MONASH University

Office Use Only

Semester One 2021
Examination Period

Faculty of Business & Economics

EXAM CODES: **ETF3231/ETF5231 - PAPER 1**

TITLE OF PAPER: **Business Forecasting**

EXAM DURATION: 2 hours 40 minutes

READING TIME: 0 minutes

THIS PAPER IS FOR STUDENTS STUDYING AT: (tick where applicable)

✓ Caulfield ☐ Clayton ☐ Parkville ☐ Peninsula
☐ Monash Extension ☐ Off Campus Learning ☐ Malaysia ☐ South Africa
☐ Other (specify)

This is an open book exam. You may consult the online textbook and the materials used during the unit. You may not communicate with any other person during the exam.

The exam contains *five* sections. *All* sections must be answered. The exam is worth 100 marks in total.

SECTION A

Write about a quarter of a page each on any **four** of the following topics. (Clearly state if you agree or disagree with each statement. No marks will be given without any justification.)

1. Narrower prediction intervals are more informative and should always be preferred.
2. The AICc should always be used to select models for forecasting.
3. An ETS model for Holt's linear trend method is a generalisation of an ETS model for simple exponential smoothing. It should therefore always be preferred as it will produce better forecasts.
4. The random walk is particularly suitable for financial non-stationary data. It is a non-stationary model due to its variance.
5. An ARIMA model with uncorrelated residuals will usually produce accurate forecasts.
6. Regression models with Fourier terms should always be used to model seasonality.

Total: 20 marks

— END OF SECTION A —

SECTION B

Figures 1, 2 and 3 relate to the number of births in the state of Victoria, Australia, from January 1980 to September 2019.

1. Using Figures 1, 2 and 3, describe the birth series in Victoria. Carefully comment on the interesting features of all three plots.

5 marks

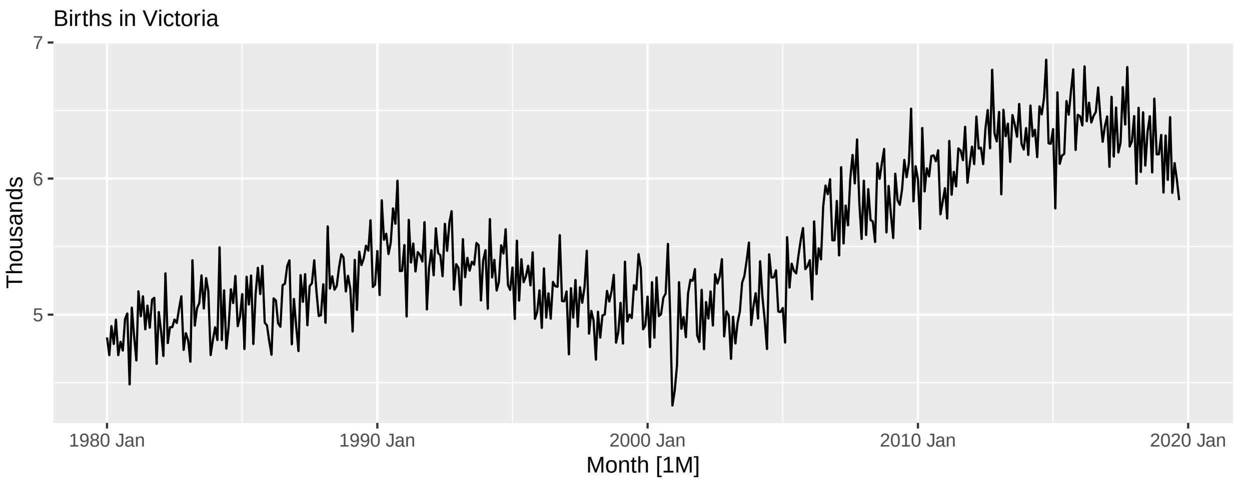


Figure 1:

