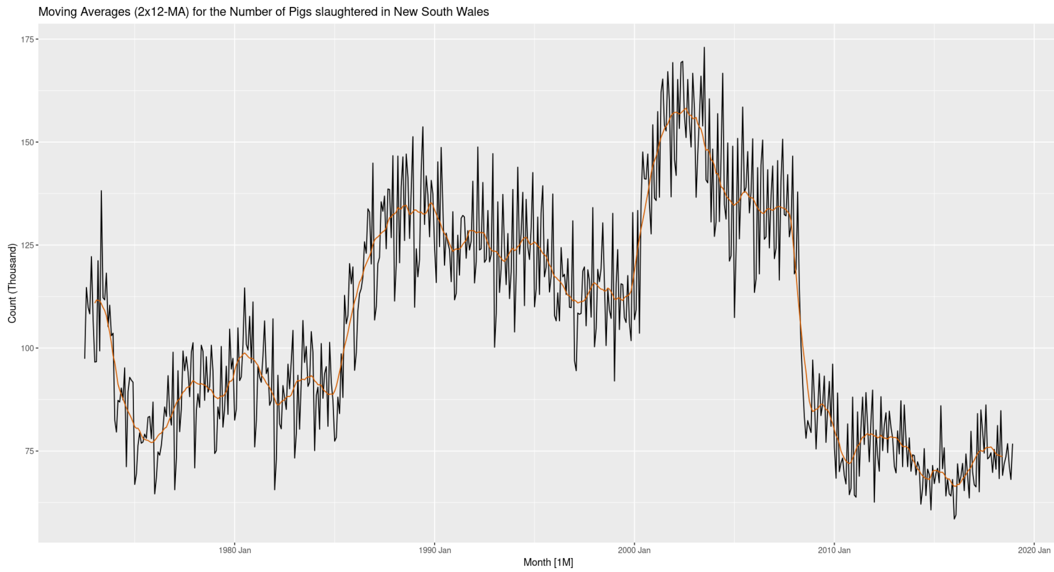


Figure 10: Moving averages (2x12 MA) for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

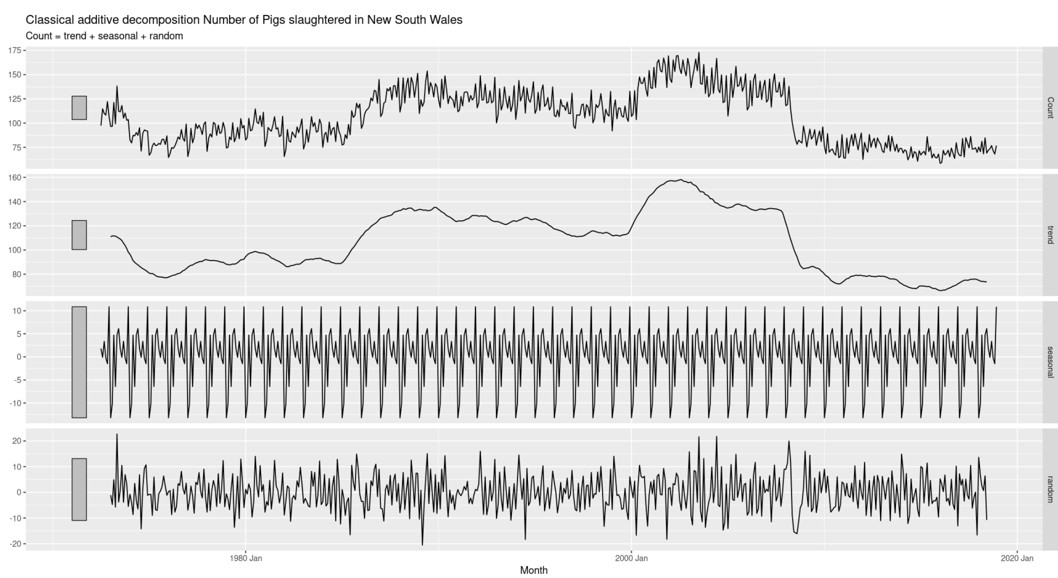


Figure 11: Classical additive decomposition for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

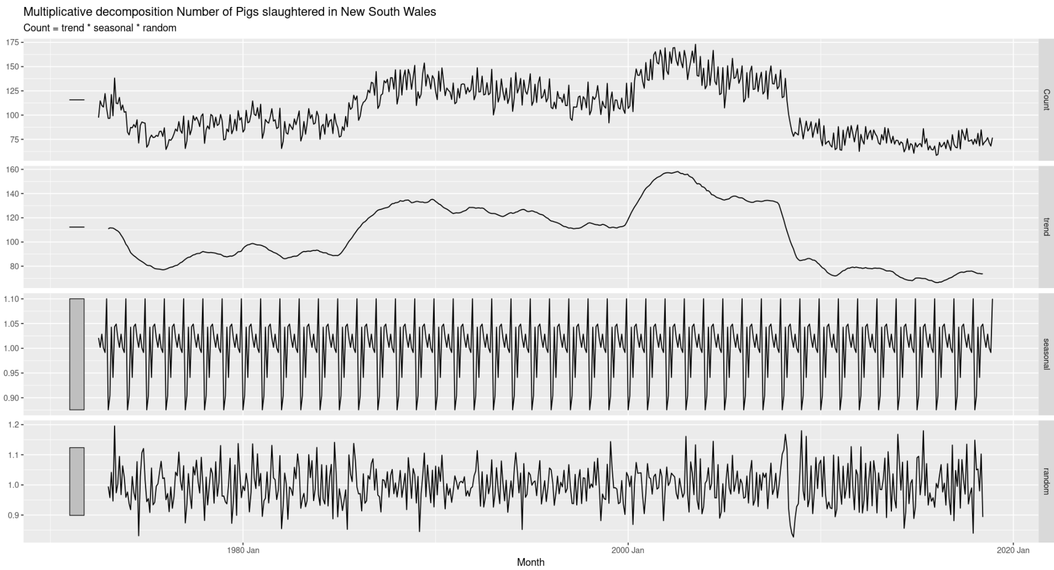


Figure 12: Multiplicative decomposition for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

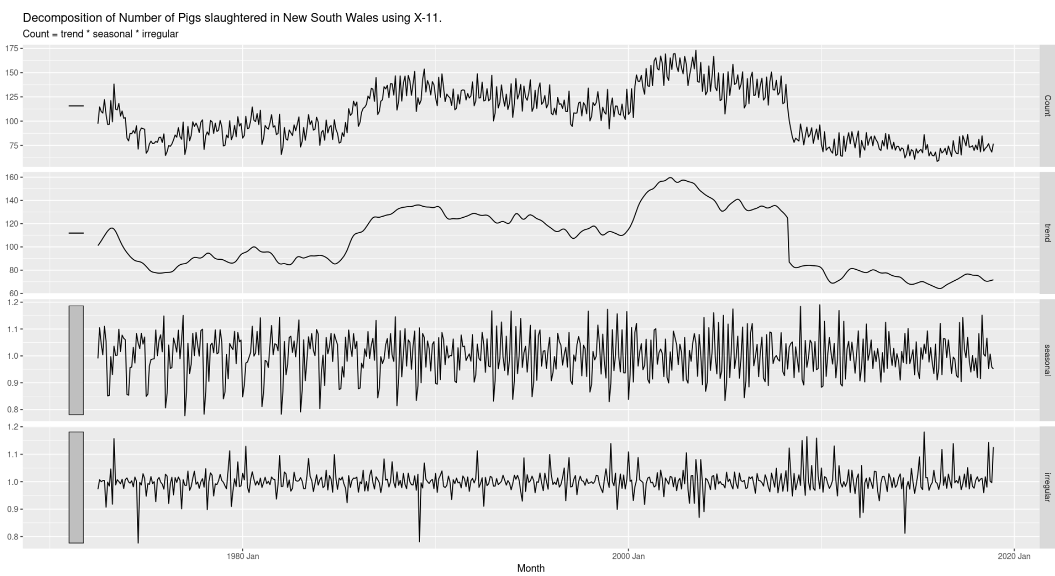


Figure 13: X-11 decomposition for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

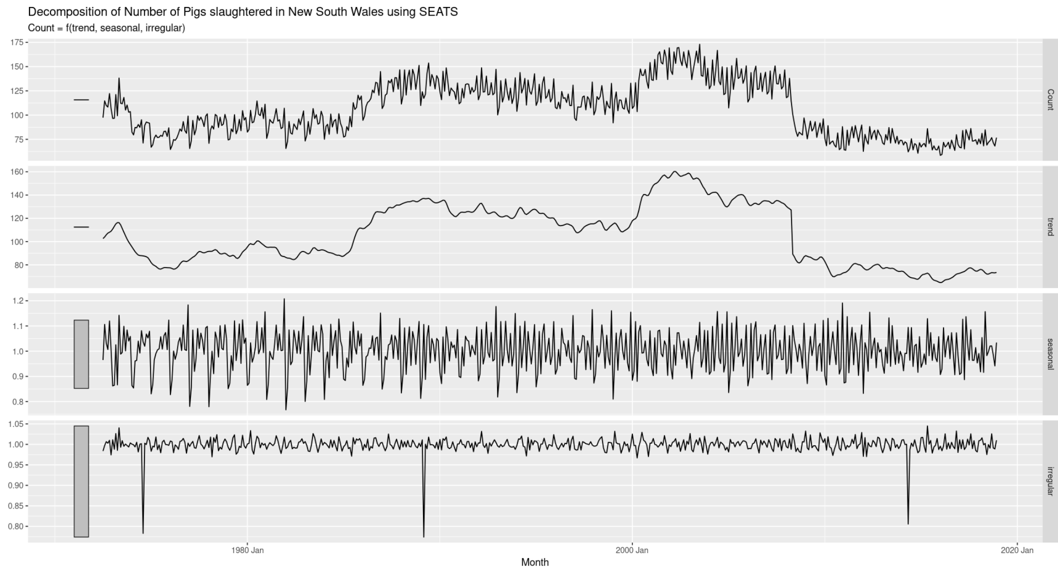


Figure 14: SEATS decomposition for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

