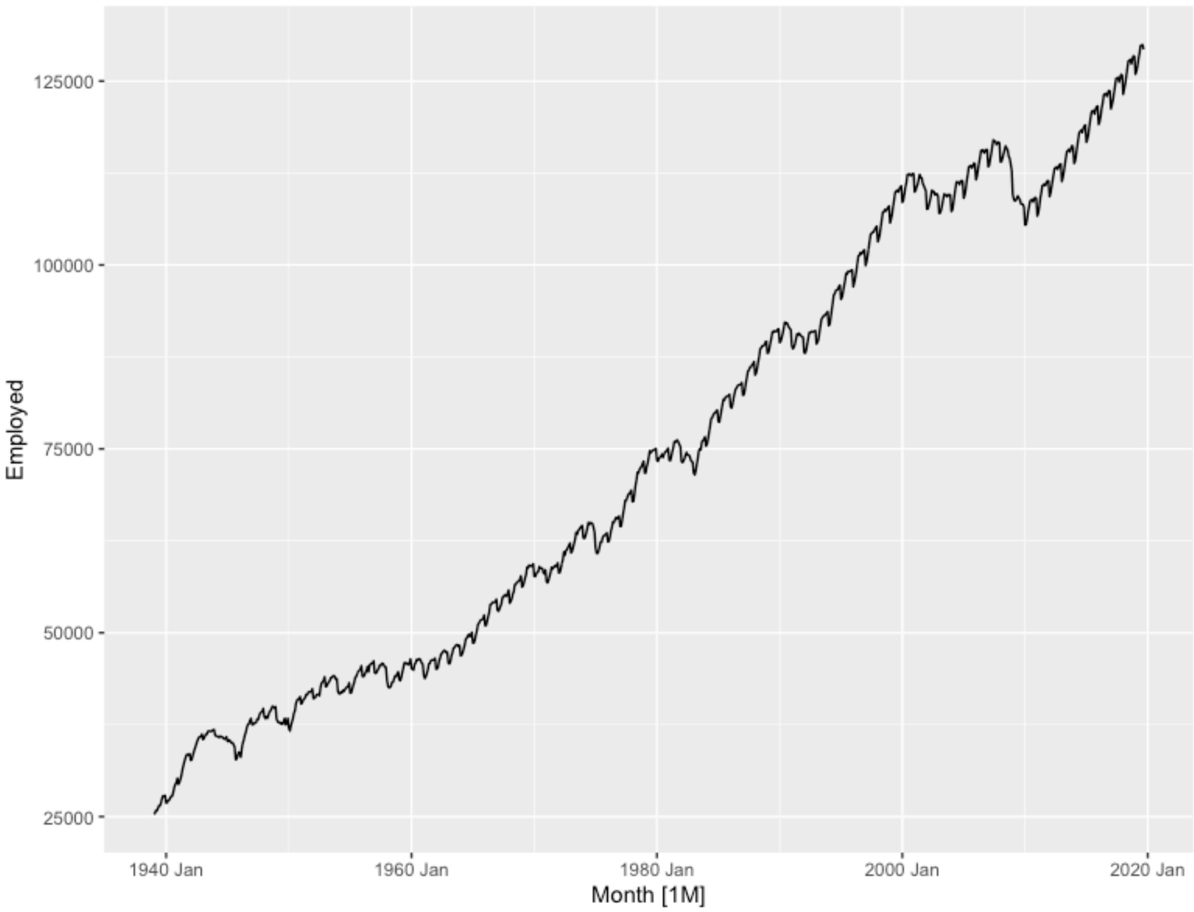
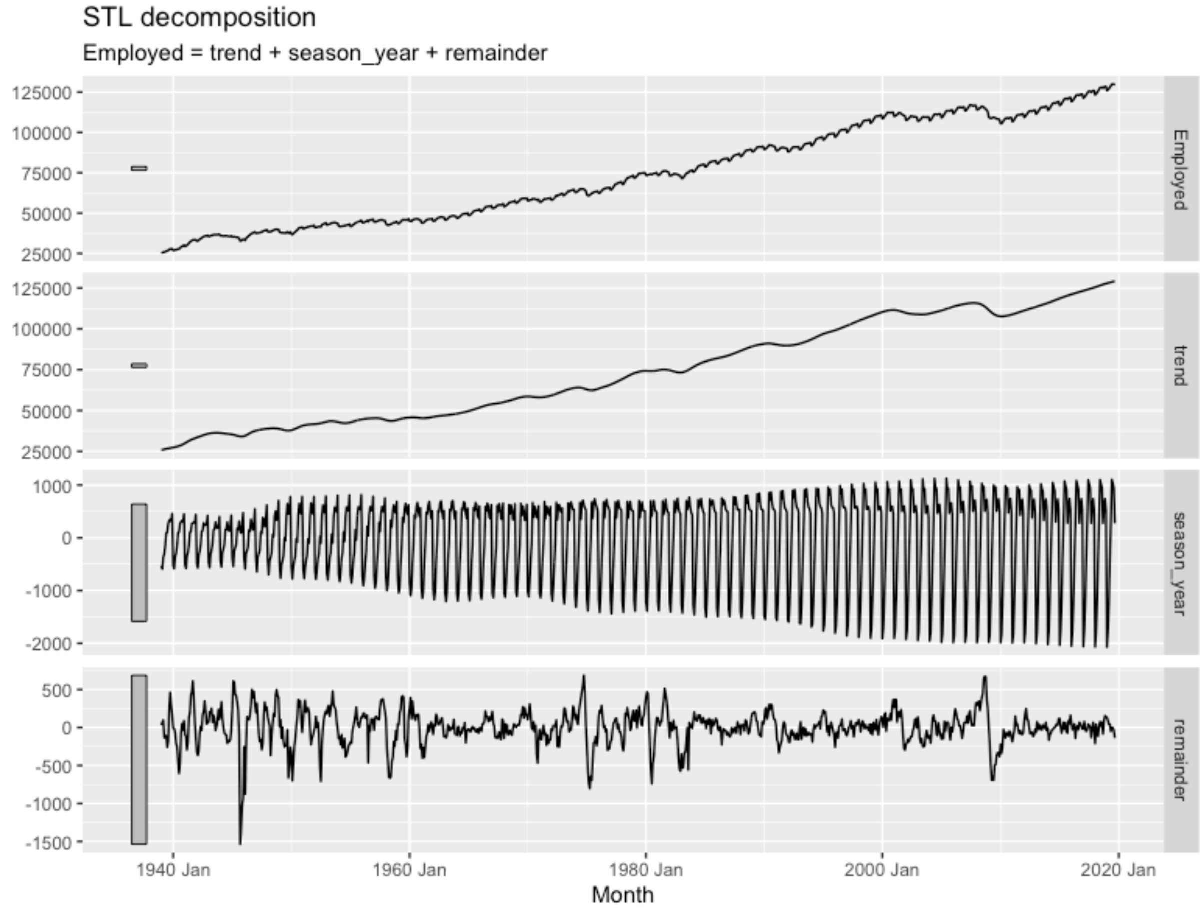
[10.a]

```
us_emp_TP <- us_employment %>%
  filter(Title == "Total Private") %>%
  select(Month, Employed)

us_emp_TP %>% autoplot(Employed)
```




```
us_emp_TP %>%
  model(STL(Employed)) %>%
  components() %>%
  autoplot()
```



[10.b]
lambda <- us_emp_TP %>% features(Employed, guerrero) %>% pull(lambda_guerrero)

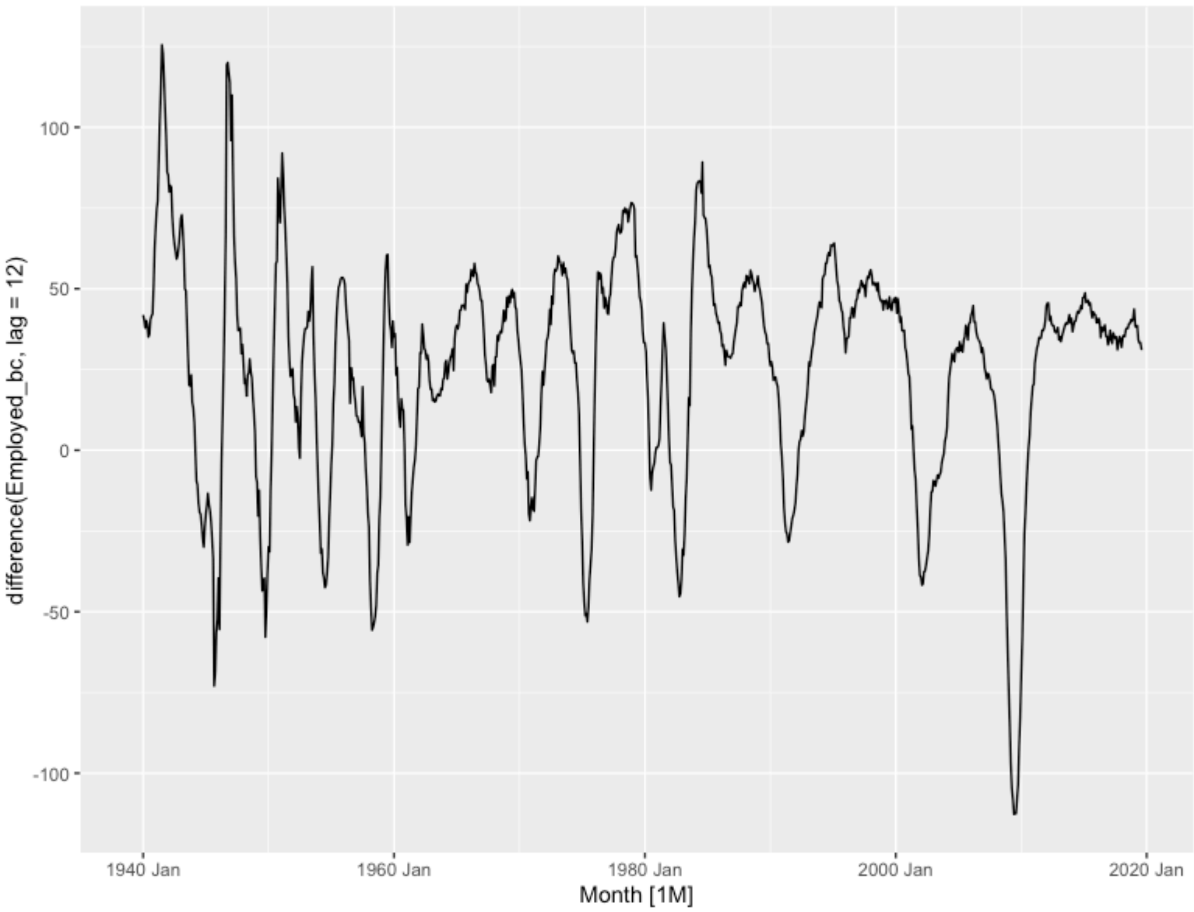
us_emp_TP <- us_emp_TP %>% mutate(Employed_bc = box_cox(Employed, lambda))