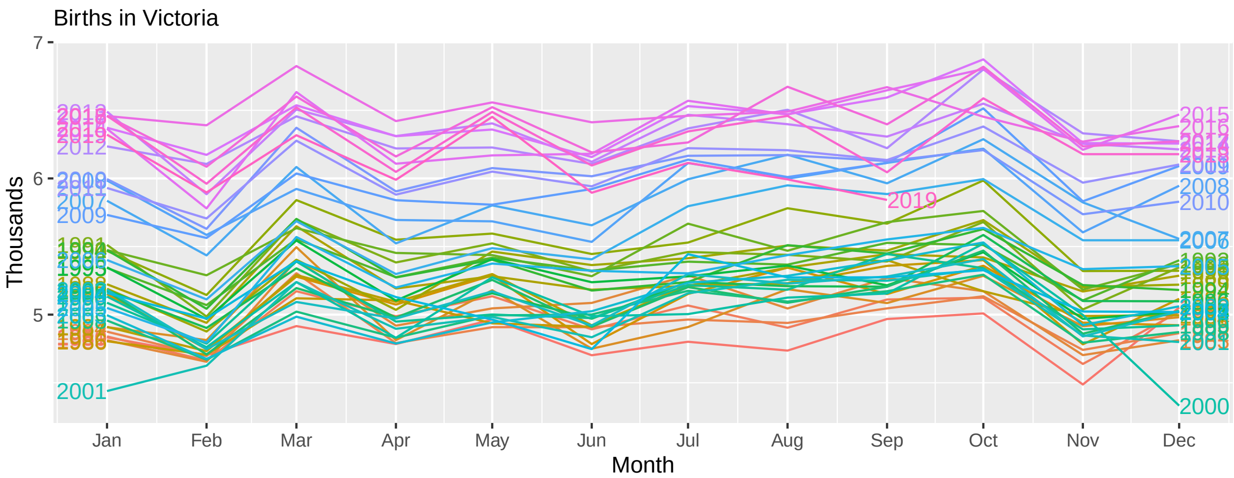
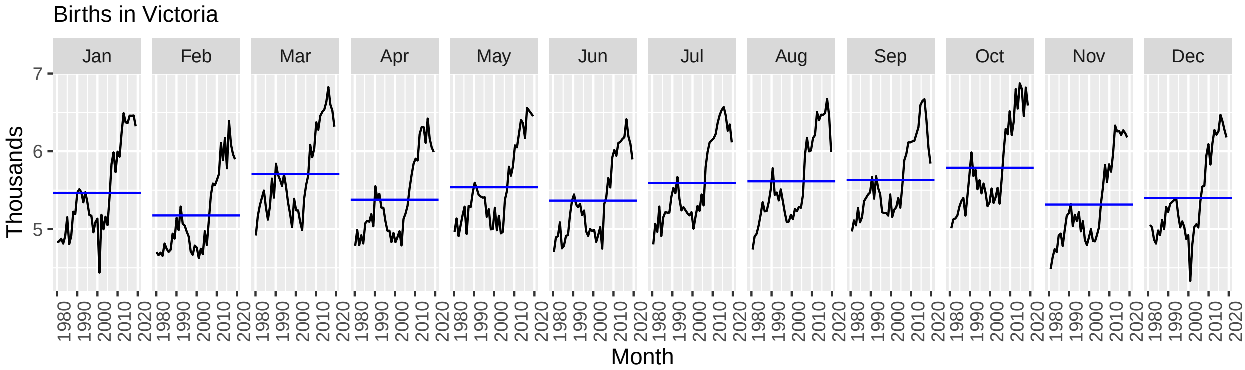


Figure 2:



Births in Victoria
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Thousands
7
6
5
1980
1990
2000
2010
2020
Month

Figure 3:

2. Using the code below, observe what is plotted in all panels of Figures 4 and 5. Are you happy with the settings and the results? Would you change anything?

   5 marks

```
births %>%
  model(STL(count)) %>%
  components() %>%
  autoplot()
```