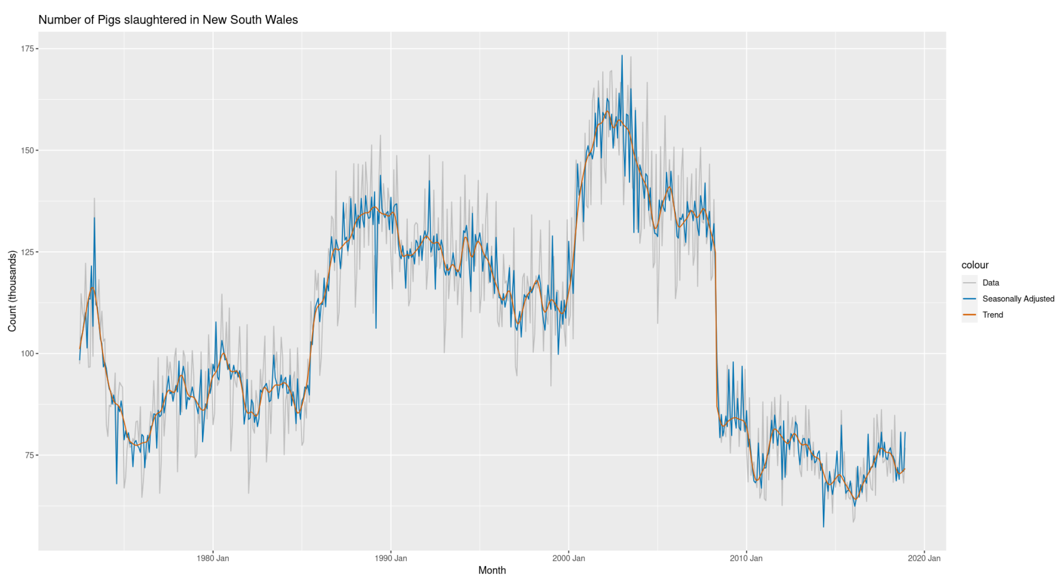
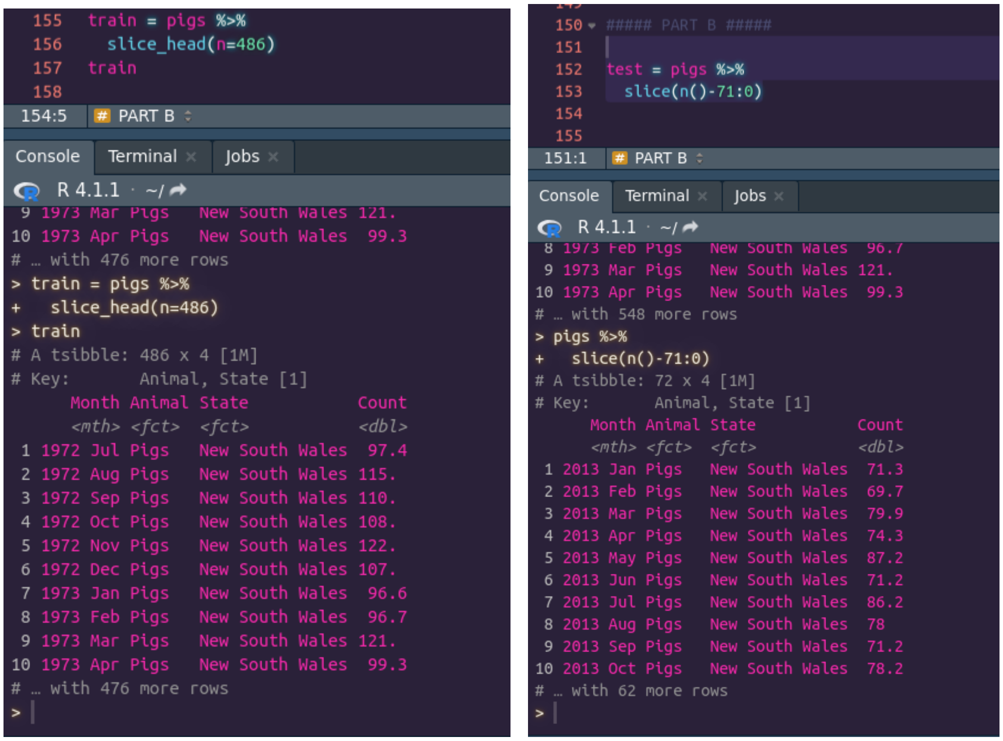


Figure 15: Trend-cycle vs seasonally adjusted data for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018.

Part b



```
155  train = pigs %>%
156    slice_head(n=486)
157  train
158
154:5   PART B

Console  Terminal  Jobs
R 4.1.1 · ~/
 9 1973 Mar Pigs   New South Wales 121.
10 1973 Apr Pigs   New South Wales  99.3
# … with 476 more rows
> train = pigs %>%
+   slice_head(n=486)
> train
# A tsibble: 486 x 4 [1M]
# Key:       Animal, State [1]
      Month Animal State           Count
      <mth> <fct>  <fct>           <dbl>
 1 1972 Jul Pigs   New South Wales  97.4
 2 1972 Aug Pigs   New South Wales 115.
 3 1972 Sep Pigs   New South Wales 110.
 4 1972 Oct Pigs   New South Wales 108.
 5 1972 Nov Pigs   New South Wales 122.
 6 1972 Dec Pigs   New South Wales 107.
 7 1973 Jan Pigs   New South Wales  96.6
 8 1973 Feb Pigs   New South Wales  96.7
 9 1973 Mar Pigs   New South Wales 121.
10 1973 Apr Pigs   New South Wales  99.3
# … with 476 more rows
>
```

```
150  ##### PART B #####
151
152  test = pigs %>%
153    slice(n()-71:0)
154
155
151:1   PART B

Console  Terminal  Jobs
R 4.1.1 · ~/
 8 1973 Feb Pigs   New South Wales  96.7
 9 1973 Mar Pigs   New South Wales 121.
10 1973 Apr Pigs   New South Wales  99.3
# … with 548 more rows
> pigs %>%
+   slice(n()-71:0)
# A tsibble: 72 x 4 [1M]
# Key:       Animal, State [1]
      Month Animal State           Count
      <mth> <fct>  <fct>           <dbl>
 1 2013 Jan Pigs   New South Wales  71.3
 2 2013 Feb Pigs   New South Wales  69.7
 3 2013 Mar Pigs   New South Wales  79.9
 4 2013 Apr Pigs   New South Wales  74.3
 5 2013 May Pigs   New South Wales  87.2
 6 2013 Jun Pigs   New South Wales  71.2
 7 2013 Jul Pigs   New South Wales  86.2
 8 2013 Aug Pigs   New South Wales  78
 9 2013 Sep Pigs   New South Wales  71.2
10 2013 Oct Pigs   New South Wales  78.2
# … with 62 more rows
>
```

Figure 16: Train test data for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. The train data consists of first 486 lines(1976-2012) while the test data is of last 72 lines (2013-2018).

Part c

Since the data show seasonality, seasonal naive method performs best. Other methods do not demonstrate any specific reactions to seasonality due to their definitions. In Fig-18 (Naive Bootstrapped) and Fig-19 (Prediction Interval), we can see other forecast methods.

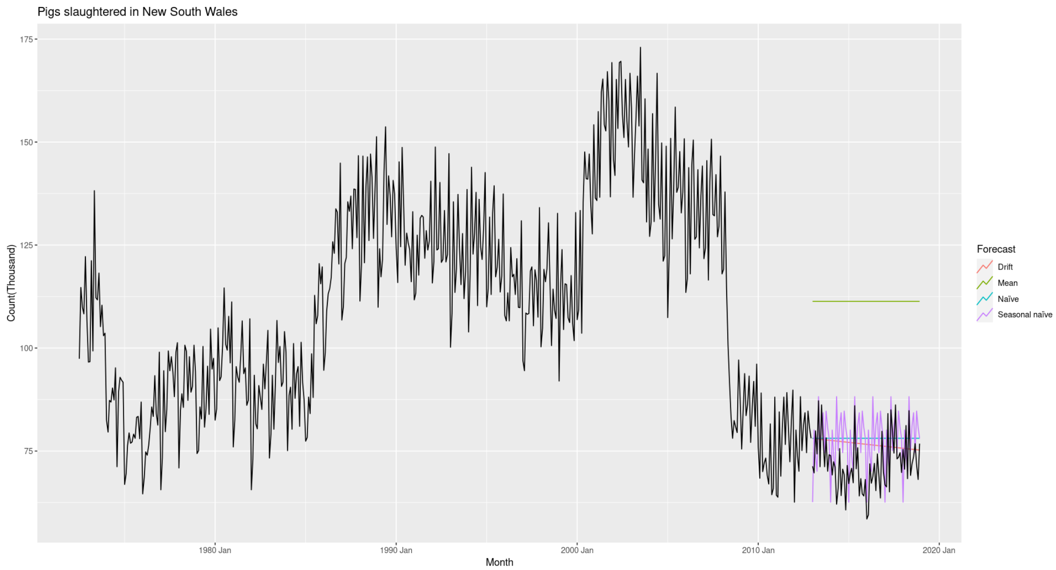


Figure 17: Different benchmark methods (MEAN, NAIVE, SNAIVE, DRIFT) for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. Seasonal naive method gives the best result since it is the only method that sensitive to seasonality in a data set.

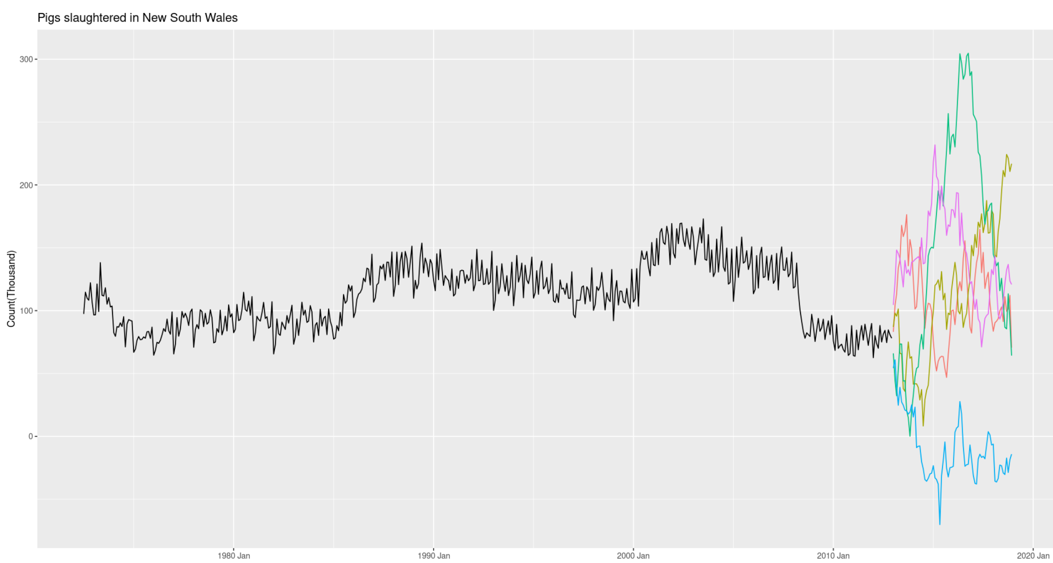


Figure 18: Naive bootstrapped for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. It gives us 5 different forecasting results for the test data.