[10.c]
Not stationary.

us_emp_TP %>% features(Employed_bc, unitroot_nsdiffs)

us_emp_TP %>% features(Employed_bc %>% difference(lag = 12), unitroot_ndiffs)

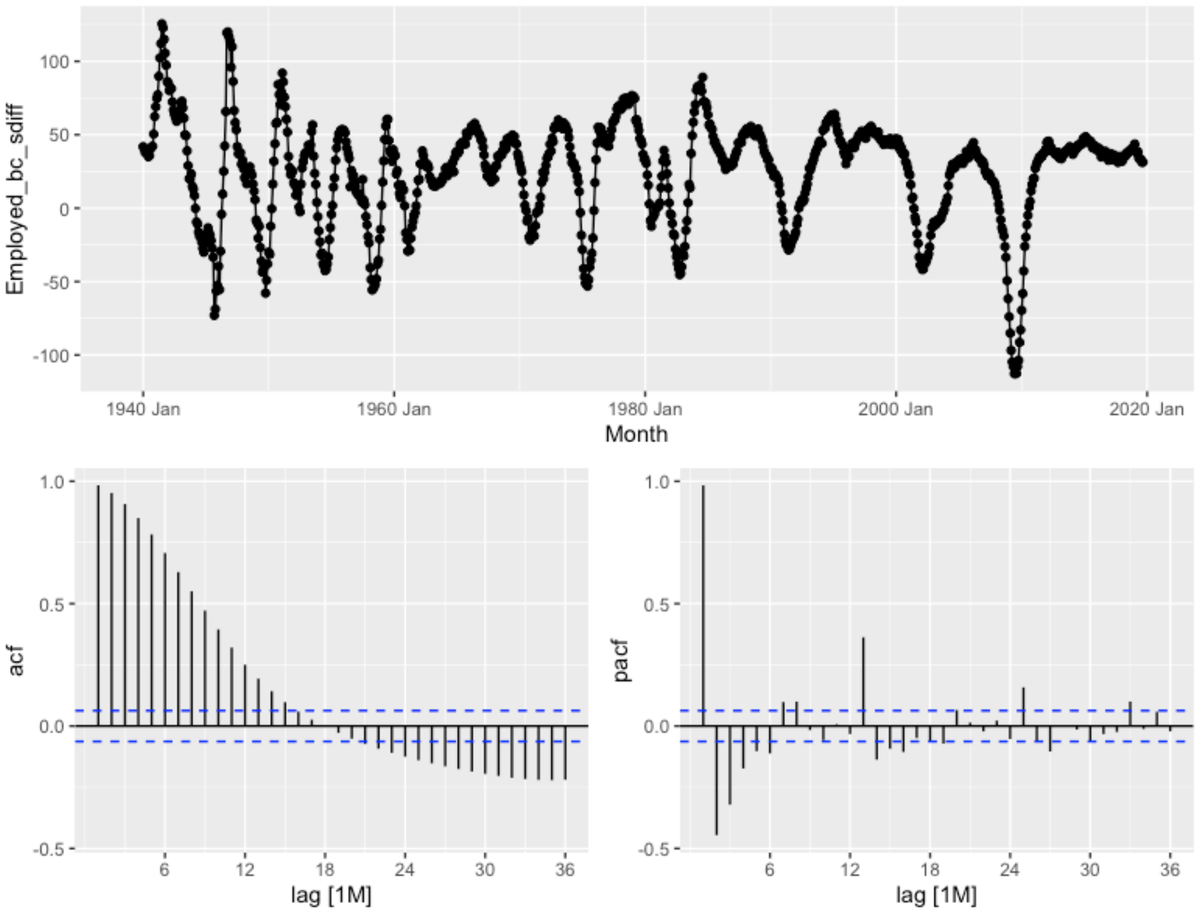
us_emp_TP %>% autoplot(difference(Employed_bc, lag = 12))




```
us_emp_TP <- us_emp_TP %>%
  mutate(Employed_bc_sdiff = difference(Employed_bc, lag = 12))
```

[10.d]

```
us_emp_TP %>% gg_tsdisplay(Employed_bc_sdiff, plot_type = "partial", lag_max = 36)
```

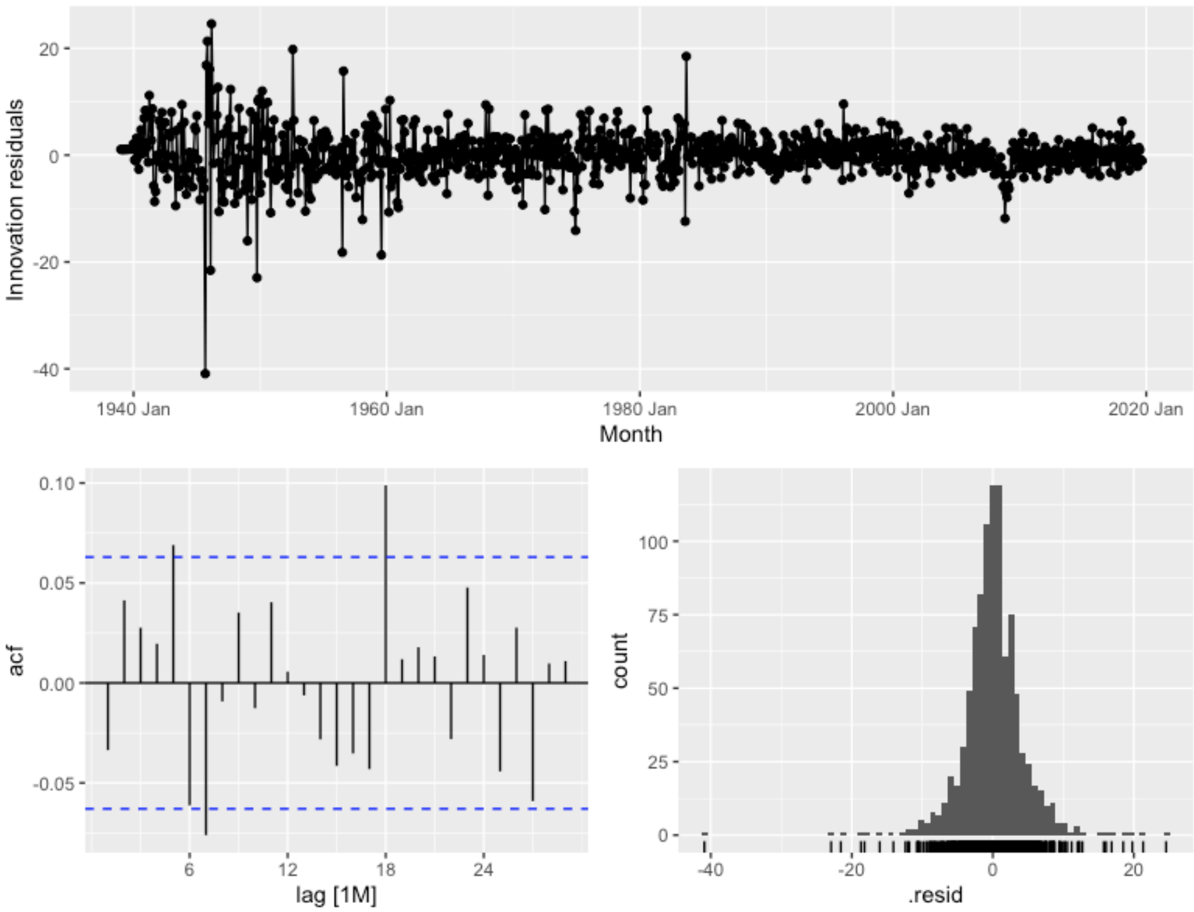



```
fit <- us_emp_TP %>%
  model(
    arima = ARIMA(box_cox(Employed, lambda)),
    arima_100010 = ARIMA(box_cox(Employed, lambda) ~ pdq(1,0,0) + PDQ(0,1,0)),
    arima_200010 = ARIMA(box_cox(Employed, lambda) ~ pdq(1,0,0) + PDQ(0,1,0)),
    arima_300010 = ARIMA(box_cox(Employed, lambda) ~ pdq(1,0,0) + PDQ(0,1,0)),
    arima_400010 = ARIMA(box_cox(Employed, lambda) ~ pdq(1,0,0) + PDQ(0,1,0))
  )

fit %>% pivot_longer(everything())

fit %>% glance() %>% arrange(AICc)
```

ARIMA model is the best.