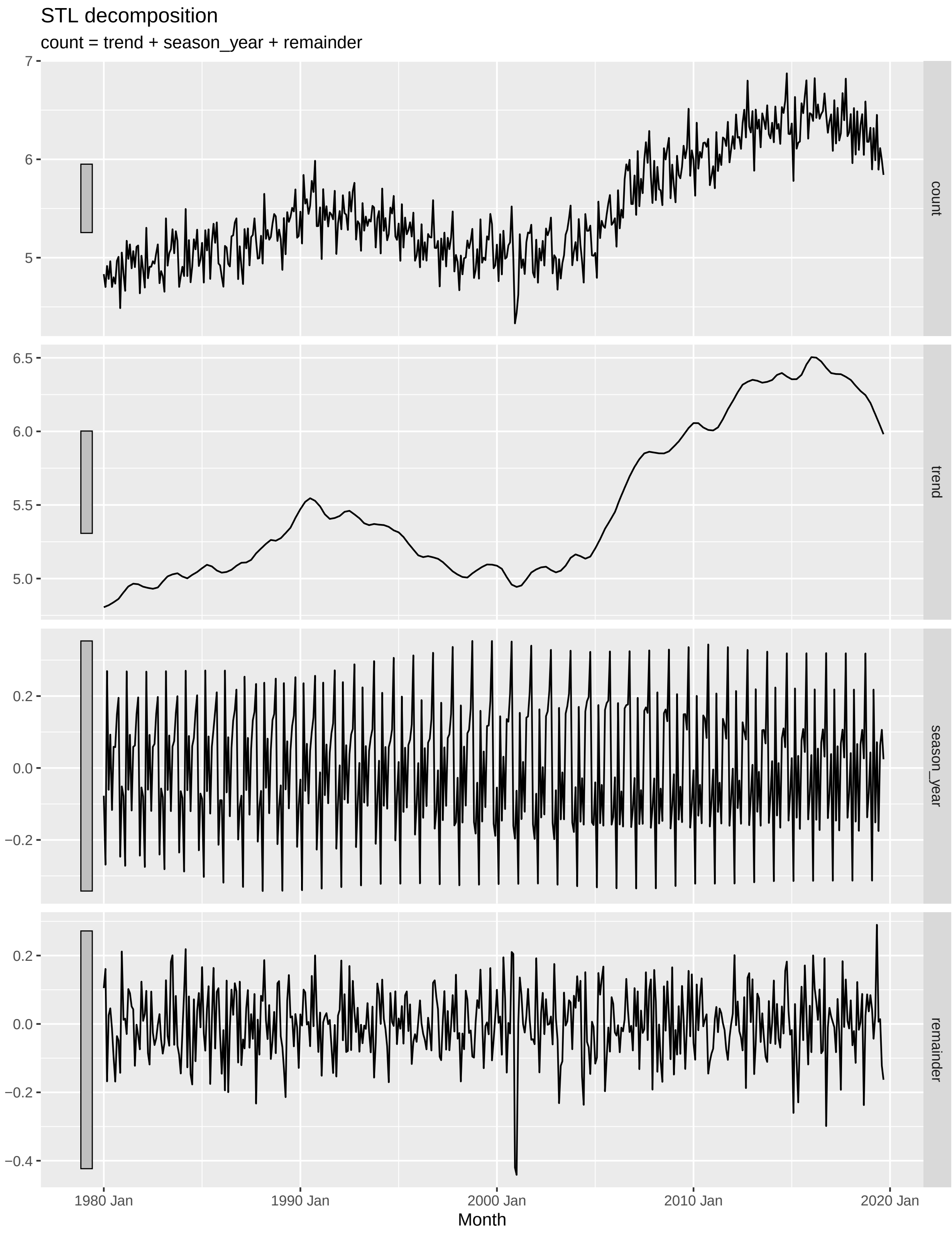


Figure 4:

```
births %>%
  model(decomposition_model(STL(count), RW(season_adjust))) %>%
  forecast(h = "2 years") %>%
  autoplot(births) +
  labs(subtitle = "Births in Victoria: Decomposition forecasts", y = "Thousands")
```

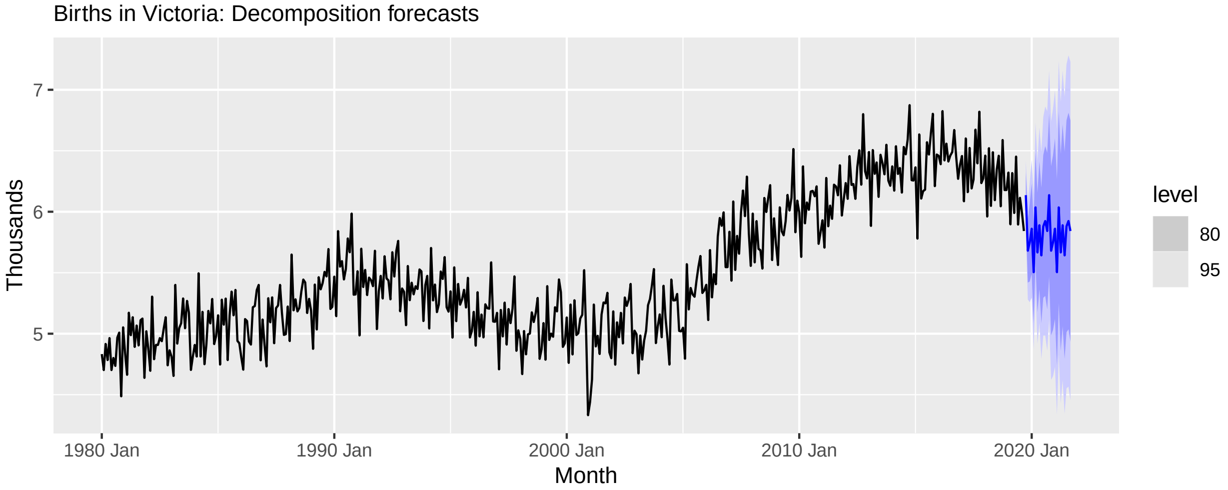


Figure 5:

3. You have been asked to provide forecasts for the next three years for the Victorian births series.

   Consider applying each of the methods and models below. Comment, in a few words each, on whether each one is appropriate for forecasting the data. No marks will be given for simply guessing whether a method or a model is appropriate without justifying your choice.

   Start your response by stating: **suitable** or **not suitable**.

   (a) Seasonal naïve method.

   (b) An STL decomposition combined with the drift method to forecast the seasonally adjusted component.

   (c) An STL decomposition on the log transformed data combined with an ETS to forecast the seasonally adjusted component.

   (d) Holt-Winters method with damped trend and additive seasonality.

   (e) ETS(A,N,A).

   (f) ETS(A,Ad,M).

   (g) ARIMA(1,1,4).

   (h) ARIMA(3,0,2)$(1,1,1)_4$.

   (i) ARIMA(1,0,2)$(2,1,0)_{12}$.

   (j) Regression with time and Fourier terms.

10 marks

Total: 20 marks

— END OF SECTION B —

SECTION C

The following R code and output concern models for the monthly number of births in Victoria, shown in Figure 1.

```
fit_ETS <- births %>%
  model(
    ets_N = ETS(count ~ trend("N")),
    ets_A = ETS(count ~ trend("A")),
    ets_Ad = ETS(count ~ trend("Ad"))
  )
fit_ETS %>%
  tidy() %>%
  pivot_wider(names_from = ".model", values_from = "estimate")
```

```
## # A tibble: 17 x 4
##    term      ets_N     ets_A     ets_Ad
##    <chr>     <dbl>     <dbl>     <dbl>
##  1 alpha  0.468     0.327     0.299
##  2 gamma  0.000204  0.000101  0.000109
##  3 l      4.80      4.84      4.86
##  4 s0    -0.0922   -0.102    -0.101
##  5 s1    -0.192    -0.178    -0.186
##  6 s2     0.276     0.294     0.292
##  7 s3     0.142     0.126     0.125
##  8 s4     0.101     0.115     0.108
##  9 s5     0.0896    0.0956    0.0870
## 10 s6    -0.146    -0.131    -0.132
## 11 s7     0.0441    0.0315    0.0411
## 12 s8    -0.134    -0.120    -0.115
## 13 s9     0.244     0.208     0.215
## 14 s10   -0.318    -0.315    -0.315
## 15 beta   NA        0.000100  0.00389
## 16 b      NA        0.00209   0.00390
## 17 phi    NA       NA         0.973
```

1. Describe what the fit_ETS object in the output above contains and explain how this is reflected in the tibble presented.

