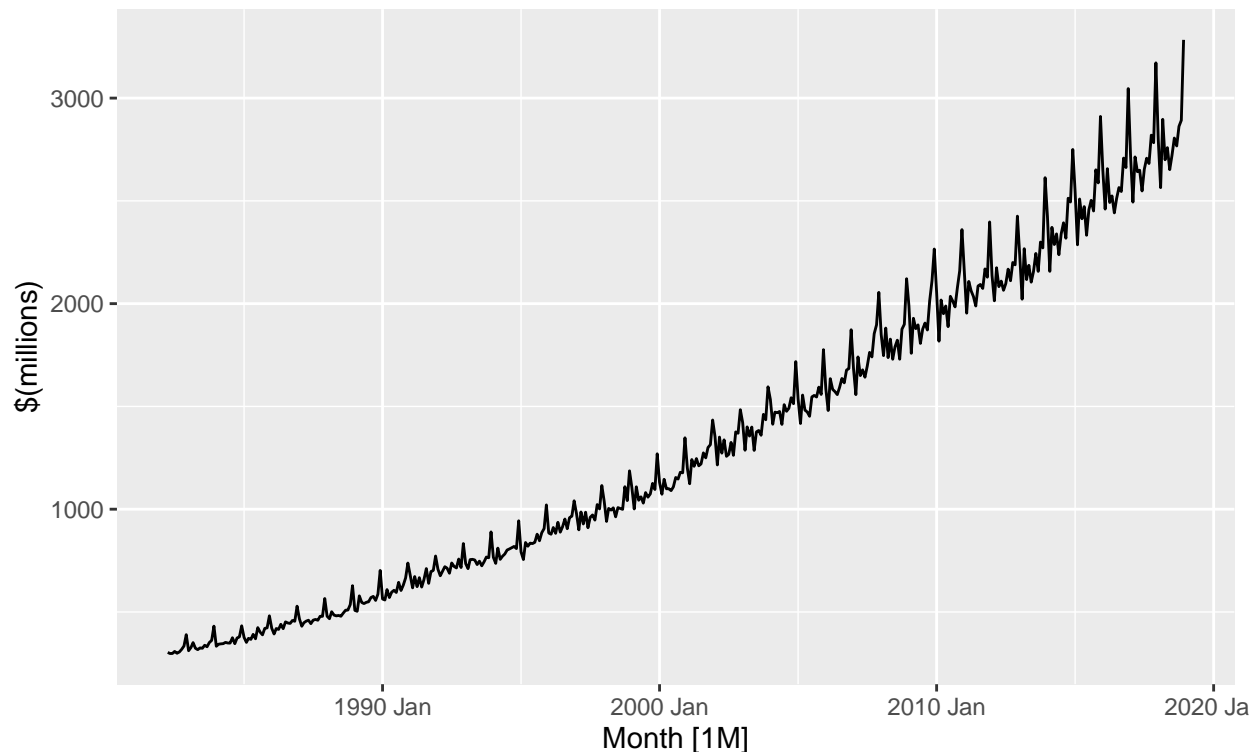# Forecasting Turnover of Markets in New South Wales