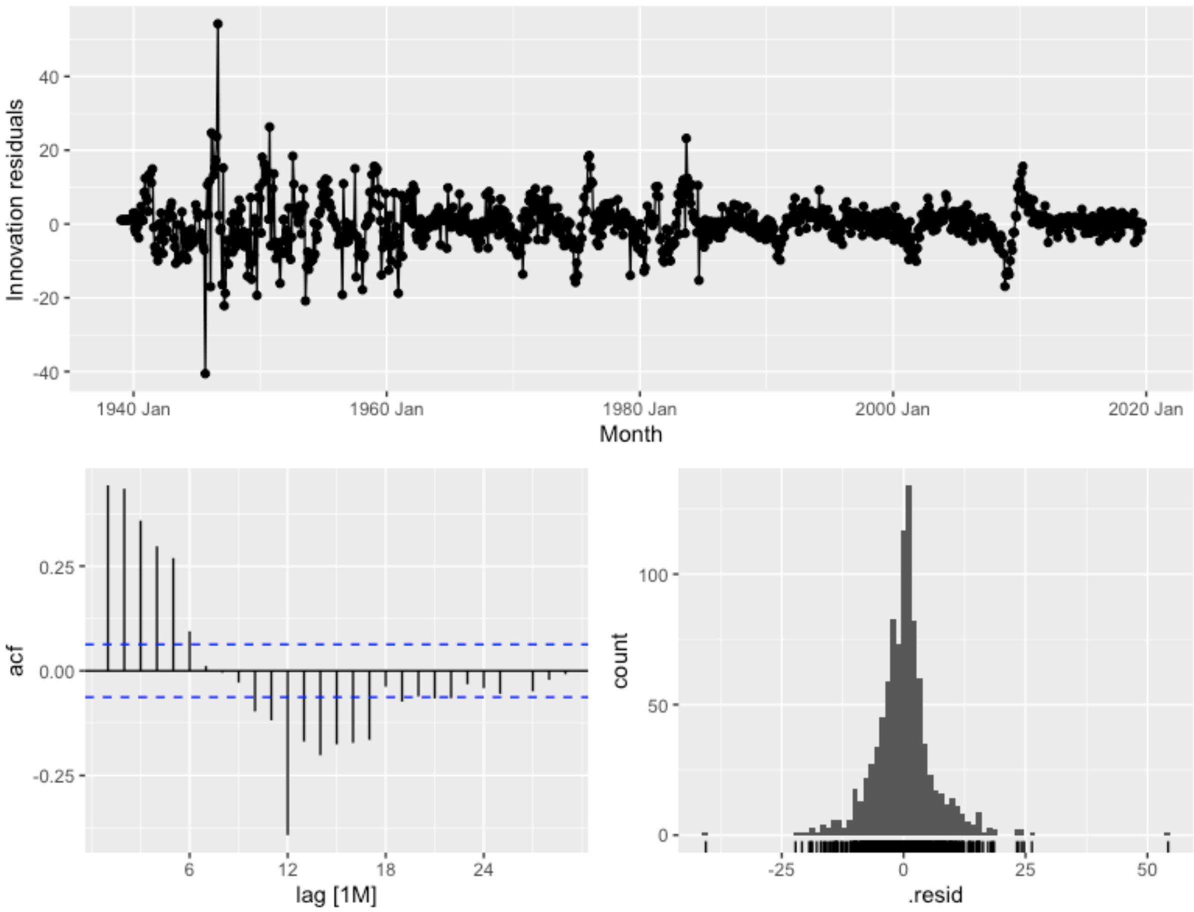
[10.e]

```
fit %>%
  select(arima) %>%
  gg_tsresiduals()
```



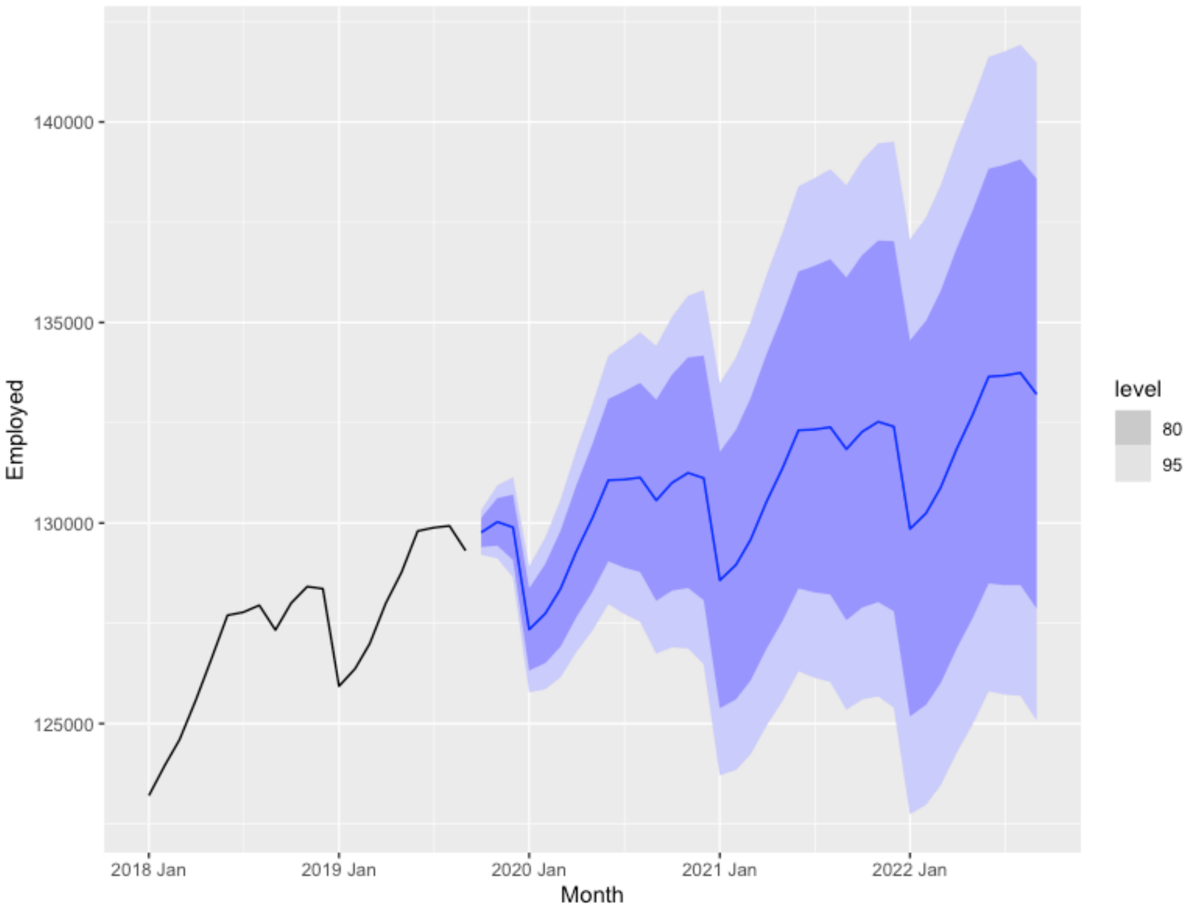
The ARIMA model has residuals which resemble white noise.

```
fit %>%
  select(arima_400010) %>%
  gg_tsresiduals()
```

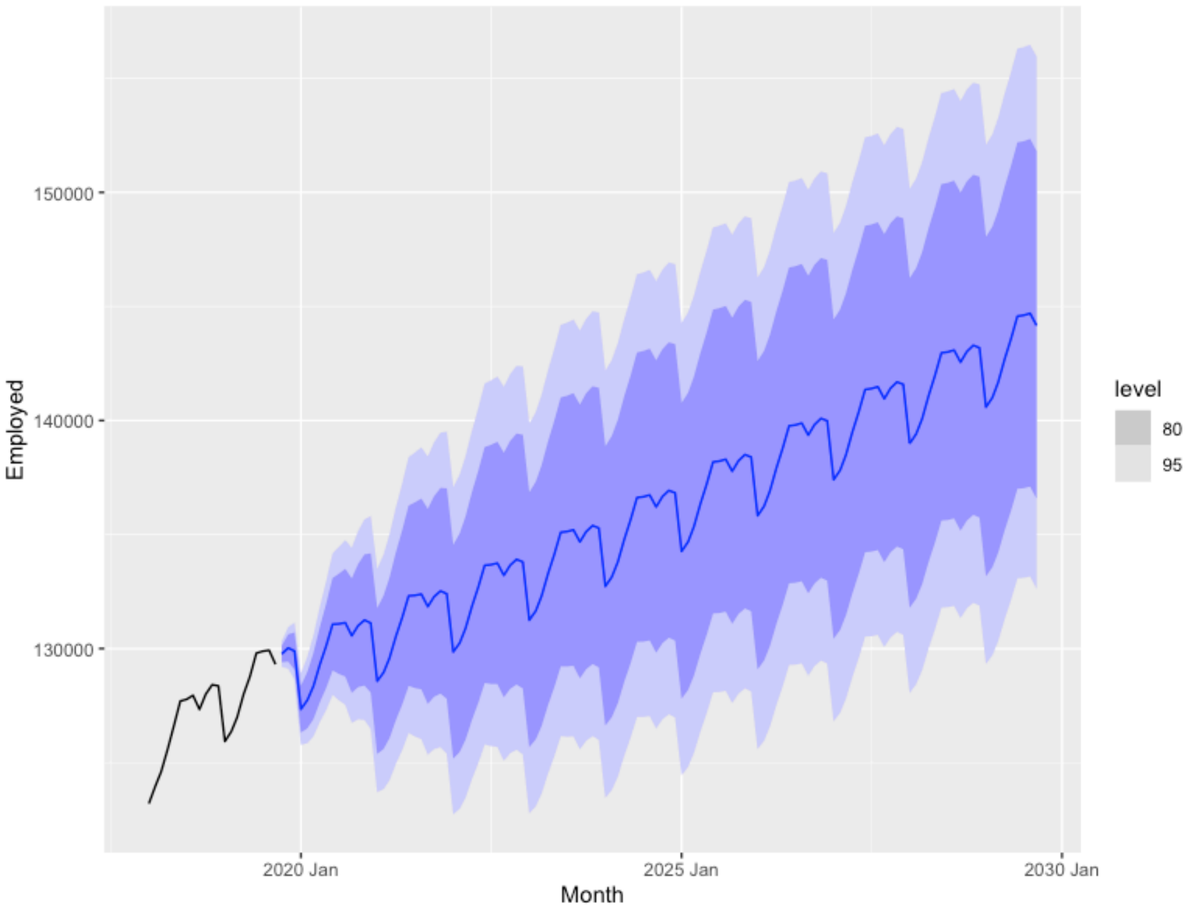


[10.f]

```
fit %>%
  select(arima) %>%
  forecast(h = "3 years") %>%
  hilo(level = 95)

fit %>%
  select(arima) %>%
  forecast(h = "3 years") %>%
  autoplot(us_emp_TP %>% filter(year(Month) >= 2018))
```