   3 marks

2. Comment on what you see in Figure 6 and how these relate to the estimated models.

   4 marks

```r
fit_ETS %>%
  components() %>%
  autoplot() +
  labs(subtitle = "Components") +
  theme(legend.position = "bottom")
```

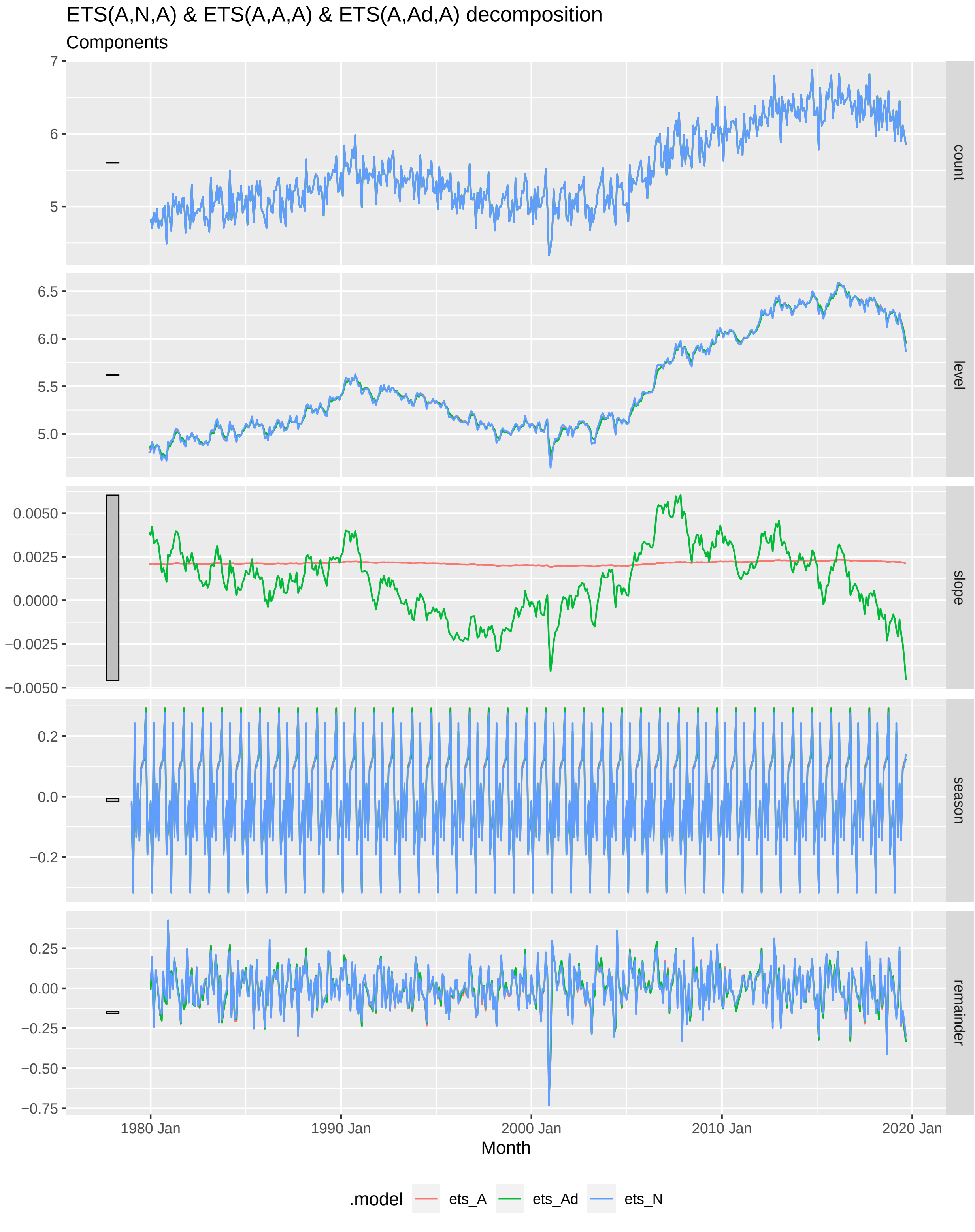


Figure 6:

3. Using the output below, comment on the fit of the models. Which model would you choose for forecasting the number of births over the next two years? Justify your choice.