Changsoo Byun

## Turnover of Supermarket and Grocery Stores
### New South Wales



#Based on the time series plot of turnover in New South Wales supermarkets and grocery stores, a clear linear trend can be observed. The trend appears to be steadily increasing over time, indicating that an additive trend should be employed in the model. Multiplicative errors are useful when the data are strictly positive. Furthermore, the plot displays a multiplicative seasonality, as the seasonal pattern appears to increase proportionally with the level of the series. Therefore, a model with multiplicative seasonality will effectively capture and project the behavior of the seasonal pattern. (M, A, M)
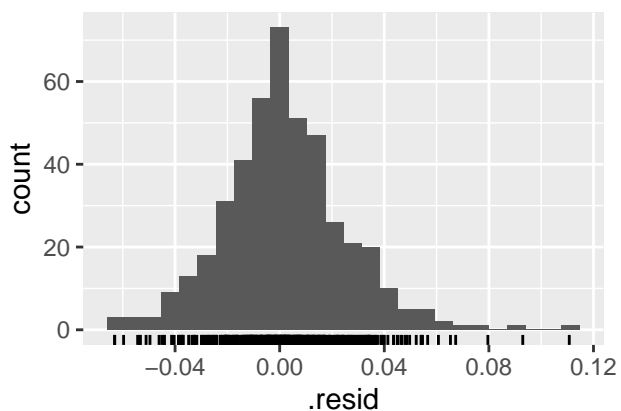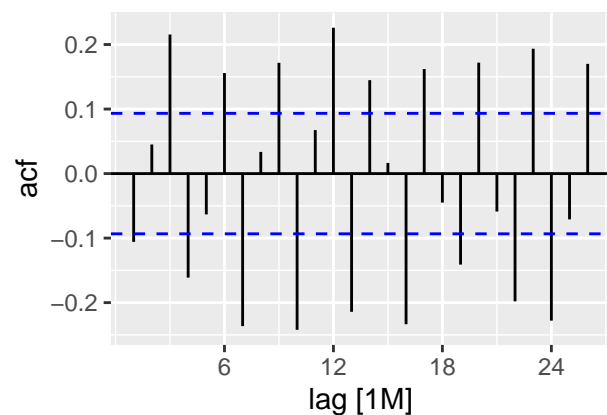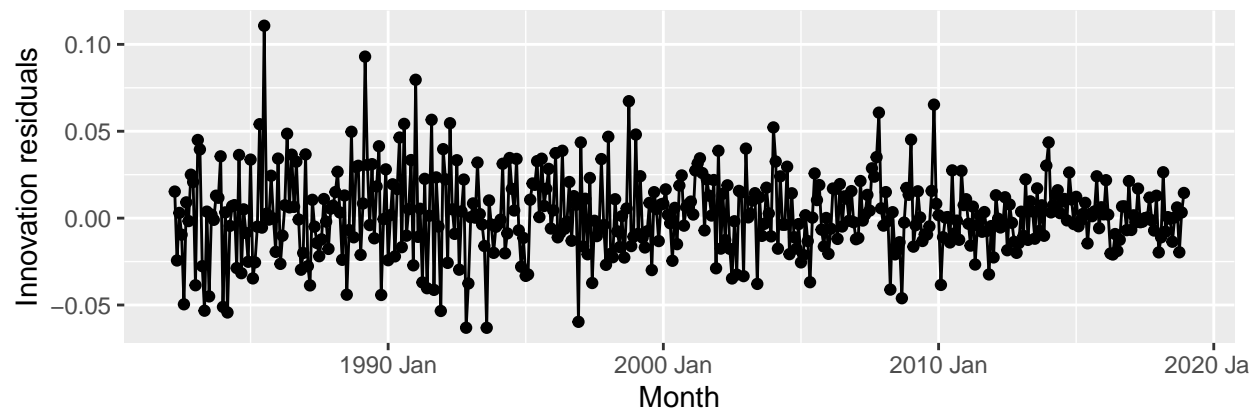
```
## Series: y
## Model: ETS(M,A,M)
##   Smoothing parameters:
##     alpha = 0.2394759
##     beta  = 0.006131541
##     gamma = 0.1417426
##
##   Initial states:
##      l[0]      b[0]      s[0]     s[-1]     s[-2]    s[-3]     s[-4]     s[-5]
##  303.8626 2.659373 1.008058 0.9526907 0.9840908 1.17473 1.018296 1.001814
##      s[-6]     s[-7]     s[-8]     s[-9]    s[-10]    s[-11]
##  0.9616378 0.9964175 0.9899021 0.9548245 0.9835982 0.9739402
##
##   sigma^2:  6e-04
##
##      AIC      AICc      BIC
## 5577.774 5579.221 5647.287
```