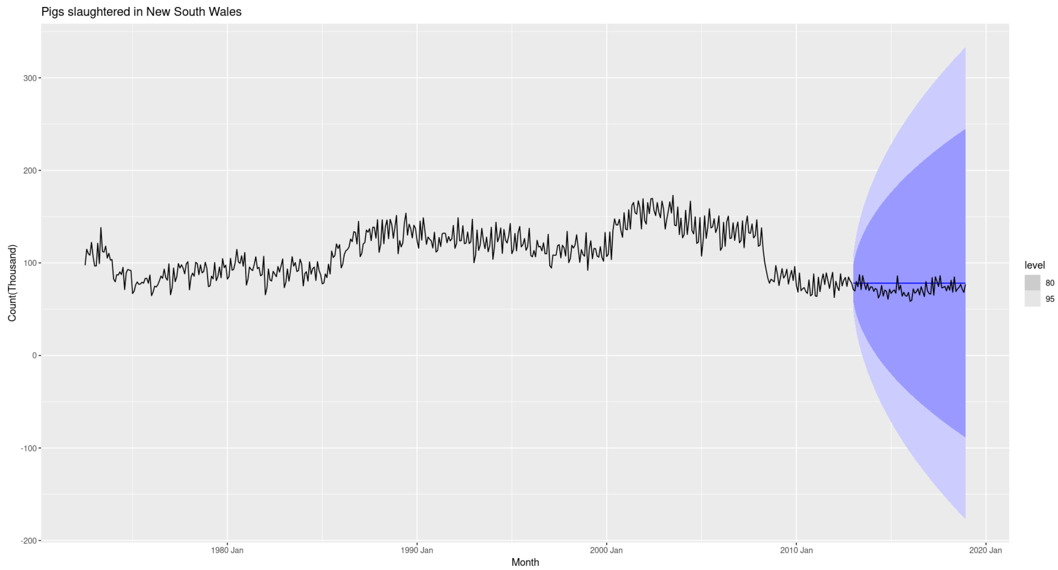


Figure 19: Prediction interval for the number of pigs slaughtered in New South Wales - Australia between years of 1976-2018. The model created by the train data and applied to the test data for 80% and 95% intervals.

Part d

Residuals can be seen in Fig-20 and Fig-21. In first figure, on the bottom left panel, we see there is a seasonality pattern in residual part. There is almost no white noise in the residual indicates that there is still useful information in the residual. In Fig-21, beside seasonality same in Fig-20, there is also an additional trend pattern in the residual.

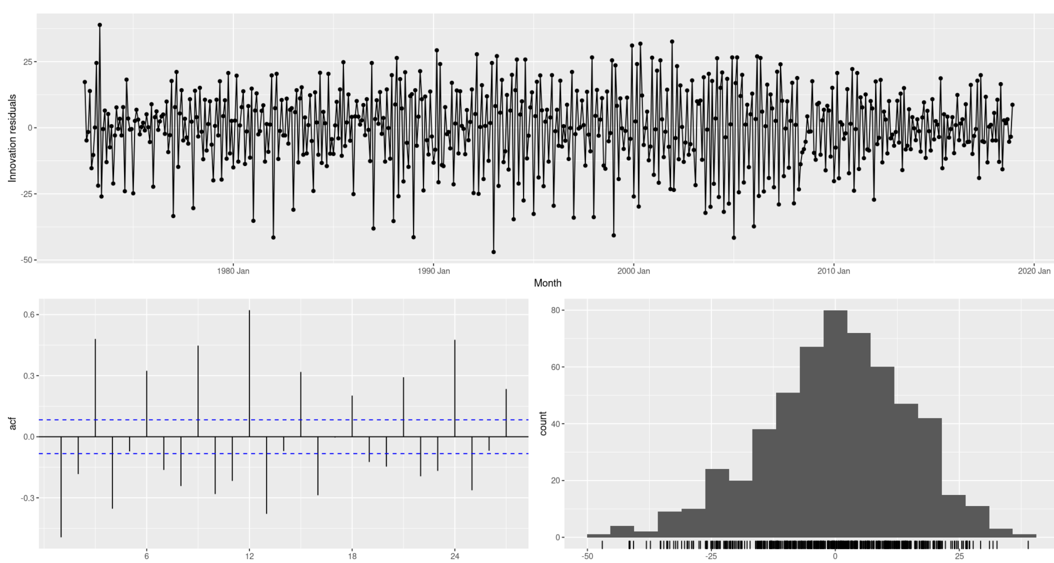


Figure 20: Time series residuals for Naive method. Note that residual still has some seasonality that cannot be considered as white noise (left bottom).

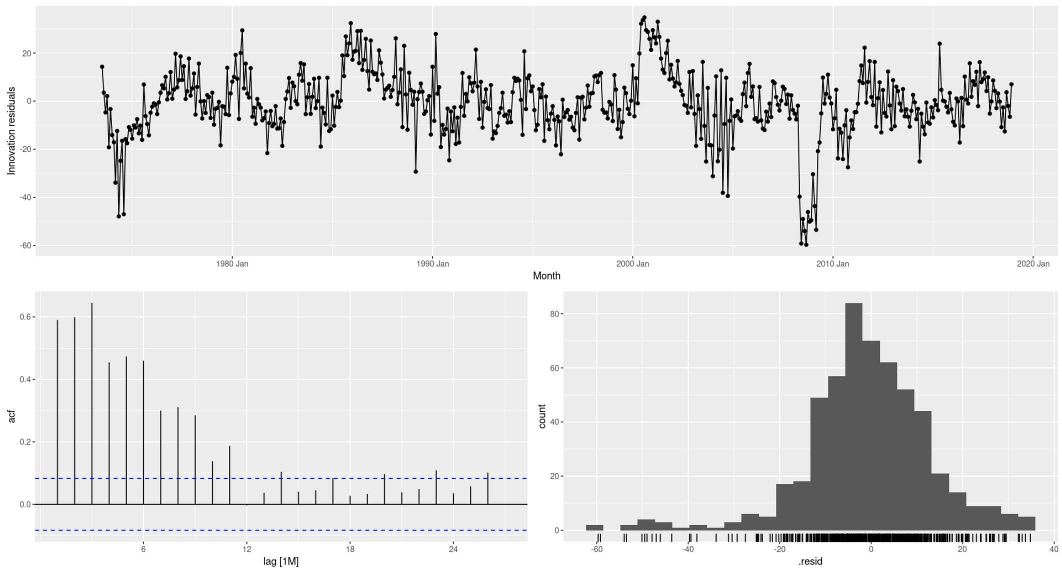


Figure 21: Time series residuals for Seasonal Naive method. Addition to the left bottom panel of Fig-20, residual has some part of trend in it (left bottom).

Problem 2

Part a

```
> ##### PART A ######
> AUS_test = AUS %>%slice_tail(n=4)
> AUS_test
# A tsibble: 4 x 8 [1Y]
# Key:       Country [1]
  Country    Year  Debt    DI Expenditure Savings Wealth Unemployment
  <chr>     <dbl> <dbl> <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
1 Australia  2013  190.  2.64        2.44    7.26   361.         5.66
2 Australia  2014  193.  2.61        2.39    7.47   375.         6.08
3 Australia  2015  202.  1.34        2.79    5.64   393.         6.06
4 Australia  2016  209.  1.90        2.14    4.62   422.         5.71
> AUS_train = AUS %>%slice_head(n=18)
> AUS_train
# A tsibble: 18 x 8 [1Y]
# Key:       Country [1]
   Country    Year  Debt     DI Expenditure Savings Wealth Unemployment
   <chr>     <dbl> <dbl>  <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
 1 Australia  1995  95.7 3.72          3.40   5.24    315.         8.47
 2 Australia  1996  99.5 3.98          2.97   6.47    315.         8.51
 3 Australia  1997 108.  2.52          4.95   3.74    323.         8.36
 4 Australia  1998 115.  4.02          5.73   1.29    339.         7.68
 5 Australia  1999 121.  3.84          4.26   0.638   354.         6.87
 6 Australia  2000 126.  3.77          3.18   1.99    350.         6.29
 7 Australia  2001 132.  4.36          3.10   3.24    348.         6.74
 8 Australia  2002 149.  0.0218        4.03  -1.15    349.         6.37
 9 Australia  2003 159.  6.06          5.04  -0.413   360.         5.93
10 Australia  2004 170.  5.53          4.54   0.657   379.         5.40
11 Australia  2005 181.  2.35          3.25  -0.275   401.         5.03
12 Australia  2006 186.  5.17          5.44   0.129   407.         4.79
13 Australia  2007 190.  6.35          4.92   1.90    401.         4.38
14 Australia  2008 184.  7.24          0.570  8.27    328.         4.23
15 Australia  2009 190.  1.05          3.25   6.14    356.         5.56
16 Australia  2010 188.  5.55          3.99   7.43    342.         5.21
17 Australia  2011 188.  3.03          2.97   8.06    318.         5.08
18 Australia  2012 189.  0.340         1.77   6.84    353.         5.22
```

Figure 22: Test and train data for Household budget in Aussieland

```
> CAN_test = CAN %>%slice_tail(n=4)
> CAN_test
# A tsibble: 4 x 8 [1Y]
# Key:       Country [1]
  Country  Year  Debt     DI Expenditure Savings Wealth Unemployment
  <chr>   <dbl> <dbl>  <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
1 Canada   2013  169.  2.19         2.58    5.26   504.         7.08
2 Canada   2014  172.  1.13         2.62    3.57   540.         6.91
3 Canada   2015  173.  2.95         2.16    4.46   536.         6.91
4 Canada   2016  182. -0.134        2.29    3.28   565.         6.99
> CAN_train = CAN %>%slice_head(n=18)
> CAN_train
# A tsibble: 18 x 8 [1Y]
# Key:       Country [1]
   Country  Year  Debt    DI Expenditure Savings Wealth Unemployment
   <chr>   <dbl> <dbl> <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
 1 Canada   1995  103. 0.641      2.24      9.31   402.         9.51
 2 Canada   1996  106. 0.215      2.96      6.79   441.         9.61
 3 Canada   1997  109. 2.11       4.90      4.65   455.         9.12
 4 Canada   1998  112. 3.09       2.74      4.95   449.         8.28
 5 Canada   1999  114. 3.39       3.94      4.75   458.         7.59
 6 Canada   2000  114. 3.45       4.06      5.30   449.         6.82
 7 Canada   2001  115. 3.79       2.44      4.86   419.         7.22
 8 Canada   2002  119. 2.14       3.99      3.00   401.         7.68
 9 Canada   2003  125. 1.53       2.75      1.98   421.         7.58
10 Canada   2004  131. 3.31       3.04      2.57   429.         7.19
11 Canada   2005  140. 1.85       3.93      1.47   462.         6.76
12 Canada   2006  144. 5.29       4.17      2.60   477.         6.33
13 Canada   2007  152. 3.44       4.39      2.15   475.         6.05
14 Canada   2008  157. 3.84       2.96      3.35   420.         6.14
15 Canada   2009  165. 2.05       0.0304    4.66   448.         8.35
16 Canada   2010  166. 3.39       3.57      4.39   467.         8.06
17 Canada   2011  169. 1.56       2.29      4.27   468.         7.52
18 Canada   2012  170. 2.19       1.91      4.82   490.         7.29
```