2 marks

```
fit_ETS %>%
  glance() %>%
  select(-AMSE, -MAE)
```

```
## # A tibble: 3 x 7
##   .model sigma2 log_lik   AIC  AICc   BIC    MSE
##   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl>  <dbl>
## 1 ets_N  0.0180   -505. 1041. 1042. 1103. 0.0174
## 2 ets_A  0.0173   -495. 1024. 1026. 1095. 0.0167
## 3 ets_Ad 0.0173   -494. 1025. 1026. 1100. 0.0167
```

4. Write down the estimated `ets_N` model in full. (You upload your answer an image if you prefer.)

3 marks

5. Use the following output to produce point forecasts for $h = 1, 4, 12$ and 13-steps ahead. You must show your full workings. (You can upload your answer as an image if you prefer.)

3 marks

```
fit_ETS %>%
  select(ets_N) %>%
  components() %>%
  tail(14)
```

```
## # A dable: 14 x 6 [1M]
## # Key:     .model [1]
## # :        count = lag(level, 1) + lag(season, 12) + remainder
##    .model    Month count level  season remainder
##    <chr>     <mth> <dbl> <dbl>   <dbl>     <dbl>
##  1 ets_N  2018 Aug  6.46  6.31  0.101     0.0808
##  2 ets_N  2018 Sep  6.04  6.12  0.142    -0.412
##  3 ets_N  2018 Oct  6.59  6.21  0.276     0.190
##  4 ets_N  2018 Nov  6.18  6.29 -0.192     0.159
##  5 ets_N  2018 Dec  6.18  6.28 -0.0923   -0.0150
##  6 ets_N  2019 Jan  6.32  6.31 -0.0146    0.0573
##  7 ets_N  2019 Feb  5.90  6.26 -0.318    -0.0904
##  8 ets_N  2019 Mar  6.32  6.17  0.244    -0.190
##  9 ets_N  2019 Apr  5.99  6.15 -0.134    -0.0503
## 10 ets_N  2019 May  6.45  6.27  0.0441    0.258
## 11 ets_N  2019 Jun  5.90  6.16 -0.146    -0.230
## 12 ets_N  2019 Jul  6.11  6.10  0.0896   -0.139
## 13 ets_N  2019 Aug  5.99  6.00  0.101    -0.207
## 14 ets_N  2019 Sep  5.84  5.86  0.142    -0.301
```

6. Produce a 1-step-ahead 80% prediction interval from the estimated `ets_N` model. (You can upload your answer as an image if you prefer.)

3 marks

7. Forecasts from the `ets_N` model are presented in Figure 7. Discuss how these differ from the forecasts shown in Figure 5. What differences do you expect to observe in longer term forecasts from the three estimated ETS models.

2 marks

```
fit_ETS %>%
  select(ets_N) %>%
  forecast(h = 24) %>%
  autoplot(births) +
  labs(subtitle = "Births in Victoria: ETS forecasts", y = "Thousands")
```

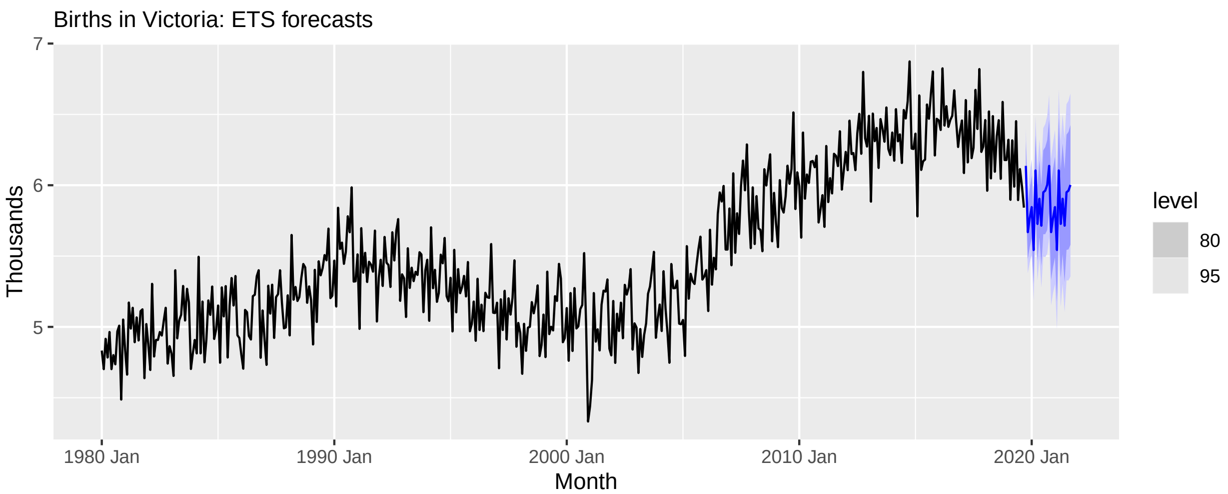


Figure 7:

Total: 20 marks

— END OF SECTION C —

SECTION D

The R code and output below relates to the monthly number of births in Victoria, shown in Figure 1.

```
births %>%
  gg_tsdisplay(difference(count, 12), plot_type = "partial")
```