## ETS(M,A,M) components

y = (lag(level, 1) + lag(slope, 1)) * lag(season, 12) * (1 + remainder)
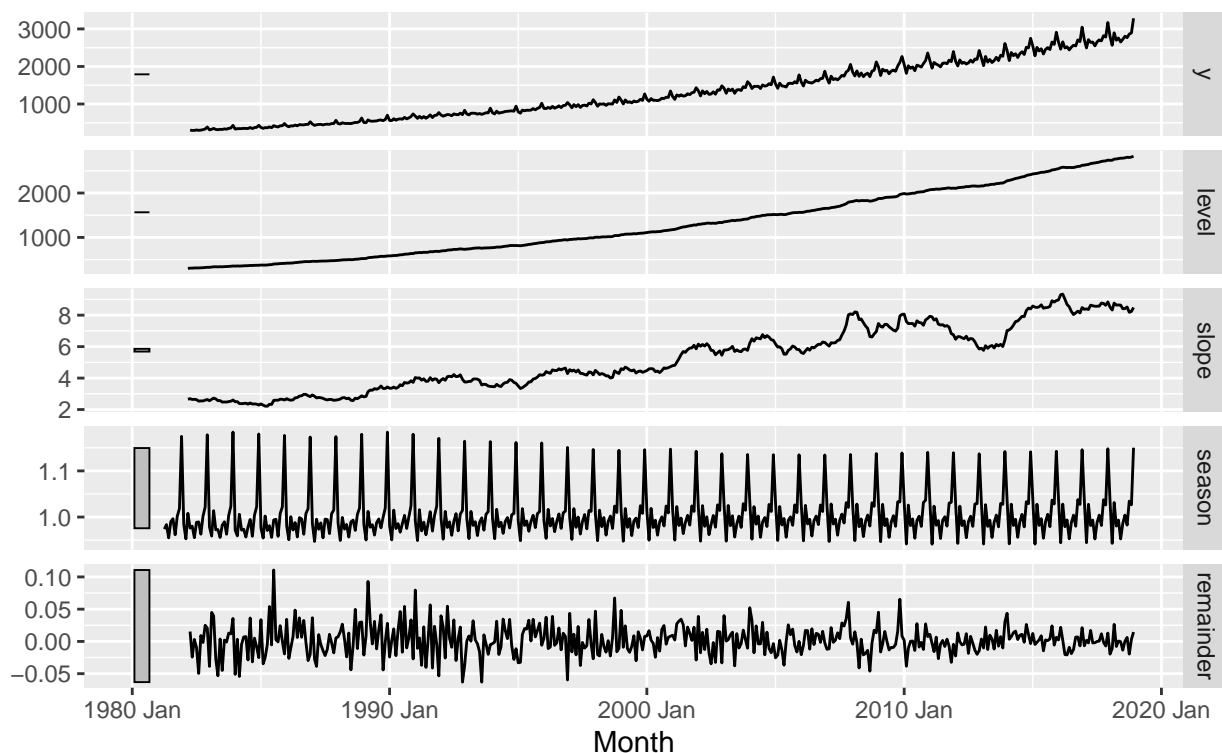


#The parameter estimates are alpha=0.2395, beta=0.0061 and gamma=0.1417. The output also returns the estimates for the intial states l[0], b[0],s[0],s[-1], s[-2], s[-3], s[-4], s[-5], s[-6], s[-7],s[-8] s[-9], s[-10] and s[-11]. The smoothing parameters is restricted to between 0 and 1, which allows the equations to be interpreted as weighted averages. The lower the AIC, AICc, and BIC values (5577.774, 5579.221, and 5647.287, respectively), the better the model fits the data.

```
## # A tibble: 1 x 3
##   .model lb_stat lb_pvalue
##   <chr>    <dbl>     <dbl>
## 1 mam       139.         0

## # A tibble: 1 x 3
##   .model bp_stat bp_pvalue
##   <chr>    <dbl>     <dbl>
## 1 mam       135.         0
```

#The null hypothesis is the spike is 0 assuming the residuals are white noise. Computed in 95% CI and 24 degrees of freedom as it is seasonal. The results of the tests are statistically significant, having enough evidence to reject the null hypothesis. Hence, residuals are not white noise.

```
fit2 <- train |>
  model(auto=ETS(y))
report(fit2)
```

```
## Series: y
## Model: ETS(M,A,M)
##   Smoothing parameters:
##     alpha = 0.2394759
##     beta  = 0.006131541
##     gamma = 0.1417426
##
##   Initial states:
##      l[0]      b[0]     s[0]     s[-1]     s[-2]   s[-3]    s[-4]     s[-5]
##  303.8626 2.659373 1.008058 0.9526907 0.9840908 1.17473 1.018296 1.001814
##      s[-6]     s[-7]     s[-8]     s[-9]    s[-10]    s[-11]
##  0.9616378 0.9964175 0.9899021 0.9548245 0.9835982 0.9739402
##
##   sigma^2:  6e-04
##
##      AIC      AICc       BIC
## 5577.774 5579.221 5647.287
```

#Model selection is using information criteria such as the AIC, AICc, and BIC. These criteria balance the trade-off between goodness-of-fit and model complexity. A lower value of these criteria suggests a better model. To illustrate this, several models, such as ETS(A,A,M), ETS(M,M,M), ETS(A,A,A), and ETS(M,A,A), could be compared to determine which one is the best fit for the data.