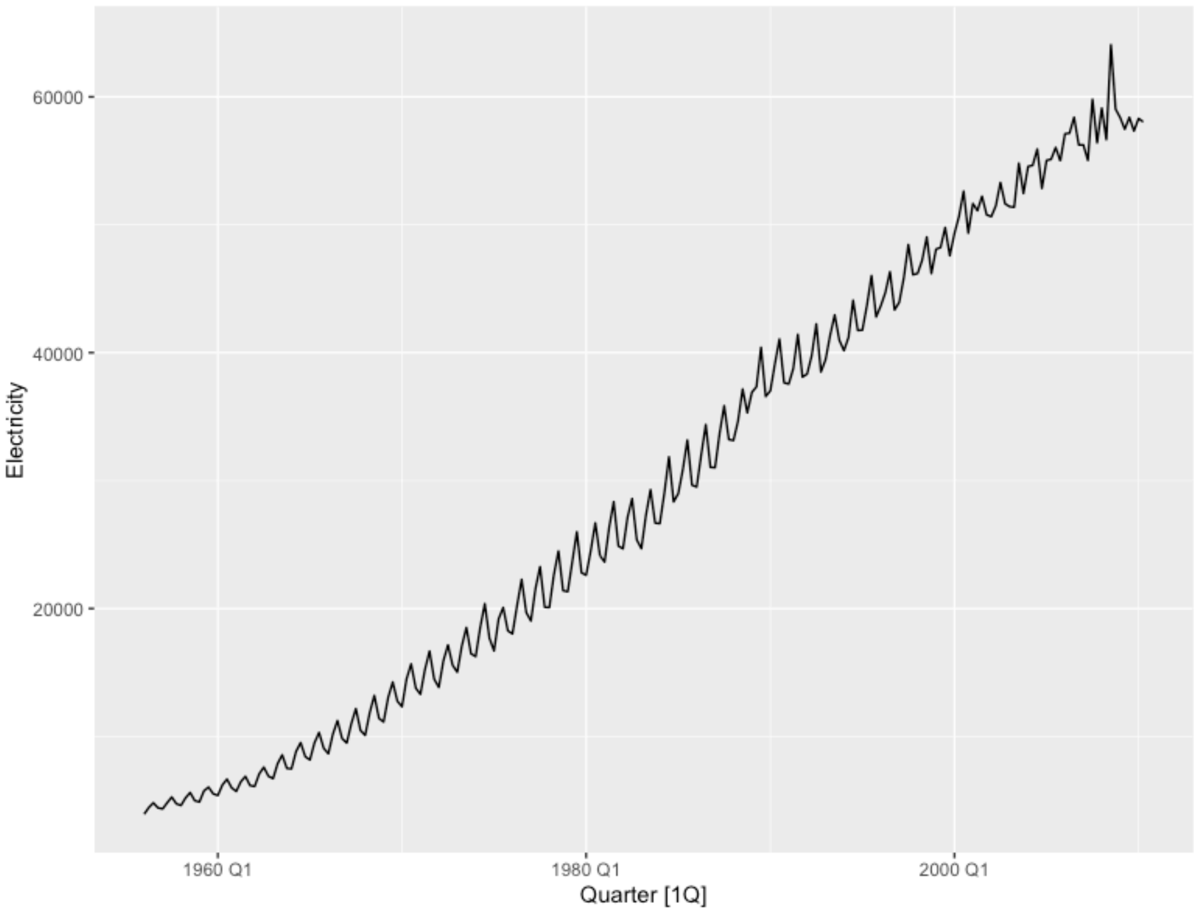



```
[10.g]
fit %>%
  select(arima) %>%
  forecast(h = "10 years") %>%
  autoplot(us_emp_TP %>% filter(year(Month) >= 2018))
```

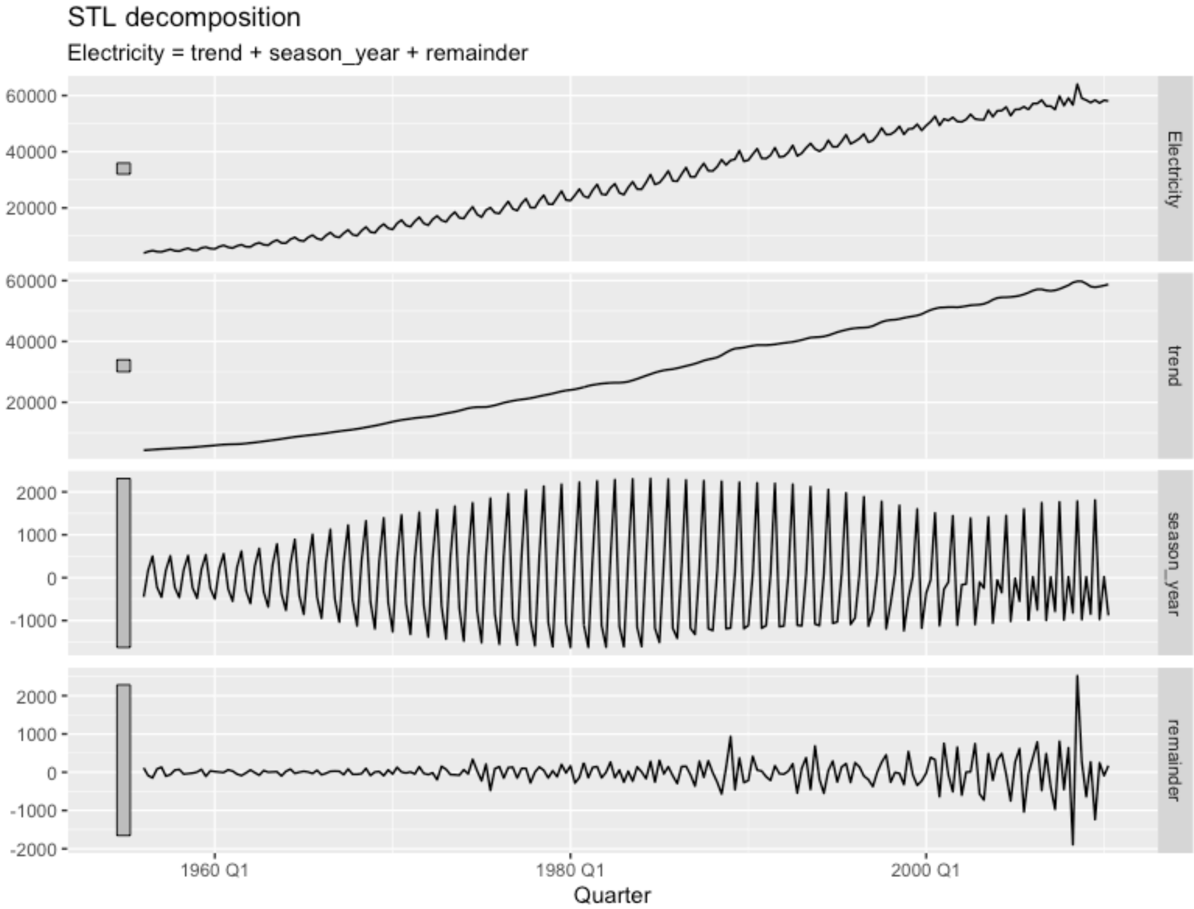


Due to the black swan event, it is not possible to check the forecast for long ranges. 2-3 years is a decent range, especially as this gives a decent near forecast. If the historical pattern is consistant and is expected to continue in the future then even 5-6 years might be considered.

[11.a]
aus_production %>% autoplot(Electricity)

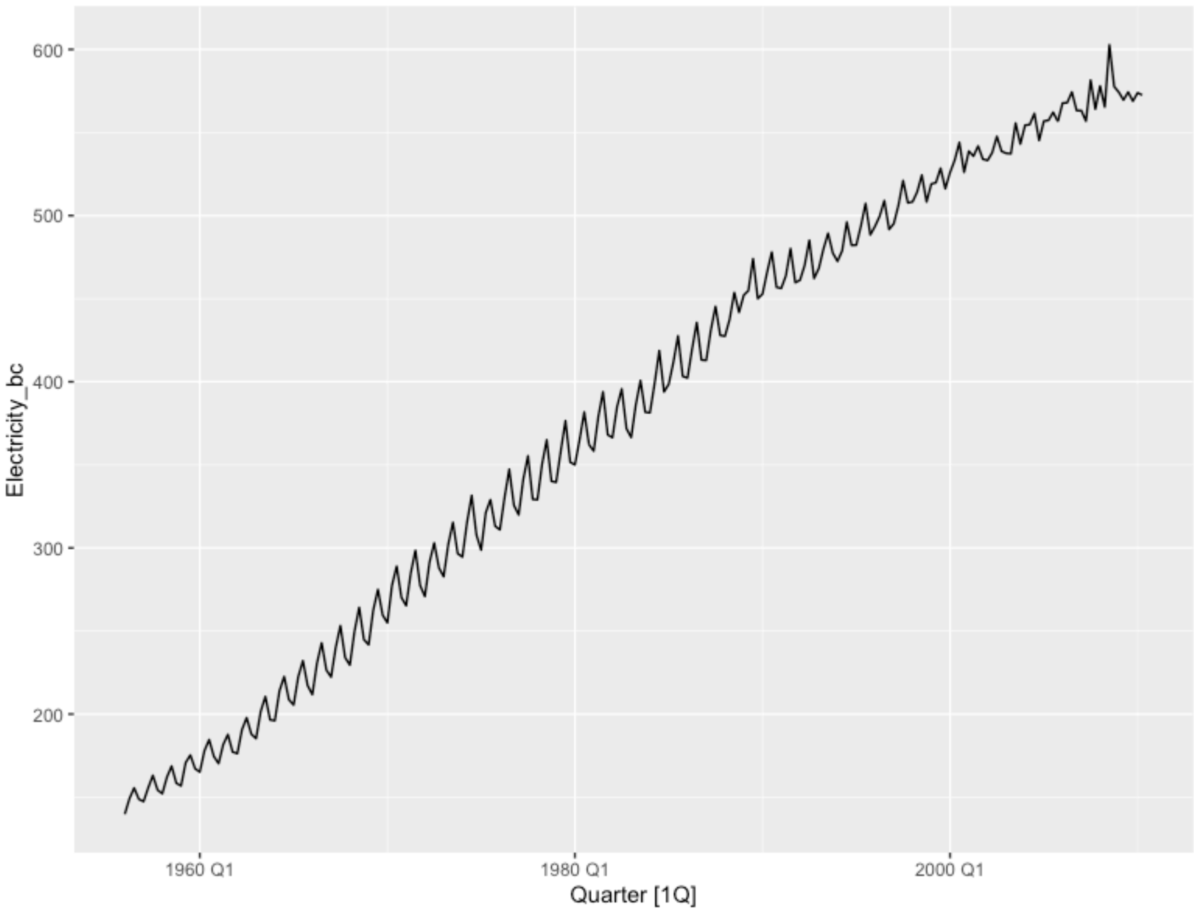



```
aus_production %>% model(STL(Electricity)) %>% components() %>% autoplot()
```