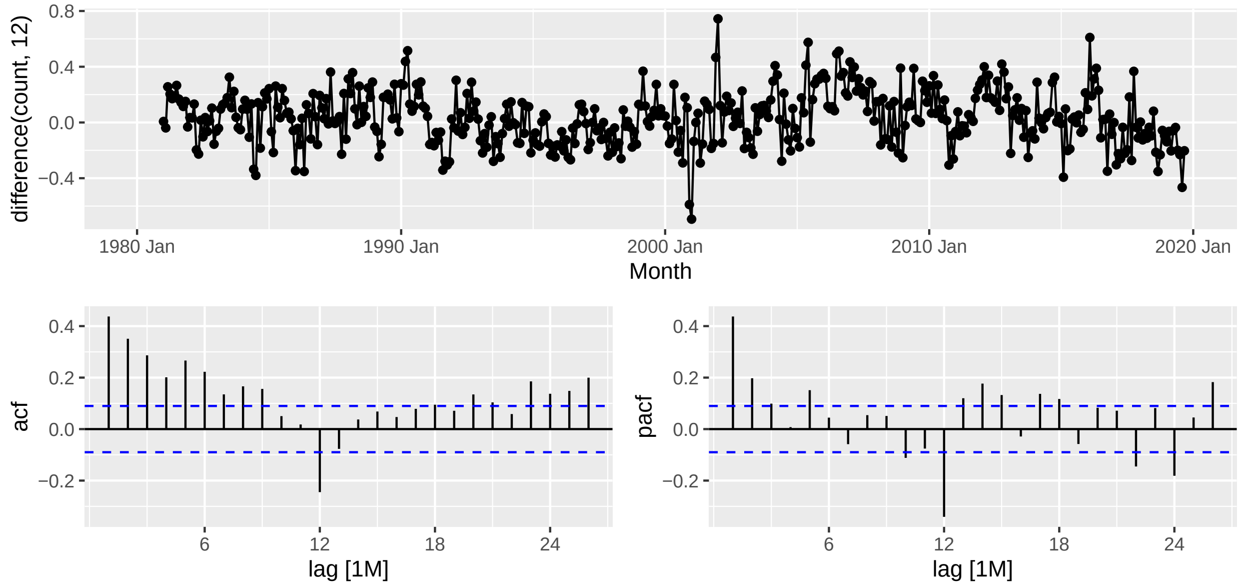


Figure 8:

```
births %>%
  mutate(sdiff_Count = difference(count, 12)) %>%
  features(sdiff_Count, unitroot_kpss)
```

```
## # A tibble: 1 x 2
##   kpss_stat kpss_pvalue
##       <dbl>       <dbl>
## 1     0.264         0.1
```

1. Using Figure 8 and the output above, describe the differencing implemented and comment on the stationarity of the resulting series.

   4 marks

2. Using the plots in Figure 8, identify a suitable ARIMA model for the monthly birth counts. Justify your choices.

   4 marks

3. Open the R file `Exam2021_for_students.R` provided to you in Moodle and run the first few lines to read in the Australian birth data and create the `births` tsibble object. Estimate the ARIMA model you have specified above. Check whether you are satisfied with the fitted model by performing some diagnostic checks of the residuals (clearly state any relevant parameters of any tests you may choose to conduct). Paste any relevant R output in the Moodle exam. (Further hints are included in the R file).

3 marks

4. An ARIMA model is estimated using the following code. Briefly explain how this model is selected and comment on the residuals shown in Figure 9. Which of the two models (from Q3 and Q4) do you prefer? Explain.

3 marks

```
fit_ARIMA <- births %>%
  model(ARIMA(count, approximation = FALSE))
fit_ARIMA %>% report()
```

```
## Series: count
## Model: ARIMA(1,0,1)(2,1,0)[12]
##
## Coefficients:
##          ar1      ma1     sar1     sar2
##       0.9737  -0.7219  -0.5795  -0.2547
## s.e.  0.0136   0.0420   0.0467   0.0459
##
## sigma^2 estimated as 0.02173:  log likelihood=229.98
## AIC=-449.96   AICc=-449.83   BIC=-429.25
```

```
fit_ARIMA %>% gg_tsresiduals()
```