Figure 23: Test and train data for Household budget in Canada

```
> JAP_test = JAP %>%slice_tail(n=4)
> JAP_test
# A tsibble: 4 x 8 [1Y]
# Key:       Country [1]
  Country  Year  Debt     DI Expenditure Savings Wealth Unemployment
  <chr>   <dbl> <dbl>  <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
1 Japan    2013  104.  0.174      2.37     0.302   582.         4.03
2 Japan    2014  105. -1.19      -0.857   -0.347   610.         3.59
3 Japan    2015  105.  1.45      -0.0324   0.815   596.         3.38
4 Japan    2016  106.  1.77       0.0604   2.56    602.         3.12
> JAP_train = JAP %>%slice_head(n=18)
> JAP_train
# A tsibble: 18 x 8 [1Y]
# Key:       Country [1]
   Country  Year  Debt     DI Expenditure Savings Wealth Unemployment
   <chr>   <dbl> <dbl>  <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
 1 Japan    1995  110.  2.15        2.50     12.3    414.         3.15
 2 Japan    1996  114.  0.611       2.09     11.0    424.         3.35
 3 Japan    1997  114.  1.08        0.693    10.9    426.         3.4
 4 Japan    1998  112.  0.667      -0.587    12.1    434.         4.11
 5 Japan    1999  114.  0.122       1.16     10.5    460.         4.68
 6 Japan    2000  115.  0.436       1.58      8.85   472.         4.72
 7 Japan    2001  117. -1.19        1.90      5.56   487.         5.03
 8 Japan    2002  115.  1.16        1.18      4.91   491.         5.38
 9 Japan    2003  115.  0.0160      0.659     4.40   511.         5.26
10 Japan    2004  113.  1.25        1.31      3.56   518.         4.72
11 Japan    2005  111.  0.927       1.24      3.45   551.         4.42
12 Japan    2006  108.  0.233       1.03      2.51   565.         4.14
13 Japan    2007  108.  1.23        0.936     2.54   526.         3.84
14 Japan    2008  109. -0.750      -1.02      2.52   509.         3.99
15 Japan    2009  107.  1.44       -0.706     4.00   529.         5.07
16 Japan    2010  107.  2.10        2.40      3.71   527.         5.05
17 Japan    2011  103.  0.421      -0.386     3.97   540.         4.58
18 Japan    2012  103.  1.19        2.03      2.71   561.         4.35
```

Figure 24: Test and train data for Household budget in Japan

```
> USA_test = USA %>%slice_tail(n=4)
> USA_test
# A tsibble: 4 x 8 [1Y]
# Key:       Country [1]
  Country  Year  Debt    DI Expenditure Savings Wealth Unemployment
  <chr>   <dbl> <dbl> <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
1 USA      2013  116. -1.47        1.45    6.61   592.         7.38
2 USA      2014  112.  3.42        2.93    7.58   596.         6.17
3 USA      2015  109.  4.07        3.72    7.83   588.         5.29
4 USA      2016  109.  1.63        2.74    6.91   609.         4.87
> USA_train = USA %>%slice_head(n=18)
> USA_train
# A tsibble: 18 x 8 [1Y]
# Key:       Country [1]
   Country  Year  Debt      DI Expenditure Savings Wealth Unemployment
   <chr>   <dbl> <dbl>   <dbl>       <dbl>   <dbl>  <dbl>        <dbl>
 1 USA      1995  94.3  3.00         2.95     7.24   472.         5.61
 2 USA      1996  96.2  2.86         3.47     6.79   485.         5.42
 3 USA      1997  97.3  3.46         3.77     6.58   510.         4.95
 4 USA      1998  98.5  5.66         5.31     7.06   527.         4.51
 5 USA      1999 103.   3.20         5.27     5.27   563.         4.22
 6 USA      2000 104.   4.69         5.08     5.03   521.         3.99
 7 USA      2001 108.   2.82         2.52     5.25   499.         4.73
 8 USA      2002 113.   3.29         2.57     6.06   461.         5.78
 9 USA      2003 121.   2.79         3.18     5.75   492.         5.99
10 USA      2004 128.   3.28         3.75     5.35   525.         5.53
11 USA      2005 136.   1.34         3.56     3.28   550.         5.07
12 USA      2006 141.   3.56         3.06     4.00   563.         4.62
13 USA      2007 144.   1.78         2.22     3.88   561.         4.62
14 USA      2008 137.   1.74        -0.212    5.18   479.         5.78
15 USA      2009 136.  -0.0516      -1.25     6.34   500.         9.27
16 USA      2010 128.   1.01         1.75     6.78   520.         9.62
17 USA      2011 121.   2.23         1.89     7.40   503.         8.95
18 USA      2012 115.   2.96         1.50     9.14   514.         8.07
```

Figure 25: Test and train data for Household budget in USA

Part b

In Fig-26-29, train and test data are used for forecasting with different benchmark methods (MEAN, NAIVE, DRIFT).

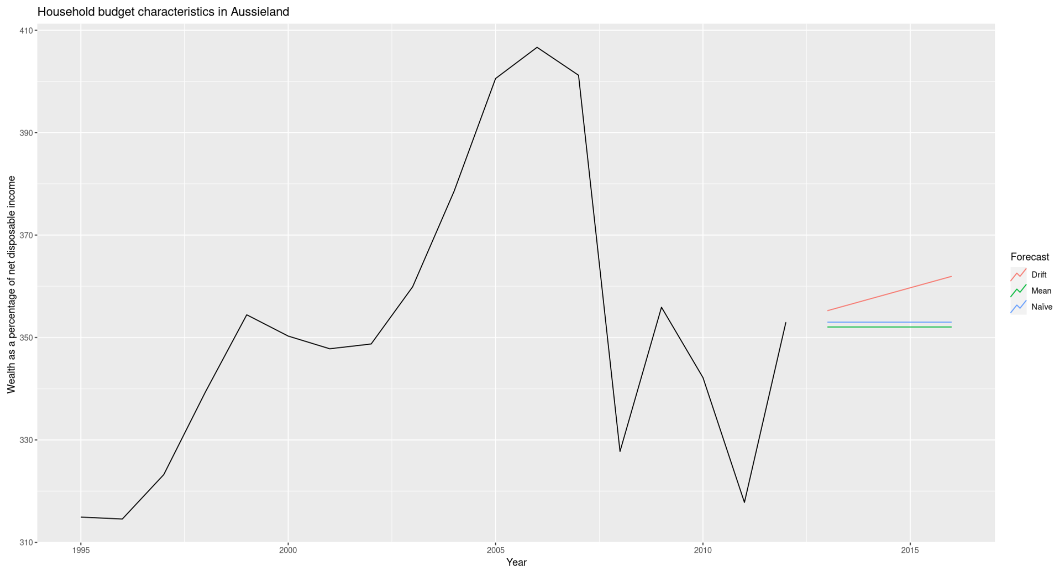


Figure 26: Benchmark methods (MEAN, NAIVE, DRIFT) for forecasting on Household budget in Aussieland.

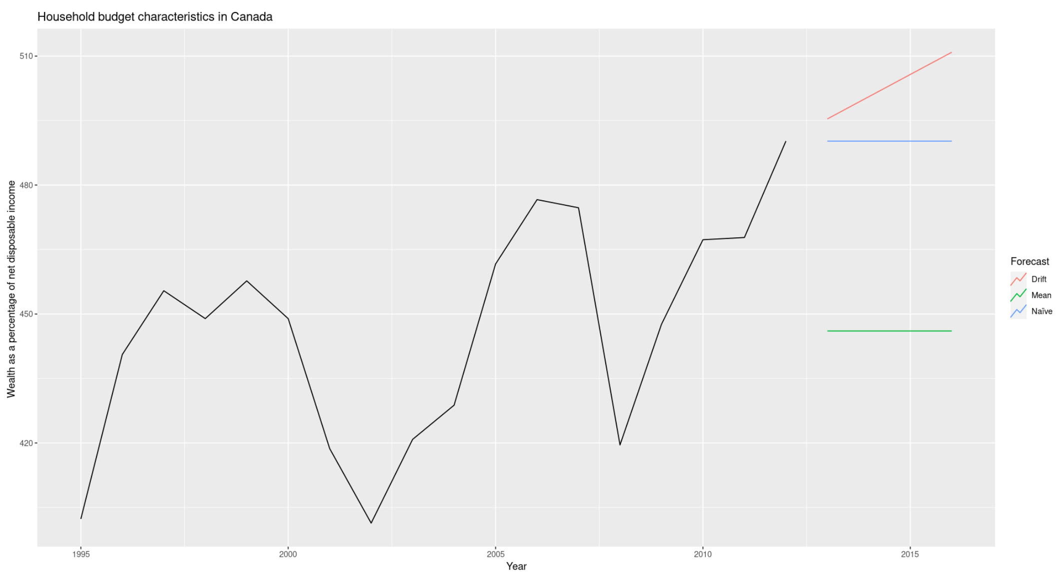


Figure 27: Benchmark methods (MEAN, NAIVE, DRIFT) for forecasting on Household budget in Canada.

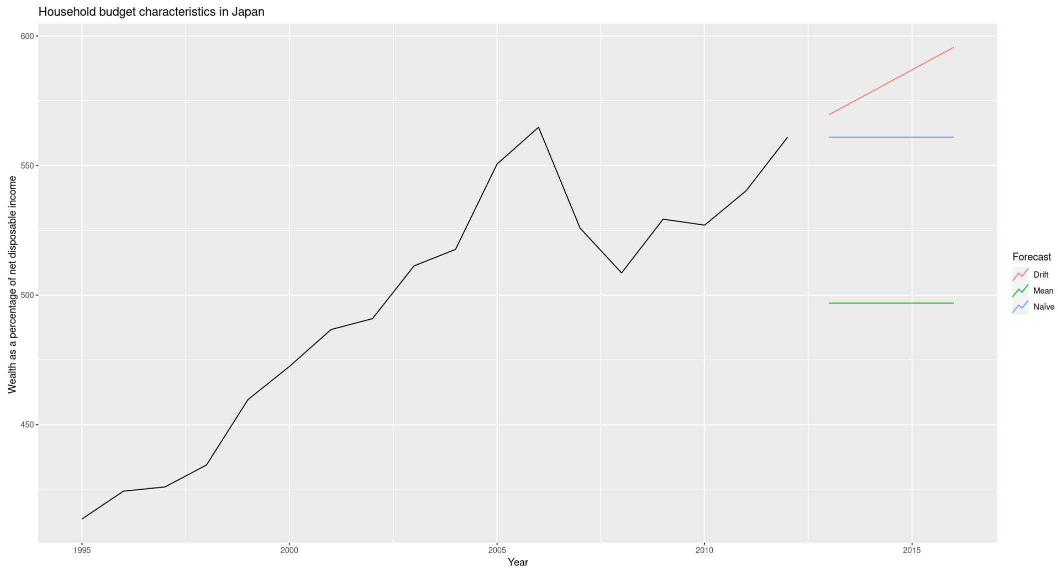


Figure 28: Benchmark methods (MEAN, NAIVE, DRIFT) for forecasting on Household budget in Japan.

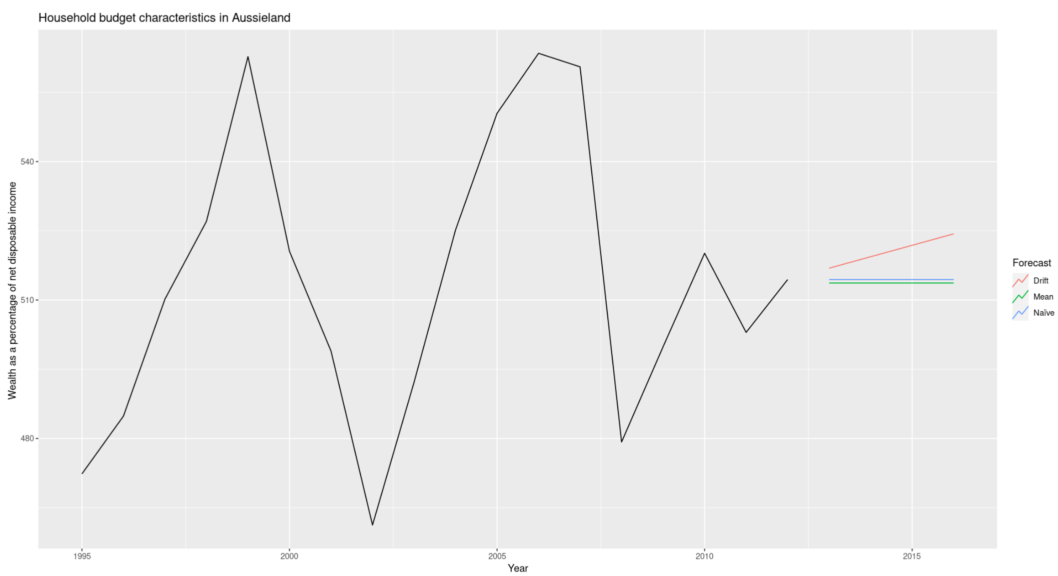


Figure 29: Benchmark methods (MEAN, NAIVE, DRIFT) for forecasting on Household budget in USA.