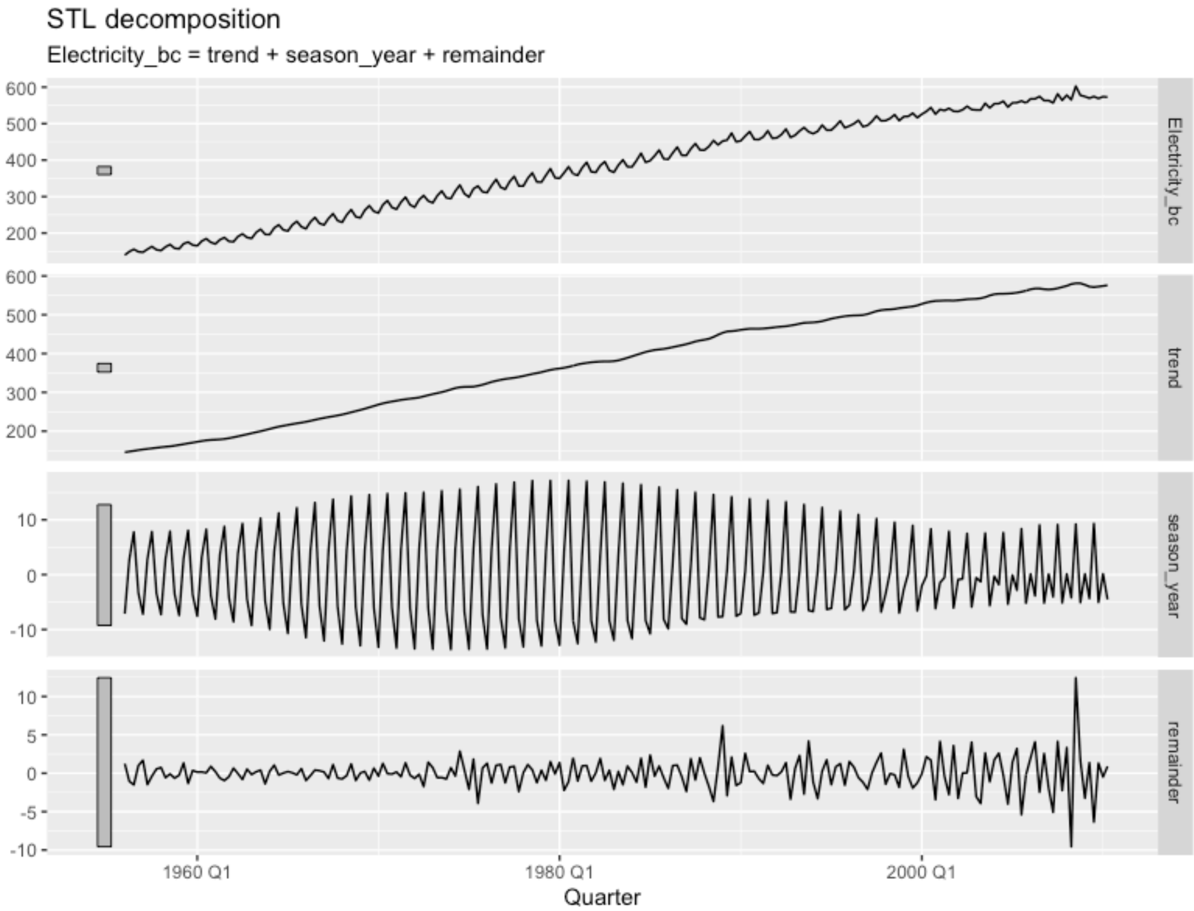


There is changing variation in the seasonal pattern, so transformation is required.

```
lambda <- aus_production %>%
  features(Electricity, guerrero) %>%
  pull(lambda_guerrero)

aus_electricity <- aus_production %>%
  select(Quarter, Electricity) %>%
  mutate(Electricity_bc = box_cox(Electricity, lambda))

aus_electricity %>% autoplot(Electricity_bc)
```

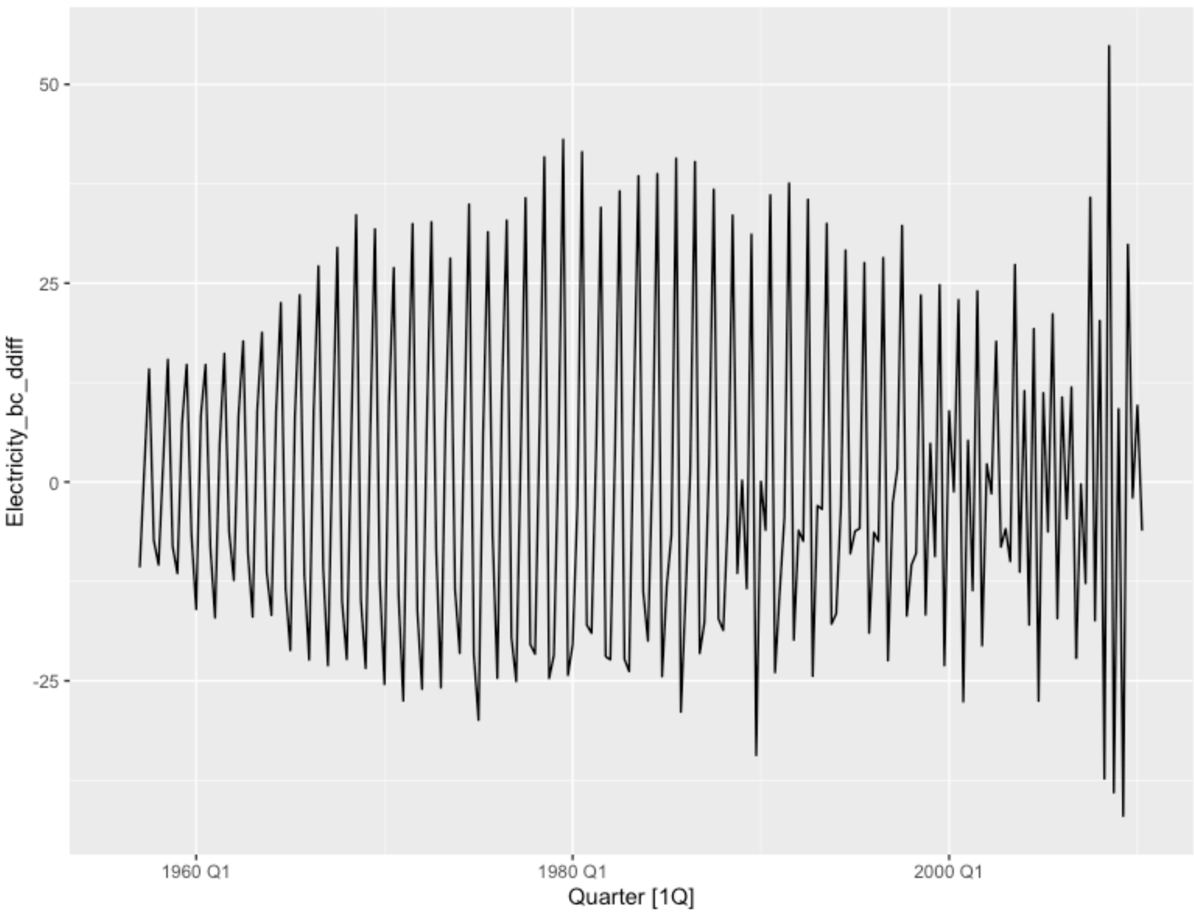



```
aus_electricity %>% model(STL(Electricity_bc)) %>% components() %>% autoplot()
```

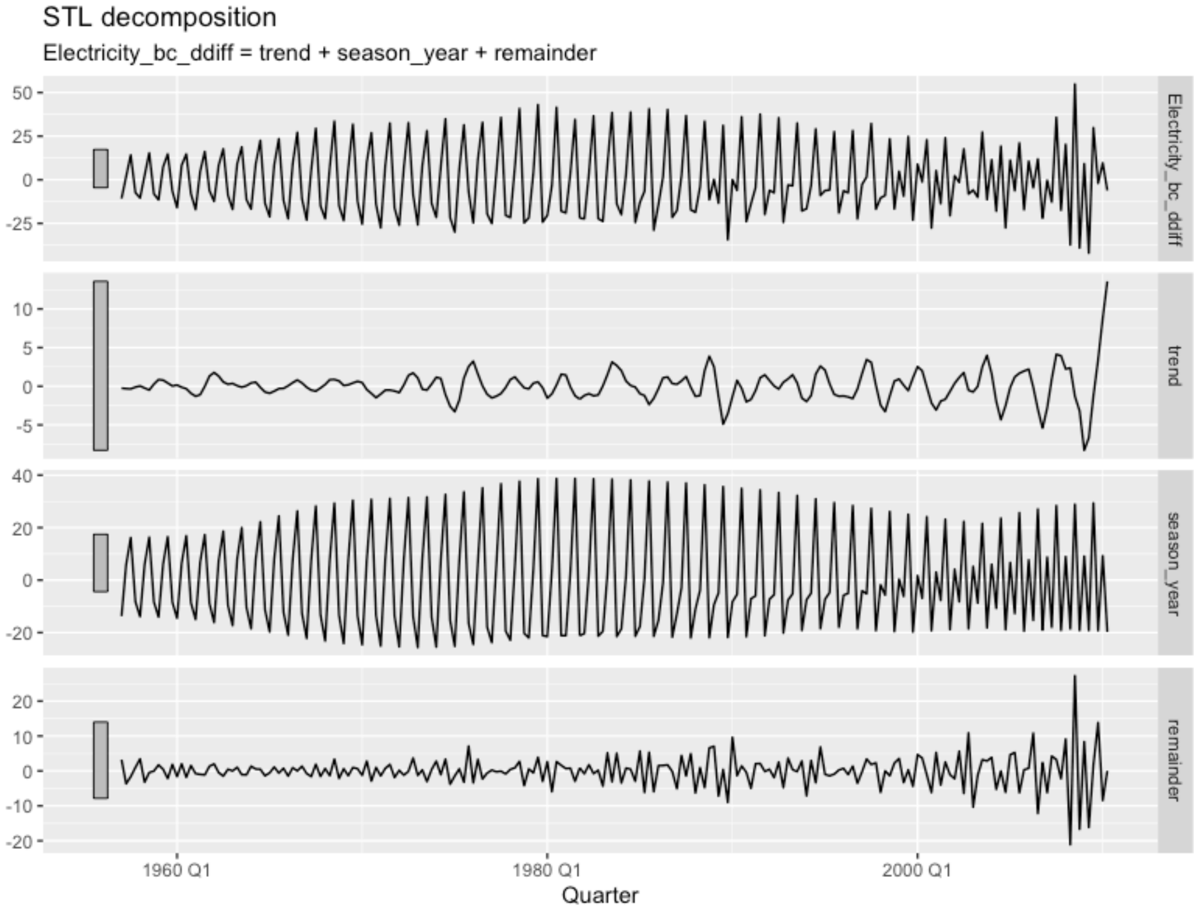


[11.b]
The data is not stationary as it has trend and seasonality.

```
aus_electricity %>% features(Electricity_bc, unitroot_nsdiffs)
aus_electricity %>% features(Electricity_bc %>% difference(lag = 4), unitroot_ndiffs)

aus_electricity <- aus_electricity %>%
  mutate(Electricity_bc_ddiff = Electricity_bc %>% difference(lag = 3) %>% difference(lag = 1))

aus_electricity %>% autoplot(Electricity_bc_ddiff)
```