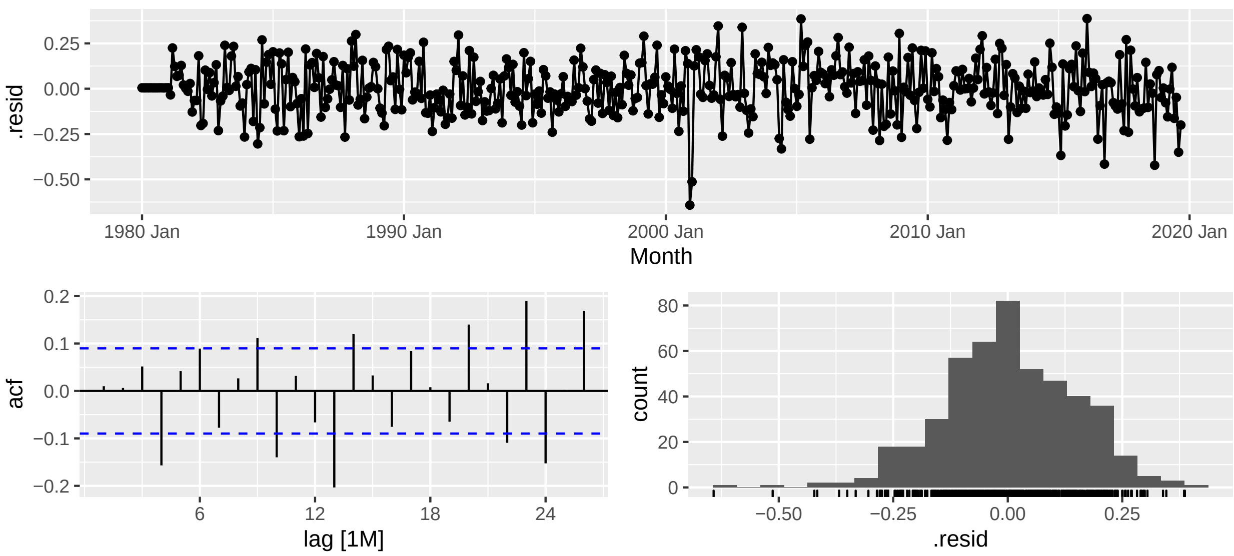


Figure 9:

5. The forecasts generated from the ARIMA model in Q4 are plotted in Figure 10 model. These seem to be trending downwards. Explain why this is the case and how you would expect the longer term forecasts from the model to behave.

2 marks

```
fit_ARIMA %>%
  forecast(h = "24 months") %>%
  autoplot(births) +
  labs(subtitle = "Births in Victoria: ARIMA forecasts", y = "Thousands")
```

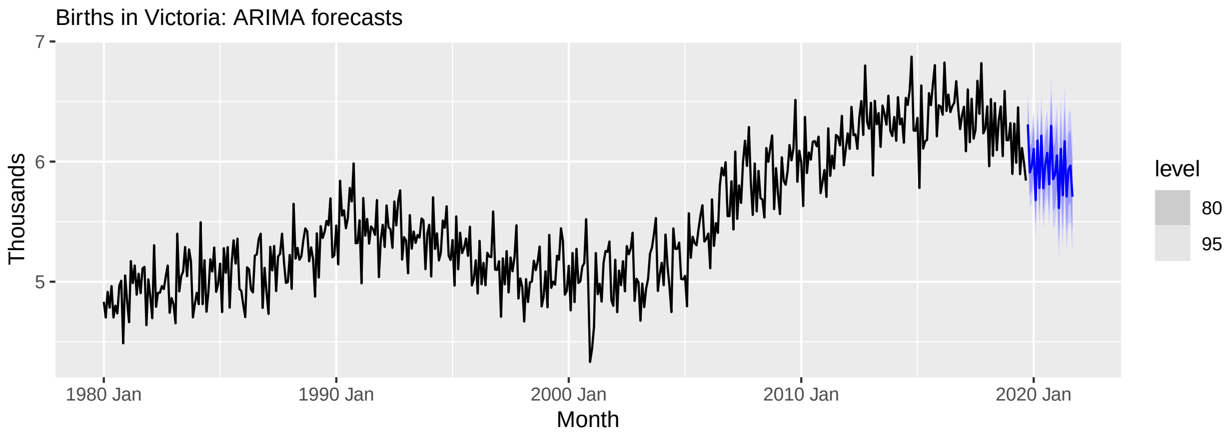


Figure 10:

6. Write out the estimated ARIMA model from Q4 in full to the point where you would be able to generate forecasts from it. There is no need to actually calculate the forecasts. (You can upload your answer as an image if you prefer.)

4 marks

Total: 20 marks

— END OF SECTION D —

SECTION E

The R code and output below relates to the monthly number of births in Victoria, shown in Figure 1.

```
knots <- yearmonth(c("1991 Jan", "2001 Jan", "2016 Jan"))

fit <- births %>%
  model(
    pw1 = ARIMA(count ~ fourier(K = 1) + trend(knots = knots)),
    pw3 = ARIMA(count ~ fourier(K = 3) + trend(knots = knots)),
    pw6 = ARIMA(count ~ fourier(K = 6) + trend(knots = knots))
  )

fit %>% pivot_longer(1:3, names_to = "Model")
```

```
## # A mable: 3 x 2
## # Key:     Model [3]
##   Model                                  value
##   <chr>                                <model>
## 1 pw1   <LM w/ ARIMA(1,0,1)(2,0,0)[12] errors>
## 2 pw3   <LM w/ ARIMA(1,0,1)(2,0,0)[12] errors>
## 3 pw6   <LM w/ ARIMA(2,0,1)(1,0,2)[12] errors>
```

```
fit %>%
  glance() %>%
  select(.model:BIC)
```

```
## # A tibble: 3 x 6
##   .model sigma2 log_lik   AIC  AICc   BIC
##   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl>
## 1 pw1    0.0209    245.  -466.  -466.  -416.
## 2 pw3    0.0209    248.  -465.  -464.  -398.
## 3 pw6    0.0159    322.  -597.  -595.  -501.
```

1. Describe the models being estimated. Which model would you select?

   2 marks

2. Comment on model `pw6`. What would this be equivalent to?

   1 marks

3. Write out in full model `pw6` and describe its features explaining the trend coefficients. (You can upload your answer as an image if you prefer.)

   10 marks

4. Comment on Figures 11 and 12 regarding the fit of the model and the forecasts generated. How could you improve the fit of the model in terms of any outliers?