Part c

According to the accuracy test, DRIFT method provides best results for different countries.

```
> accuracy(AUS_fc, AUS)
# A tibble: 3 × 11
  .model Country   .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
  <chr>  <chr>     <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Drift  Australia Test   29.1  35.5  29.1  7.23  7.23  1.73  1.48 0.210
2 Mean   Australia Test   35.7  42.3  35.7  8.89  8.89  2.12  1.76 0.216
3 Naïve  Australia Test   34.7  41.5  34.7  8.64  8.64  2.06  1.73 0.216
> accuracy(CAN_fc, CAN)
# A tibble: 3 × 11
  .model Country .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE    ACF1
  <chr>  <chr>   <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 Drift  Canada  Test   33.3  37.2  33.3  6.09  6.09  1.73  1.57 -0.229
2 Mean   Canada  Test   90.4  92.9  90.4 16.7  16.7   4.69  3.92 -0.0799
3 Naïve  Canada  Test   46.2  51.0  46.2  8.46  8.46  2.40  2.15 -0.0799
> accuracy(JAP_fc, JAP)
# A tibble: 3 × 11
  .model Country .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE   ACF1
  <chr>  <chr>   <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl>
1 Drift  Japan   Test   14.7  17.9  14.7  2.44  2.44 0.943 0.967 -0.229
2 Mean   Japan   Test  100.  101.  100.  16.8  16.8  6.45  5.46  -0.534
3 Naïve  Japan   Test   36.3  37.8  36.3  6.06  6.06 2.34  2.04  -0.534
> accuracy(USA_fc, USA)
# A tibble: 3 × 11
  .model Country .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE   ACF1
  <chr>  <chr>   <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl>
1 Drift  USA     Test   75.9  76.2  75.9  12.7  12.7  2.88  2.43 -0.561
2 Mean   USA     Test   82.9  83.3  82.9  13.9  13.9  3.15  2.65 -0.423
3 Naïve  USA     Test   82.1  82.5  82.1  13.8  13.8  3.12  2.63 -0.423
```

Figure 30: Accuracy tests for Household budget in Aussieland, Canada, Japan and USA.

Part d

In Fig-31-34, time series residuals applied for different countries. In every figures, residuals are purely white noise and there are no data out of the pre-determined limits.

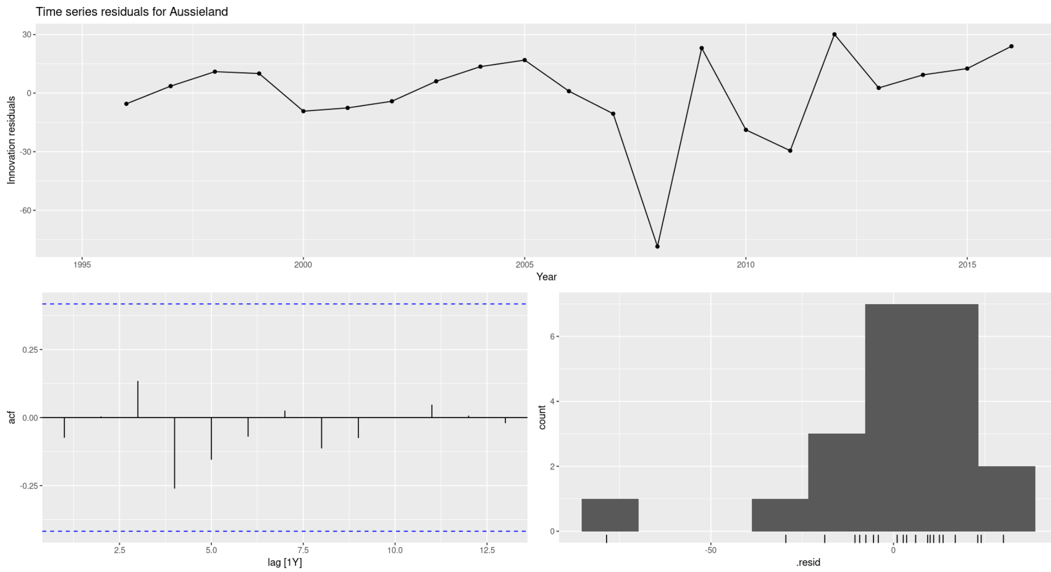


Figure 31: Drift residuals for Household budget in Aussieland.

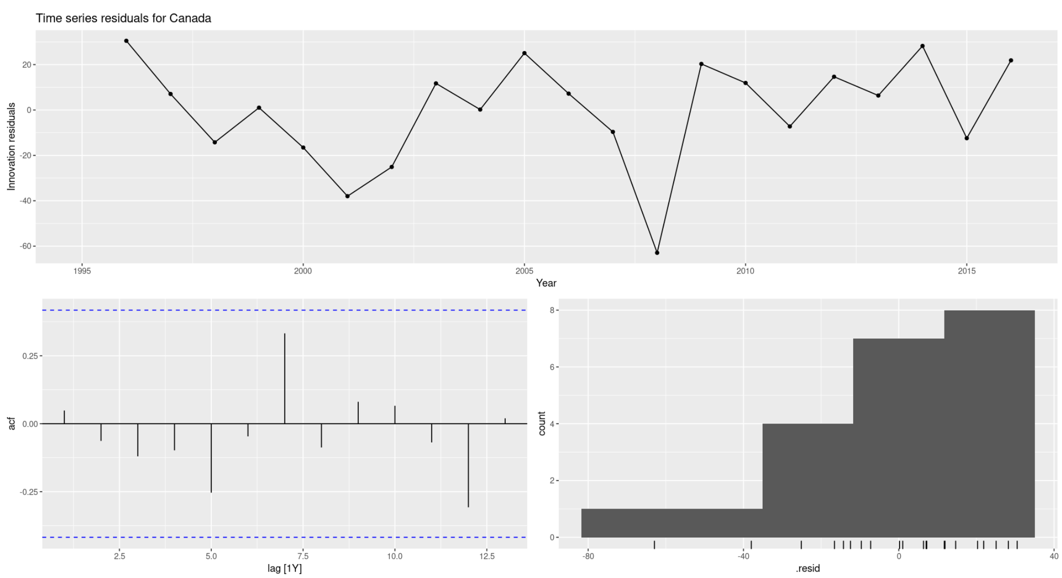


Figure 32: Drift residuals for Household budget in Canada.

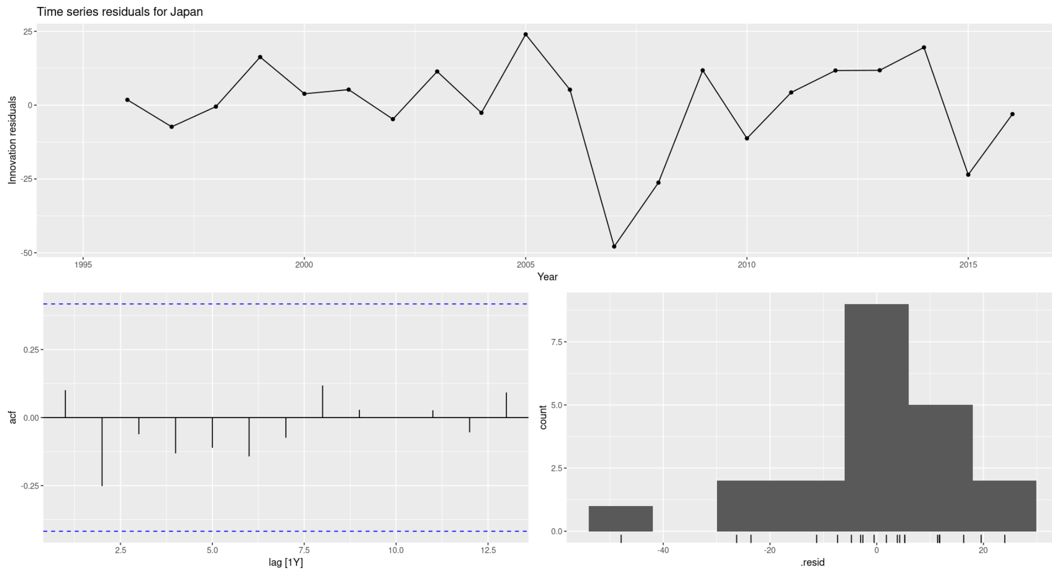


Figure 33: Drift residuals for Household budget in Japan.

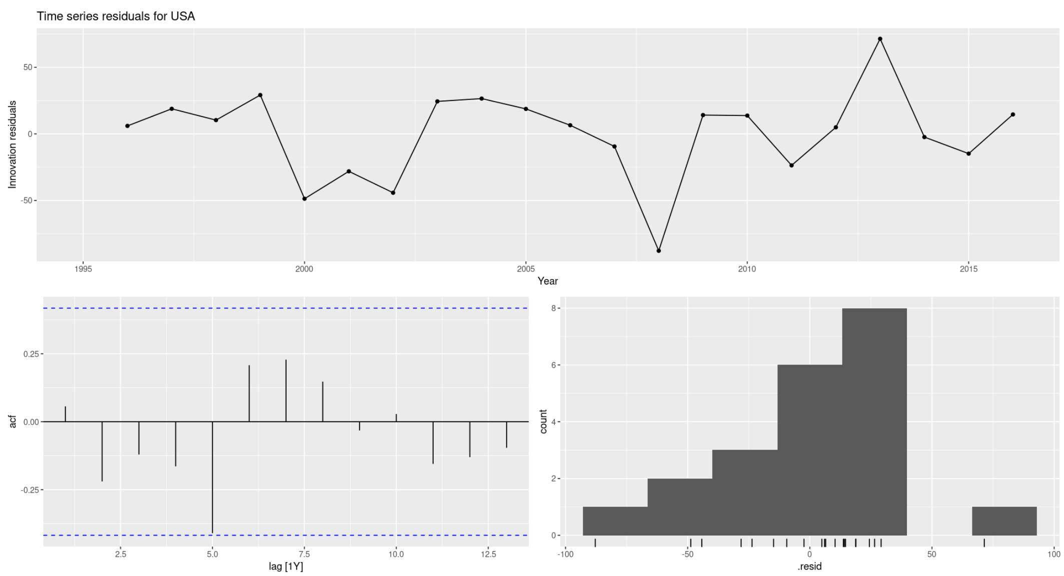


Figure 34: Drift residuals for Household budget in USA.