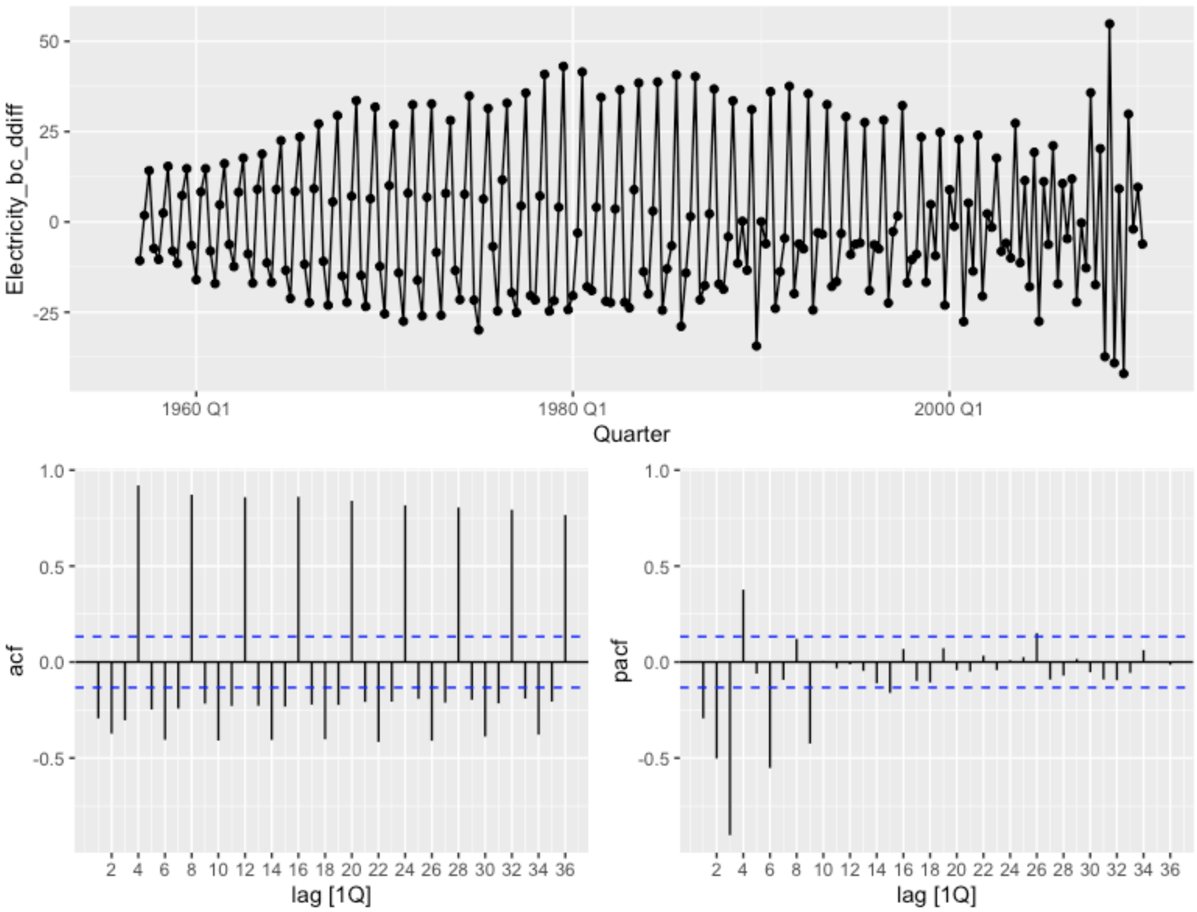



```
aus_electricity %>%
  filter(!is.na(Electricity_bc_ddiff)) %>%
  model(STL(Electricity_bc_ddiff)) %>%
  components() %>% autoplot()
```



[11.c]

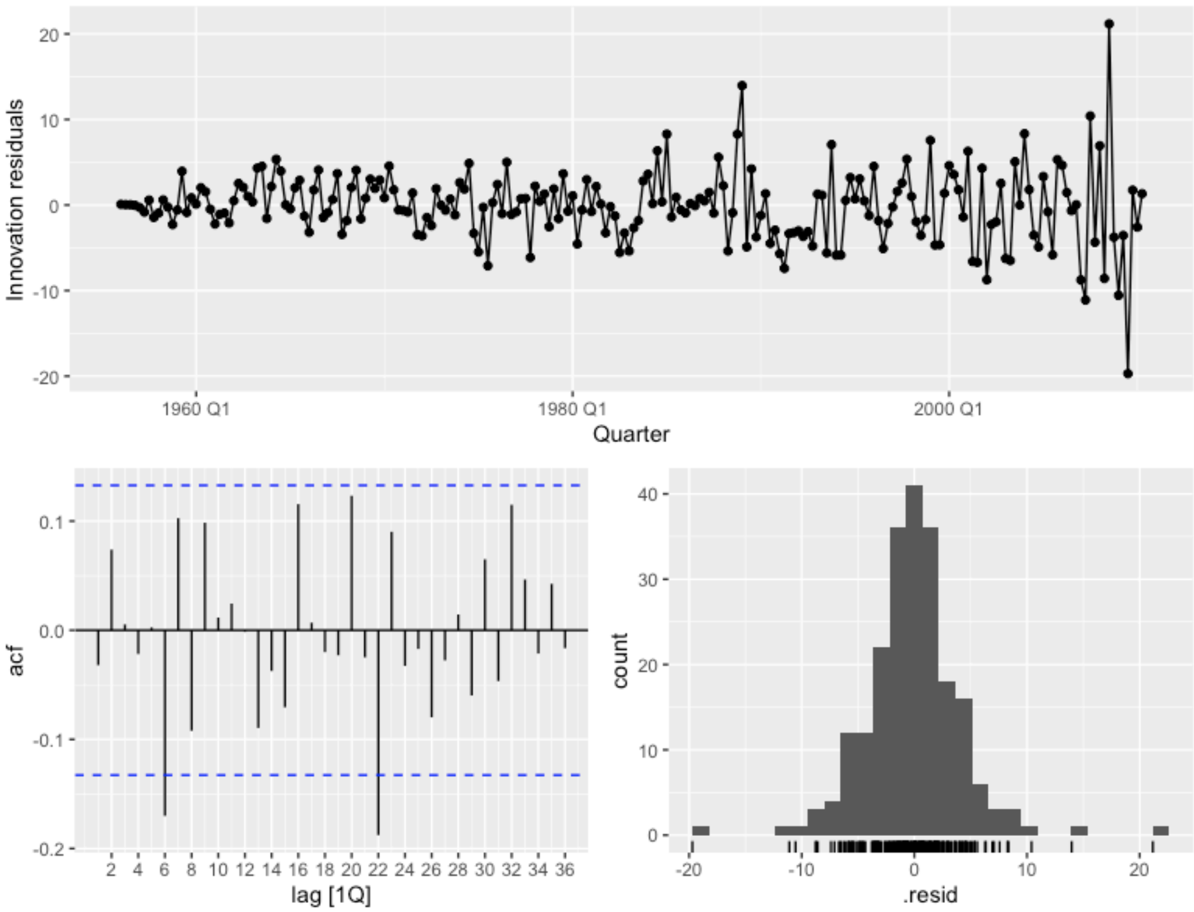
```
aus_electricity %>%
  gg_tsdisplay(Electricity_bc_ddiff, plot_type = "partial", lag_max = 36)
```




```
# Possible Model = (1,1,1)(1,1,1)[4]

fit <- aus_electricity %>%
  model(
    arima = ARIMA(box_cox(Electricity, lambda)),
    arima_111111 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,1) + PDQ(1,1,1)),
    arima_111011 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,1) + PDQ(0,1,1)),
    arima_011111 = ARIMA(box_cox(Electricity, lambda) ~ pdq(0,1,1) + PDQ(1,1,1)),
    arima_011011 = ARIMA(box_cox(Electricity, lambda) ~ pdq(0,1,1) + PDQ(0,1,1)),
    arima_111110 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,1) + PDQ(1,1,0)),
    arima_110111 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,0) + PDQ(1,1,1)),
    arima_110110 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,0) + PDQ(1,1,0)),
    arima_110112 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,0) + PDQ(1,1,2)),
    arima_112011 = ARIMA(box_cox(Electricity, lambda) ~ pdq(1,1,2) + PDQ(0,1,1))
  )

fit %>% pivot_longer(everything())

fit %>% glance() %>% arrange(AIC)
# The best model is 111011.
```

[11.d]
If not, try to find another ARIMA model which fits better.