   4 marks

5. Would you trust this model for forecasting over the models estimated in earlier sections? In particular consider the forecasts from the decomposition approach (Figure 5), your selected ETS model (Figure 7), your selected ARIMA model (Figure 10) and `pw6`? (Hint: think about and comment on uncertainty and short, medium and longer term forecasts).

3 marks

```r
fc <- fit %>% forecast(h = 24)
births %>%
  autoplot(count) +
  geom_line(
    data = fitted(fit %>% select("pw6")),
    aes(y = .fitted, colour = .model)
  ) +
  autolayer(fc %>% filter(.model == "pw6"), level = 95) +
  labs(subtitle = "Births in Victoria: Regression forecasts", y = "Thousands")
```

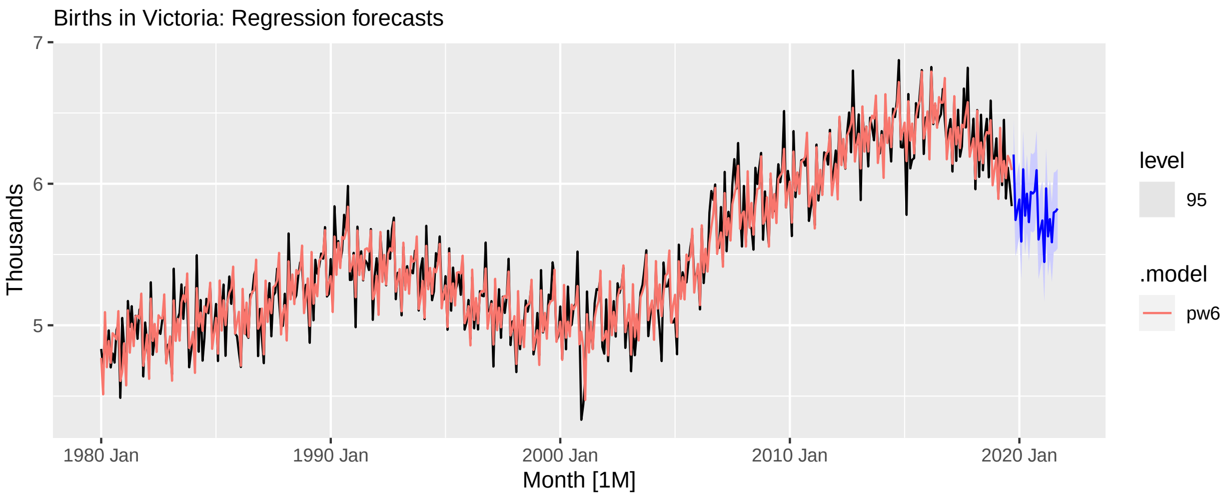


Figure 11:

```r
fit %>%
  select("pw6") %>%
  gg_tsresiduals()
```

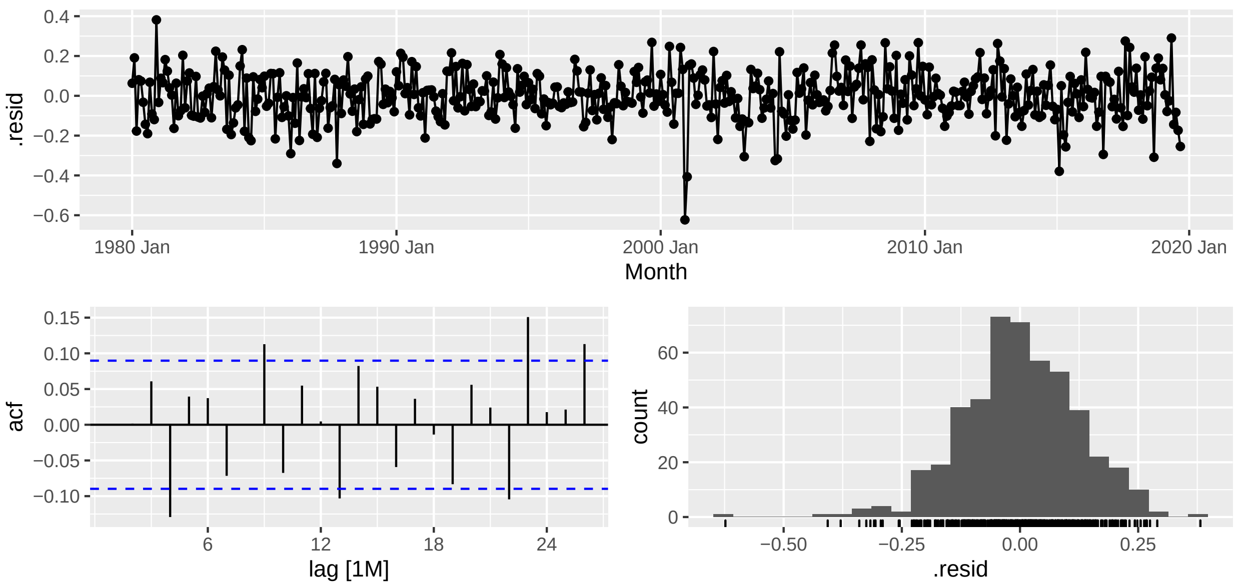


Figure 12:

Total: 20 marks

— END OF SECTION E —