Problem 3

In Fig35-40, STL decomposition applied for different seasonalities (1,10,25,50,100 and periodic).

Part a

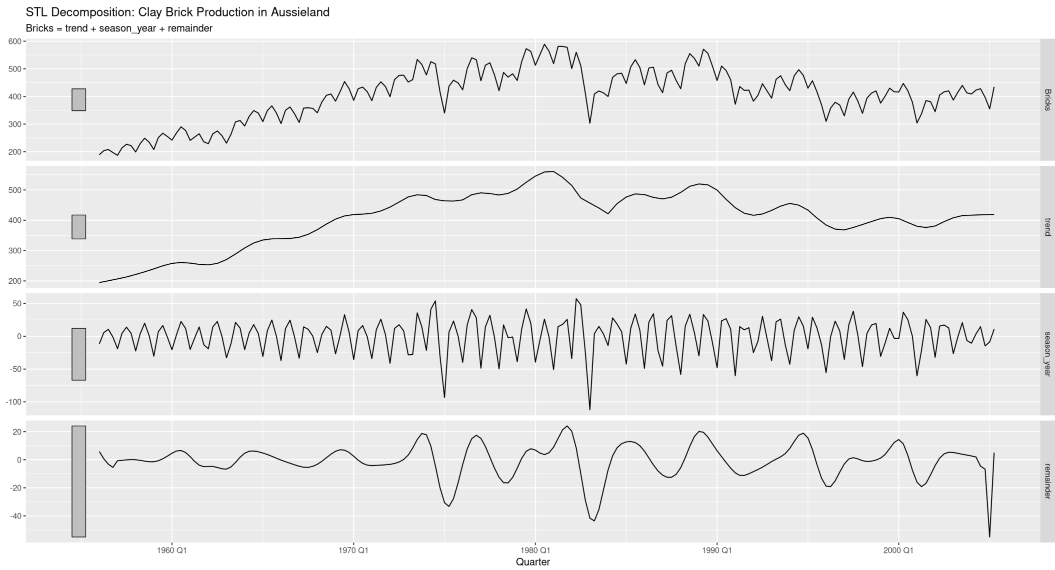


Figure 35: STL decomposition plot (Season window = 1) for Australian quarterly clay brick production between 1956–2005.

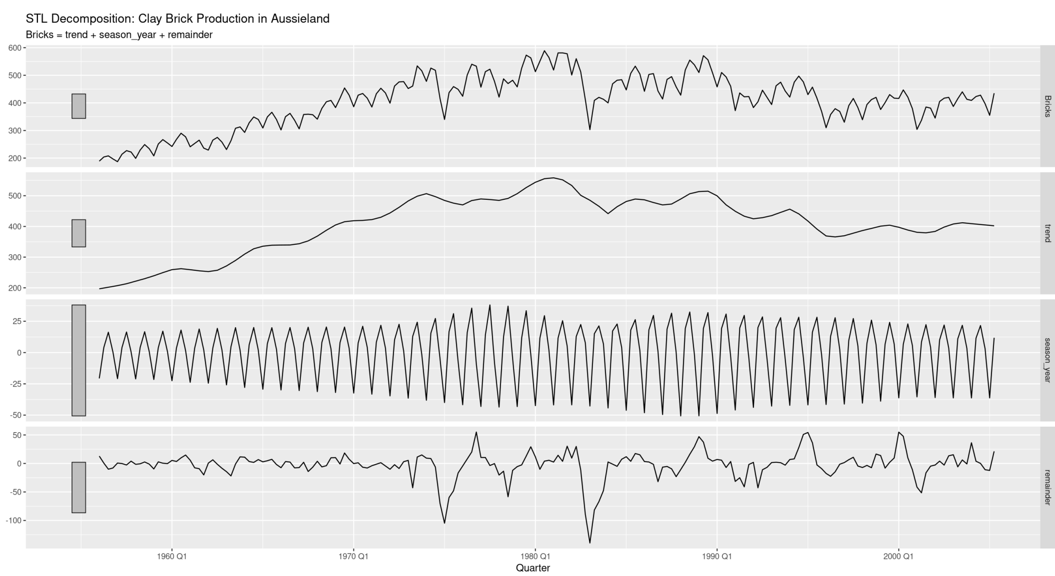


Figure 36: STL decomposition plot (Season window = 10) for Australian quarterly clay brick production between 1956–2005.

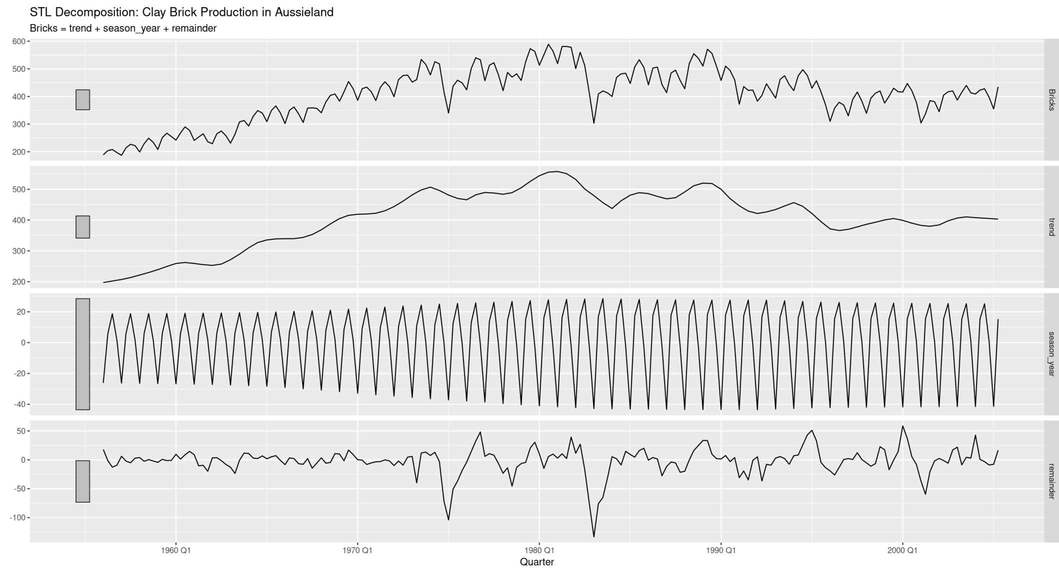


Figure 37: STL decomposition plot (Season window = 25) for Australian quarterly clay brick production between 1956–2005.

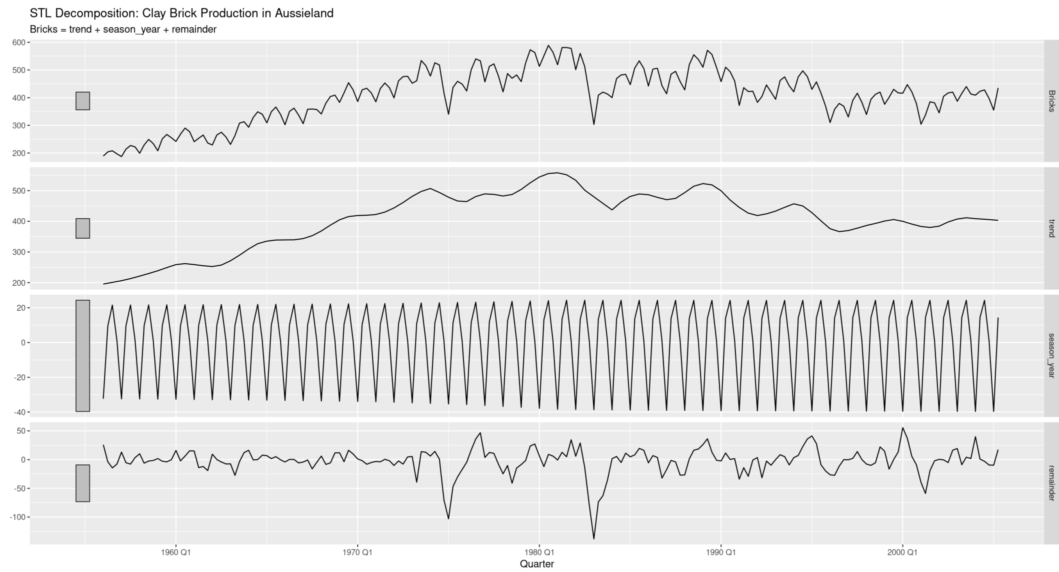


Figure 38: STL decomposition plot (Season window = 50) for Australian quarterly clay brick production between 1956–2005.

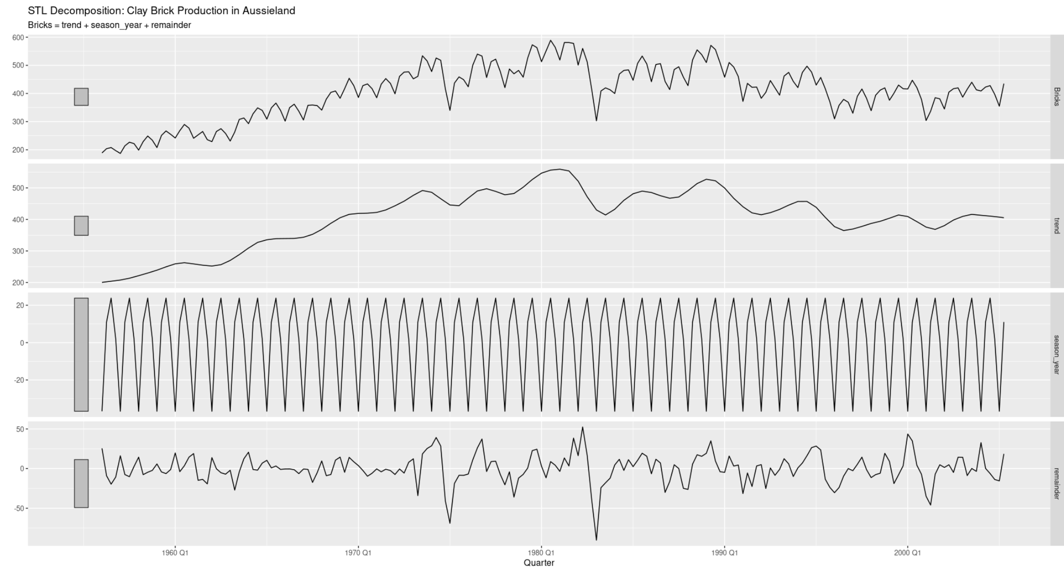


Figure 39: STL decomposition plot (Season window = ”periodic”) for Australian quarterly clay brick production between 1956–2005.