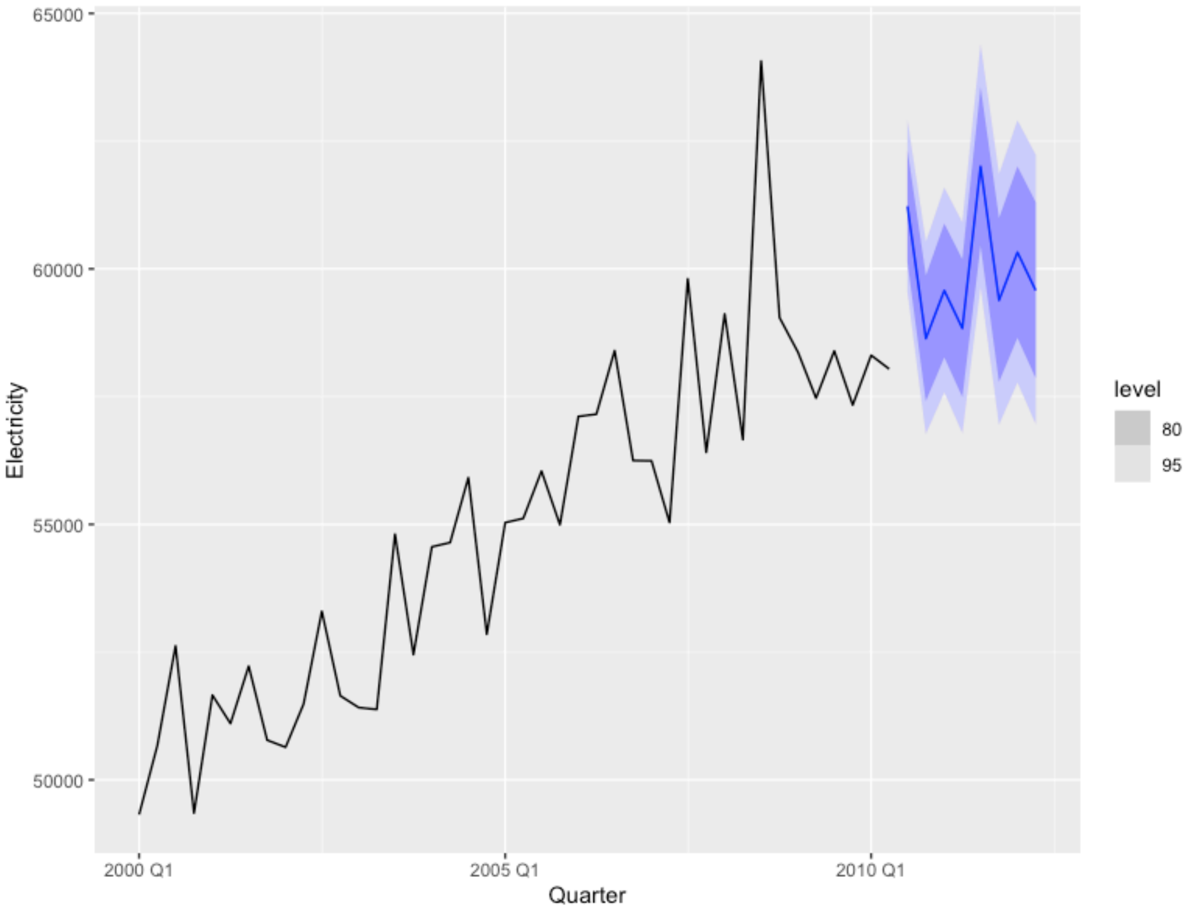
```
fit %>%
  select(arima_111011) %>%
  gg_tsresiduals(lag = 36)
```



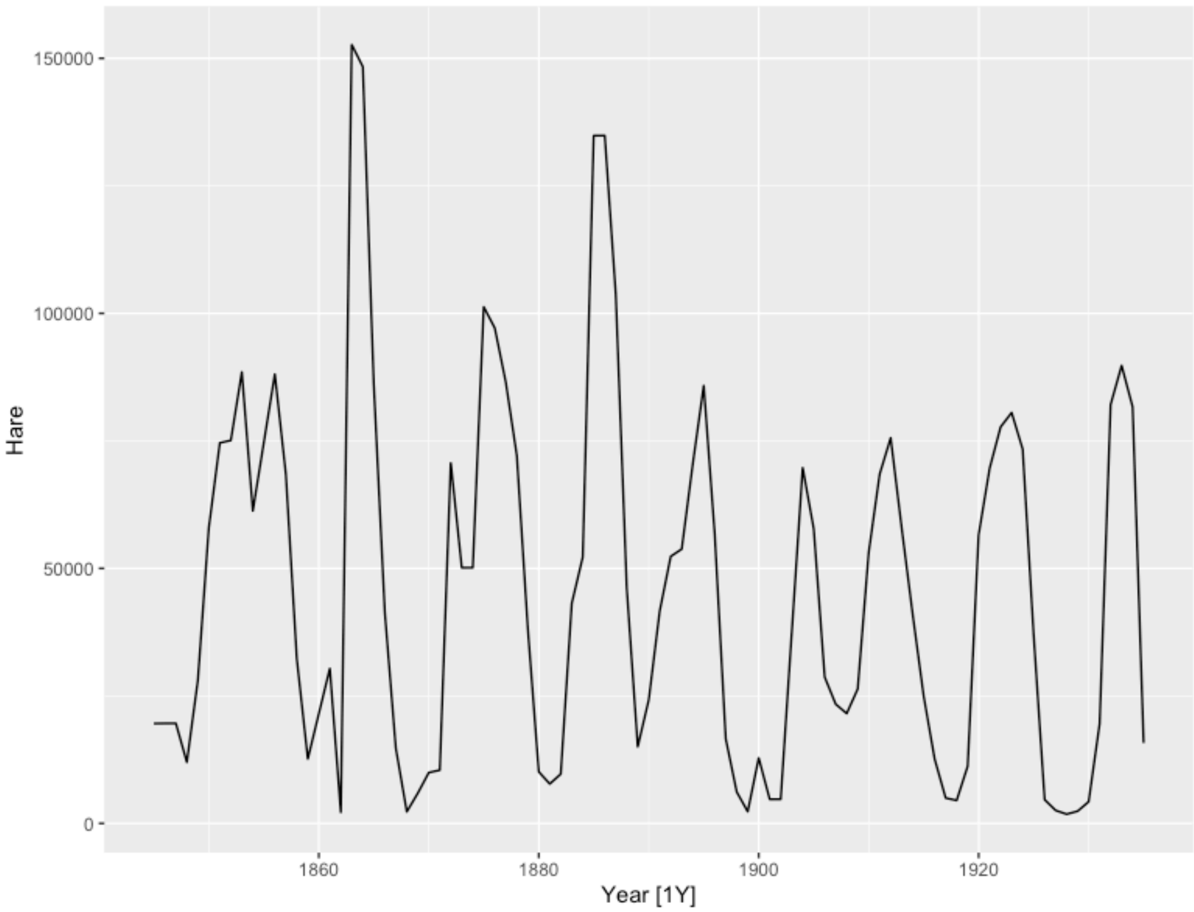
The residuals resemble white noise. There are spikes at lag 6 and 22 but they are far back that they can be ignored.

[11.e]

```
fit %>%
  select(arima_111011) %>%
  forecast(h = "24 months") %>%
  autoplot(aus_electricity %>% filter(year(Quarter) >= 2000))
```

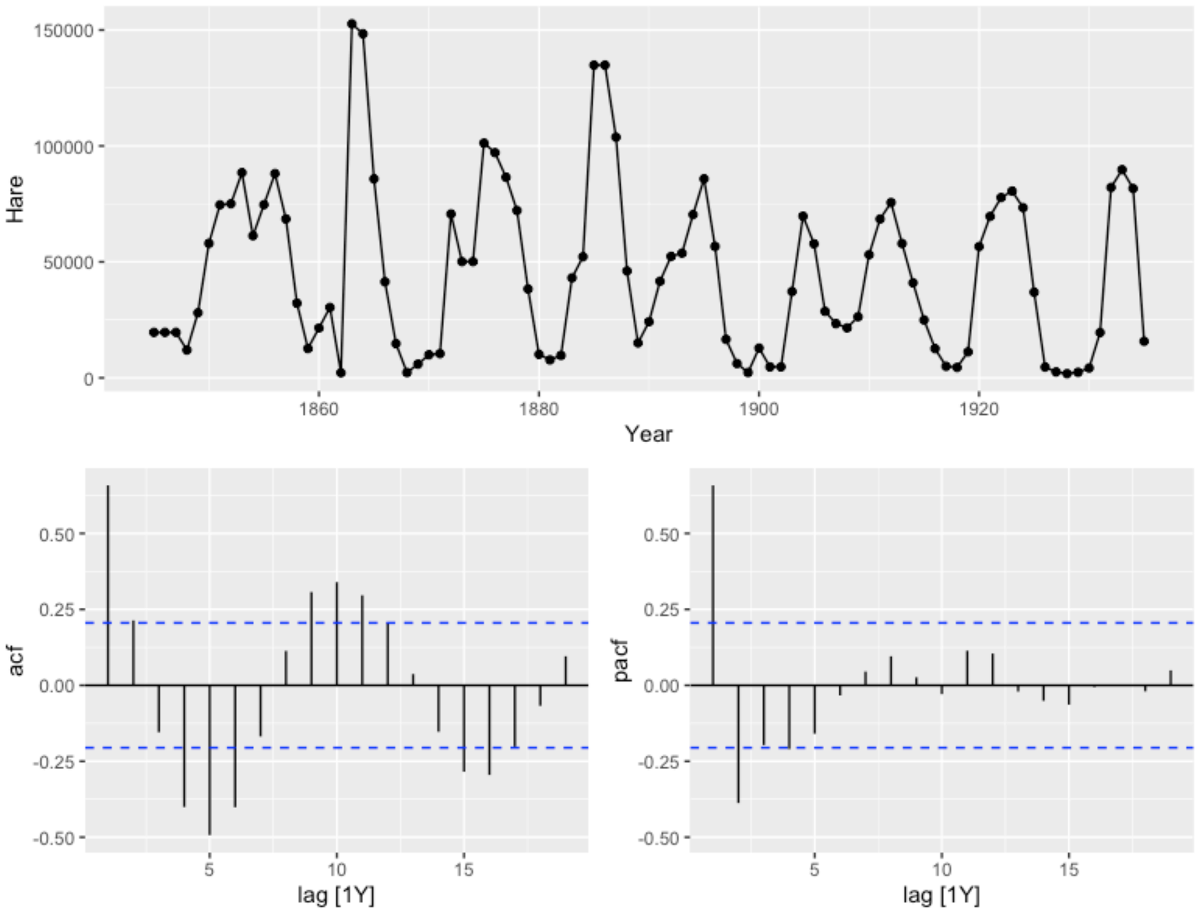


[15.a]
pelt %>% autoplot(Hare)



[15.b]
Model - ARIMA(4,0,0)

[15.c]
pelt %>% gg_tsdisplay(Hare, plot_type = "partial")



The ACF plot is sinusoidal and the PACF has spike at lag 4 but none after. Therefore the plot is correct.