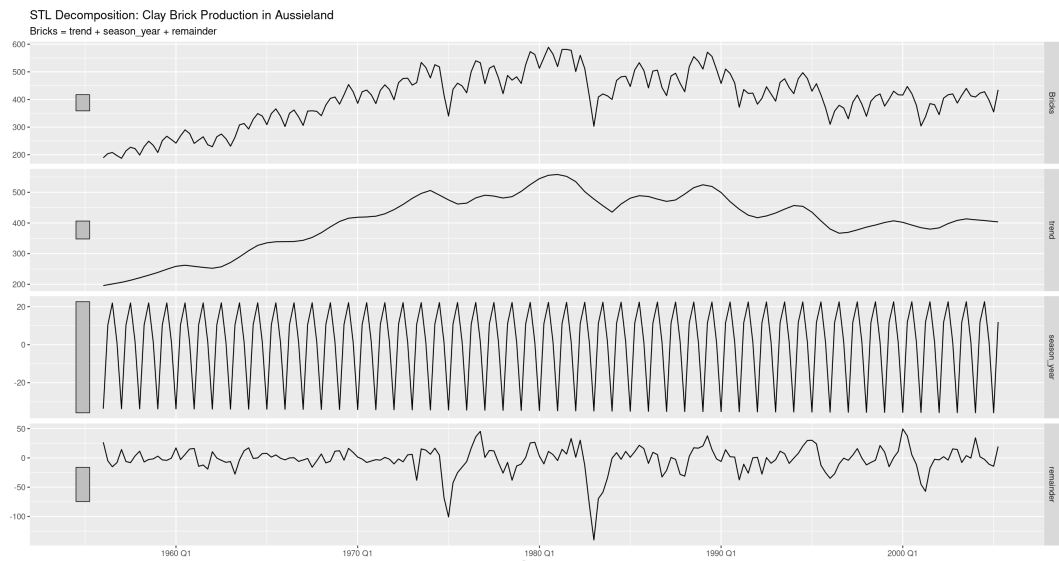


Figure 40: STL decomposition plot (Season window = 100) for Australian quarterly clay brick production between 1956–2005.

0.1 Part b

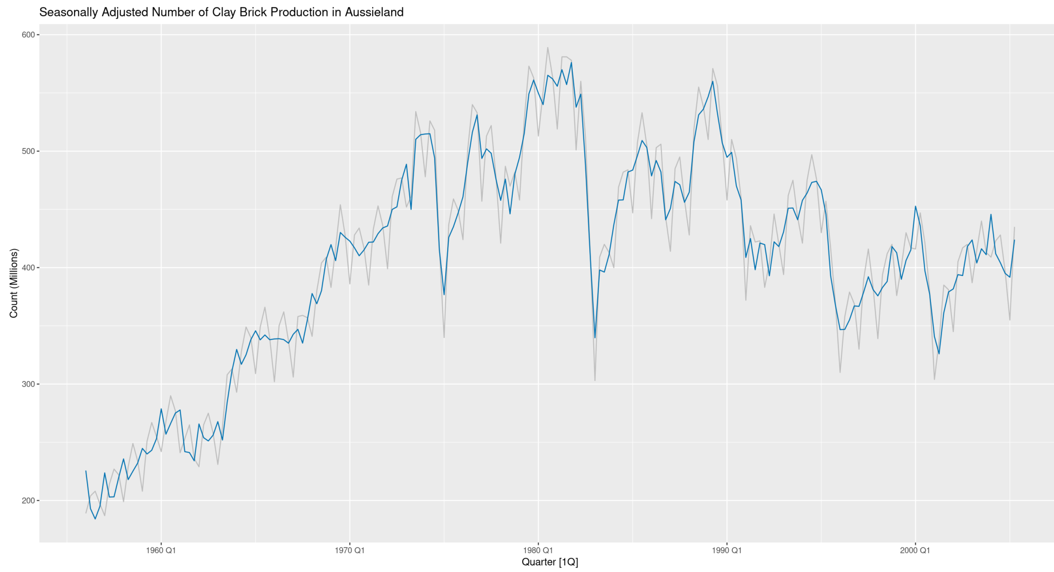


Figure 41: Seasonally adjusted plot for Australian quarterly clay brick production between 1956–2005.

0.2 Part c

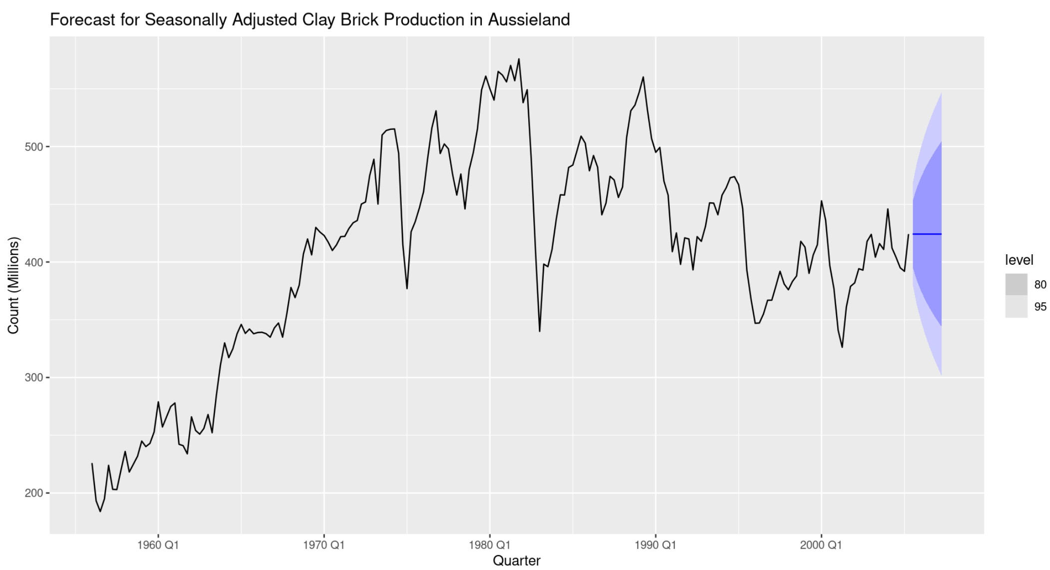


Figure 42: Naive method forecasting on seasonally adjusted data for Australian quarterly clay brick production between 1956–2005.

0.3 Part d

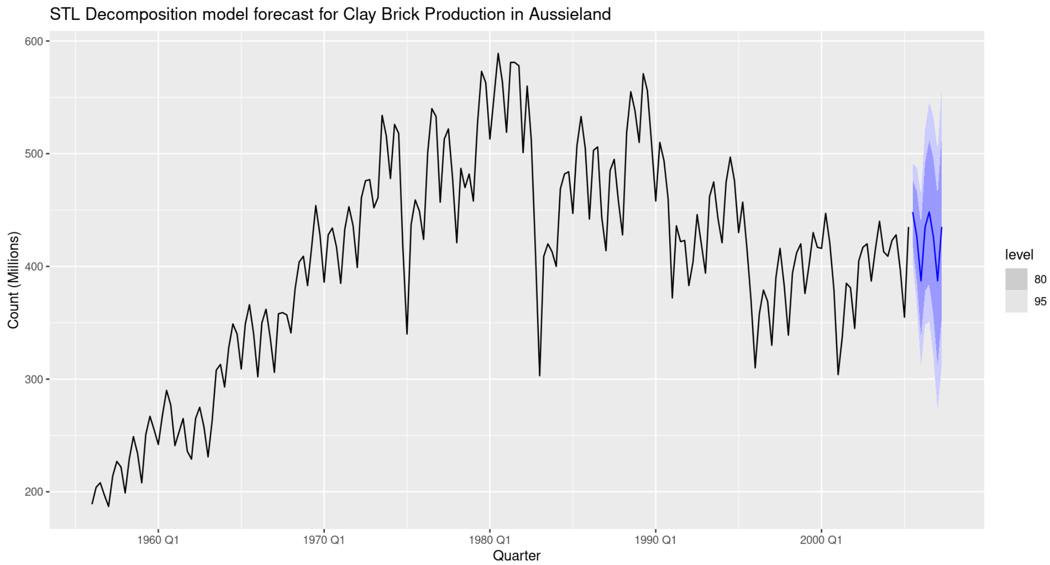


Figure 43: STL decomposition model forecasting on seasonally adjusted data for Australian quarterly clay brick production between 1956–2005.

0.4 Part e

Lags 1,8,21 in Fig-44 and lags 2,12,16,20 in Fig-45 do not look like white noise.

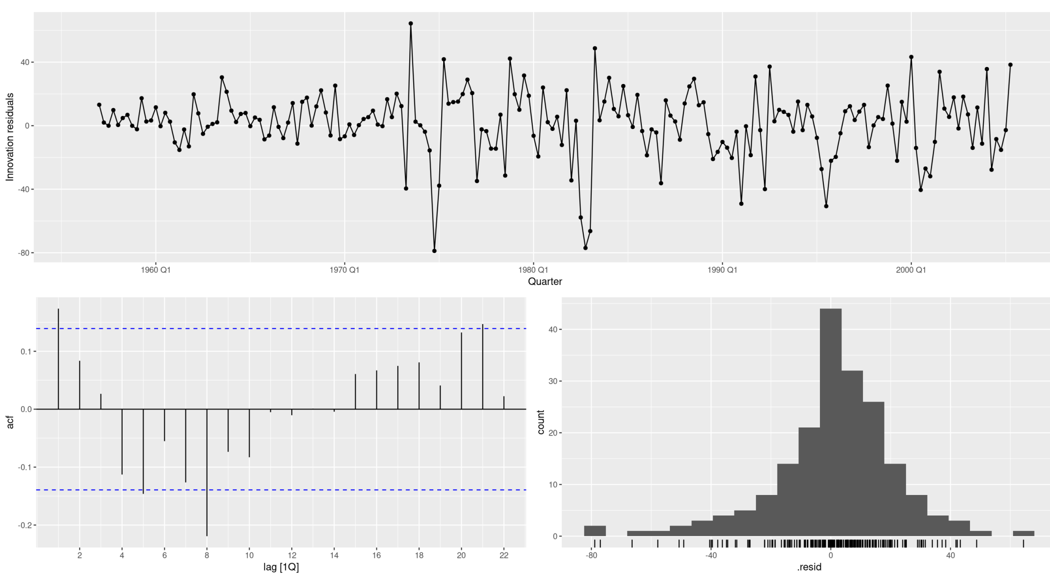


Figure 44: Residuals for decomposition model forecasting on seasonally adjusted data

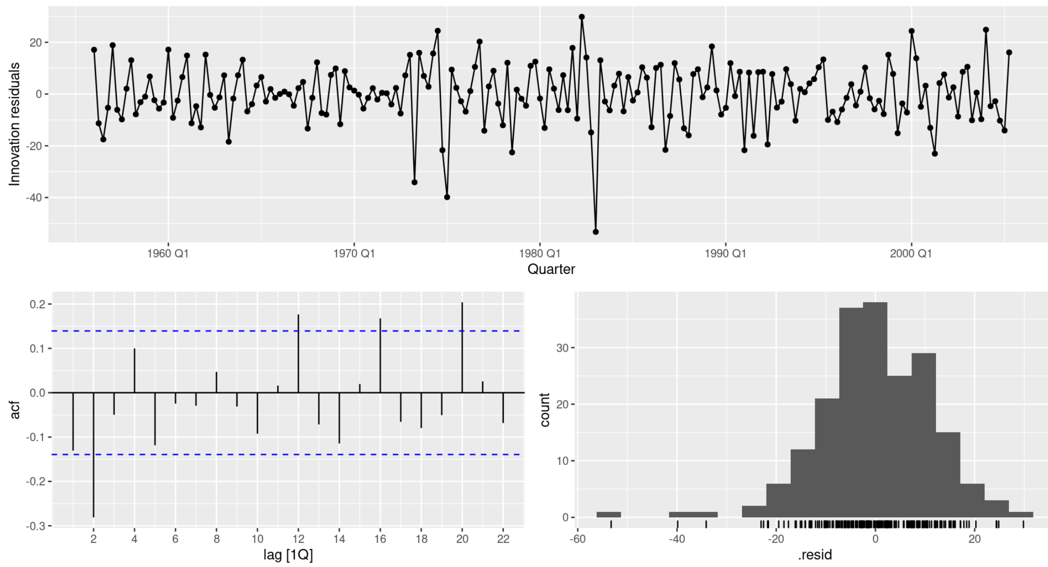


Figure 45: Residuals for STL decomposition forecasting on seasonally adjusted data

Part f

In Fig-46, unlike the previous ones, we used robust method. However, there seems there is no different from non-robust forecastings.