[15.d]

$$y[t] = c + \phi[1]y[t-1] + \phi[2]y[t-2] + \phi[3]y[t-3] + \phi[4]y[t-4] + \varepsilon[t]$$
$$= 30993 + 0.82y[t-1] - 0.29y[t-2] - 0.01y[t-3] - 0.22y[t-4]$$

$$y[1936] = 30993 + 0.82y[1935] - 0.29y[1934] - 0.01y[1933] - 0.22y[1932]$$
$$= 30993 + 0.82*15760 - 0.29*81660 - 0.01*89760 - 0.22*82110$$
$$= 1273$$

Similarly 1937 = 6903, 1938 = 18161, 1939 = 40403

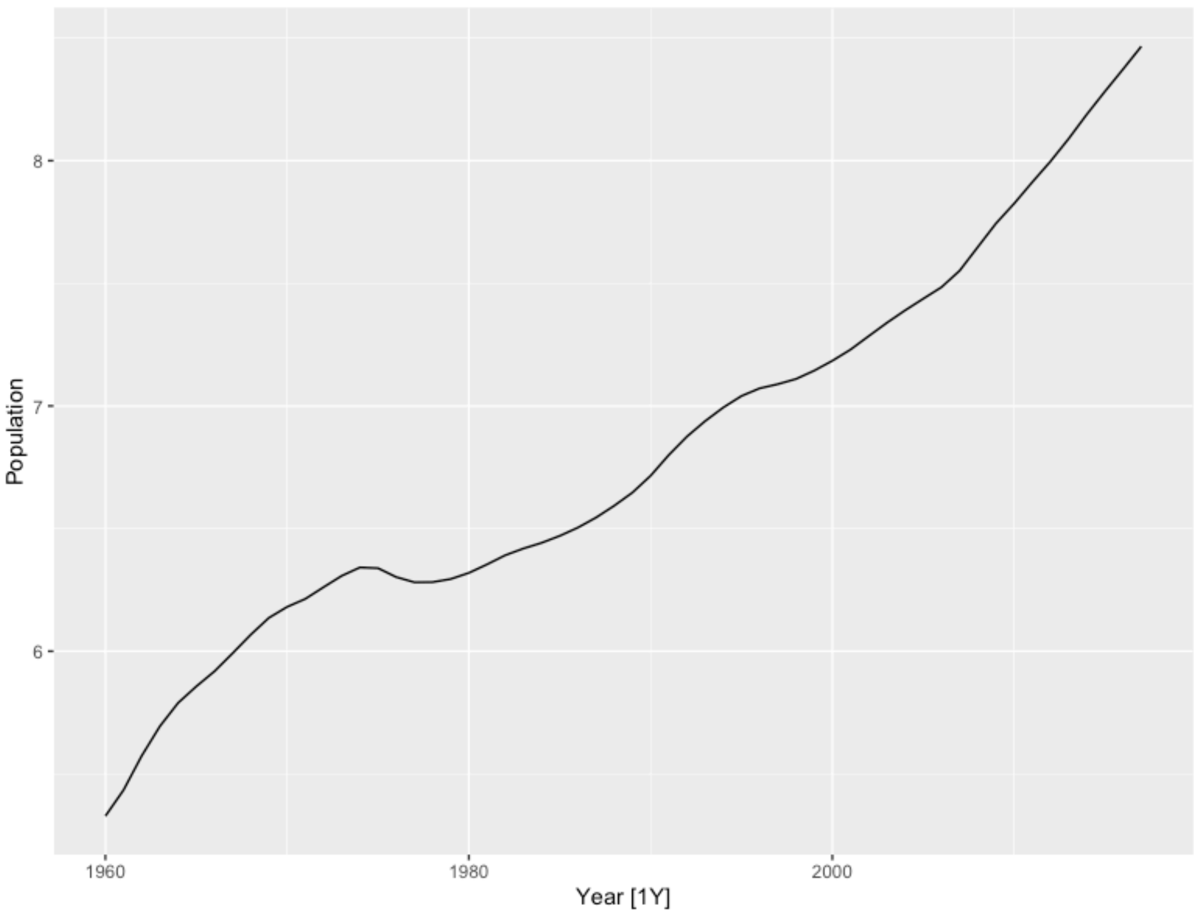
[15.e]
```
pelt %>%
  model(ARIMA(Hare ~ pdq(4,0,0))) %>%
  report()
```

The forecasted numbers are slightly different as the manually calculated numbers are using rounded off coefficients rather than the actuals ones. Also, as the future forecasted numbers are based on the past forecasted values, the difference in the past values get compounded as the future values are calculated.

[16.a]
```
popu_switz <- global_economy %>%
  filter(Country == "Switzerland") %>%
  select(Year, Population) %>%
```

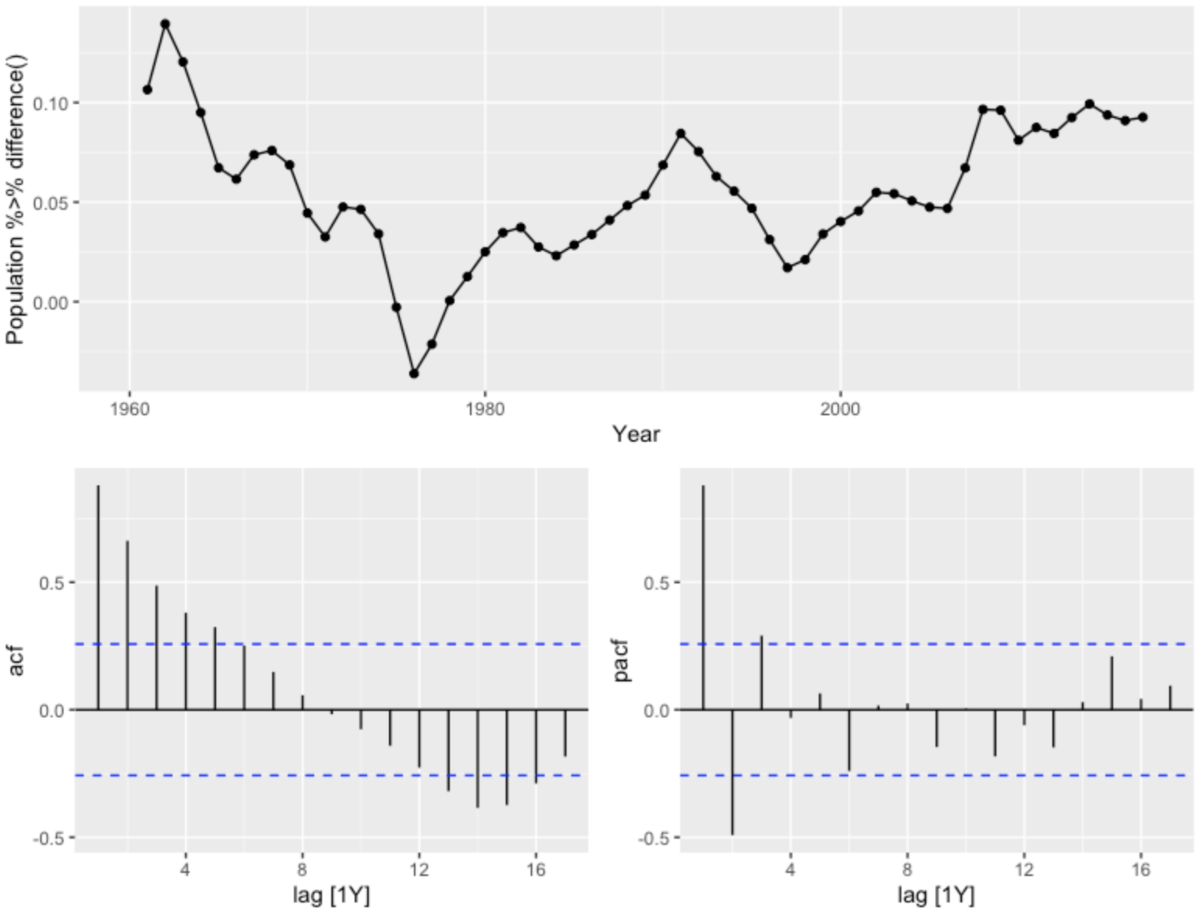
mutate(Population = Population/1e6)

popu_switz %>% autoplot(Population)



[16.b]
Model - ARIMA(3,1,0)

[16.c]
popu_switz %>% gg_tsdisplay(Population %>% difference(), plot_type = "partial")



ACF plot is sinusoidal and the PACF has a spike at lag 3 and none after. Therefore the model is appropriate.

[16.d]

$$y[t] = c + y[t-1] + \phi[1](y[t-1] - y[t-2]) + \phi[2](y[t-2] - y[t-3]) + \phi[3](y[t-3] - y[t-4]) + \varepsilon[t]$$
$$= 0.0053 + y[t-1] + 1.64*(y[t-1] - y[t-2]) - 1.17(y[t-2] - y[t-3]) + 0.45(y[t-3] - y[t-4])$$

$$y[2018] = 0.0053 + y[2017] + 1.64*(y[2017] - y[2016]) - 1.17(y[2016] - y[2015]) + 0.45(y[2015] - y[2014])$$
$$= 0.0053 + 8.47 + 1.64*(8.47 - 8.37) - 1.17(8.37 - 8.28) + 0.45(8.28 - 8.19)$$
$$= 8.5745$$

Similarly 2019 = 8.6747, 2020 = 8.7670

[16.e]

```
popu_switz %>%
  model(ARIMA(Population ~ pdq(3,1,0))) %>%
  forecast(h = 3) %>%
  hilo() %>%
  select(Year, .mean, "95%")
```

2018 = 8.5585, 2019 = 8.6475, 2020 = 8.7317

The forecasted numbers are slightly lower as the manually calculated numbers are using rounded off coefficients rather than the actual ones. Also, as the future forecasted numbers are based on the past forecasted values, the differences in the past values get compounded as the future values are calculated.