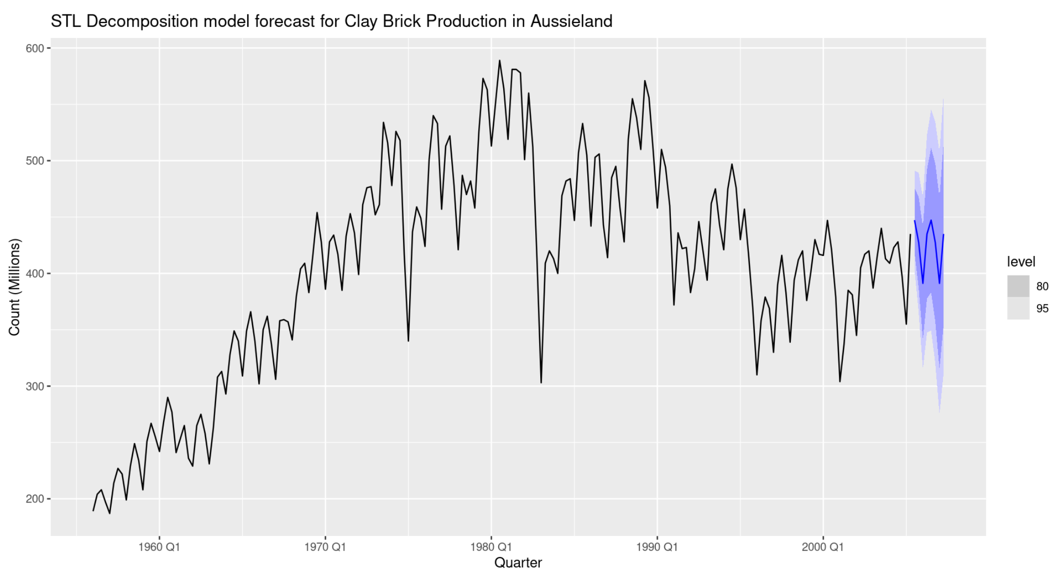


Figure 46: Robust STL decomposition model forecasting on seasonally adjusted data for Australian quarterly clay brick production between 1956–2005.

Part g

In Fig-48 and Fig-50, residual analyses show that decomposition method is better representative for our data. In Fig-50, there is a noticeable seasonal component in it.

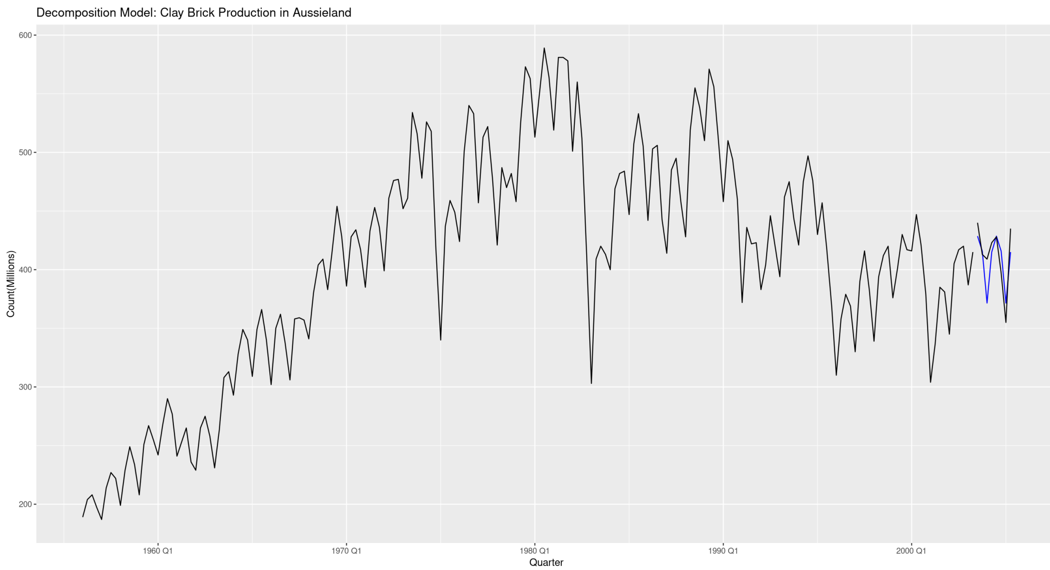


Figure 47: Decomposition model forecasting plot for last 2 years of Australian quarterly clay brick production between 1956–2005.

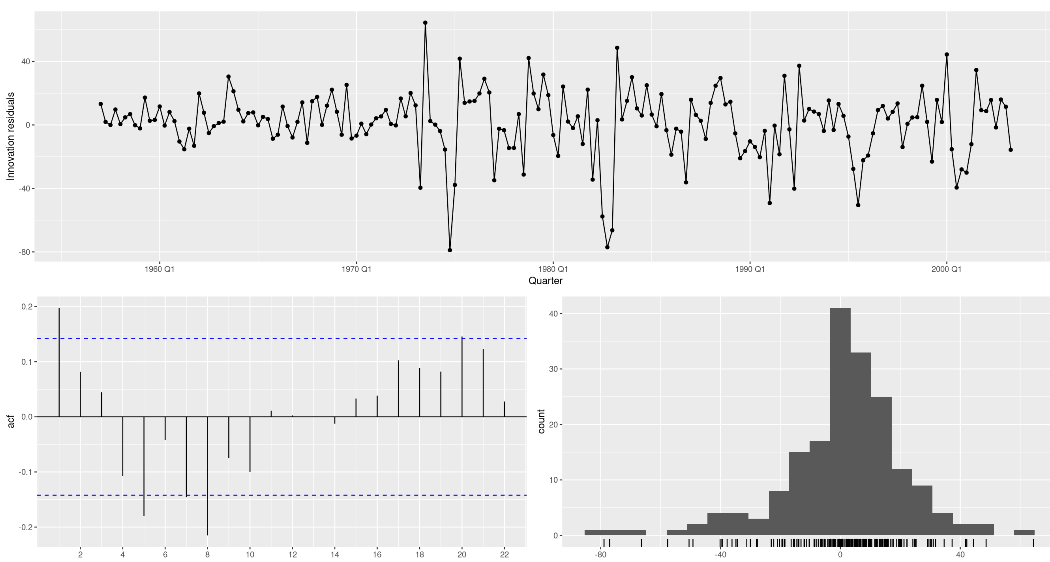


Figure 48: Residuals for Decomposition model forecasting.

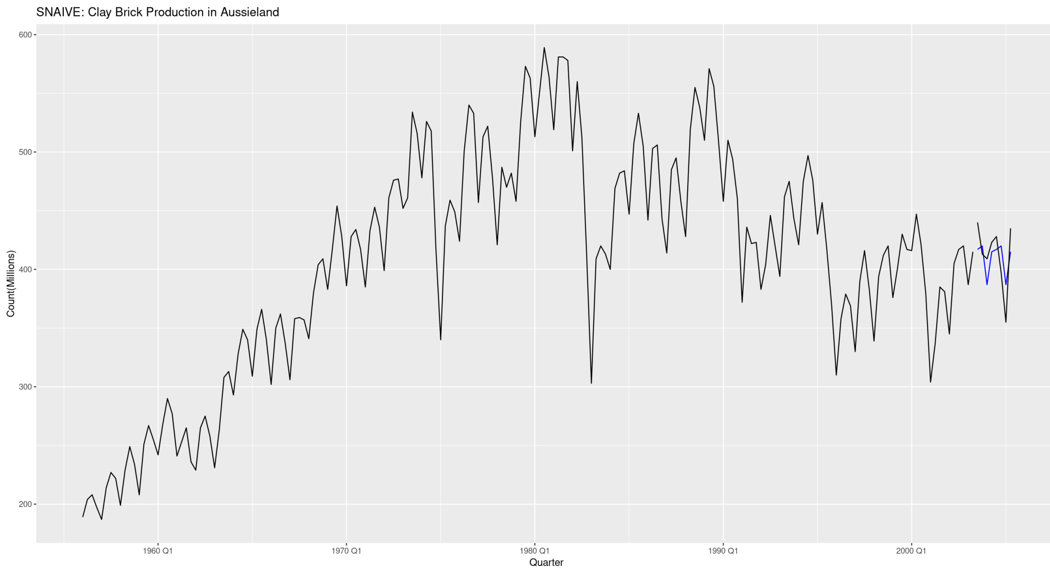


Figure 49: Seasonal naive forecasting plot for last 2 years of Australian quarterly clay brick production between 1956–2005.

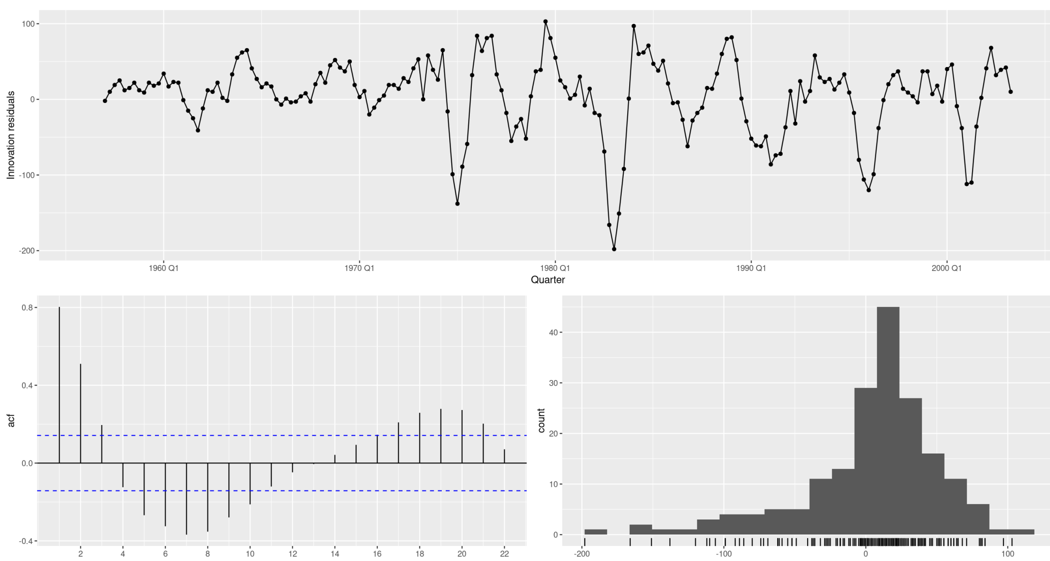


Figure 50: Residuals for Seasonal naive forecasting