## Question 5

```
fit2 <- train |>
  model(
    ETS_auto = ETS(y),
    ETS_AAA = ETS(y ~ error("A") + trend("A") + season("A")),
    ETS_damped = ETS(y ~ error("A") + trend("Ad")),
    ETS_forbidden = ETS(y ~ error("A") + trend("Ad") + season("M"))
  )


glance(fit2)
```

```
## # A tibble: 4 x 9
##   .model            sigma2 log_lik   AIC  AICc   BIC   MSE  AMSE     MAE
##   <chr>              <dbl>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
## 1 ETS_auto        0.000569  -2772. 5578. 5579. 5647.  724.  783.  0.0175
## 2 ETS_AAA          960.      -2849. 5731. 5733. 5801.  925. 1019. 22.6
## 3 ETS_damped      1030.      -2864. 5763. 5765. 5837.  991. 1118. 23.3
## 4 ETS_forbidden    709.      -2781. 5599. 5600. 5672.  682.  754. 19.9
```

#I got the identical ETS model/ Among the evaluated ETS models, ETS(M,A,M) which is auto exhibits the lowest AIC score, indicating it is the most appropriate choice as the ETS model. However, ETS_forbidden shows the best fit among the other models based on MSE and has the second lowest AIC and AICc scores, which are the preferred criteria for selecting the best ETS model. Therefore, ETS(A,Ad,M) could be a plausible alternative ETS model to consider.

## Question 6

```r
fc <- fit |>
  forecast(h="2 years")

fc |>
  autoplot(
    my_series
    )+
  labs(y= "$(millions)",
       title= "Forecasts for montly turnover of supermarket and grocery stores")
```
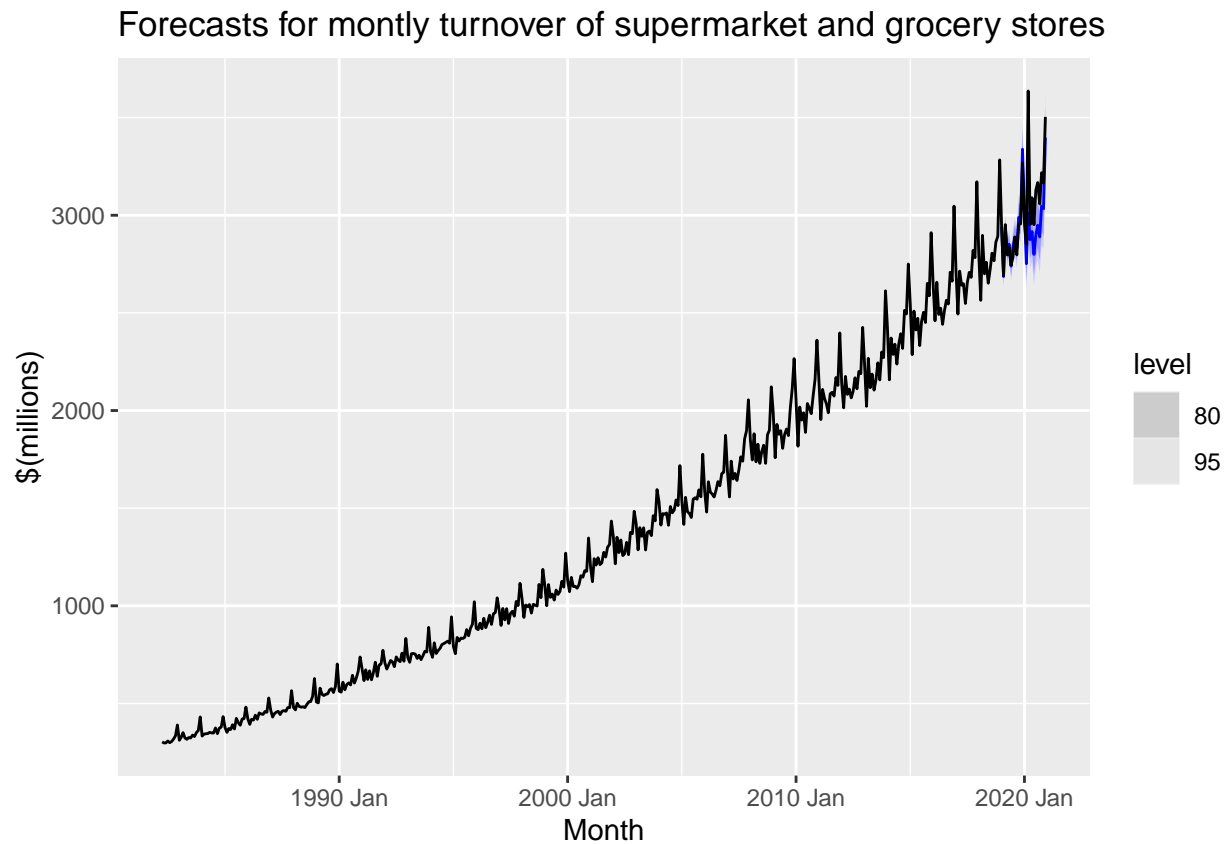
### Forecasts for montly turnover of supermarket and grocery stores
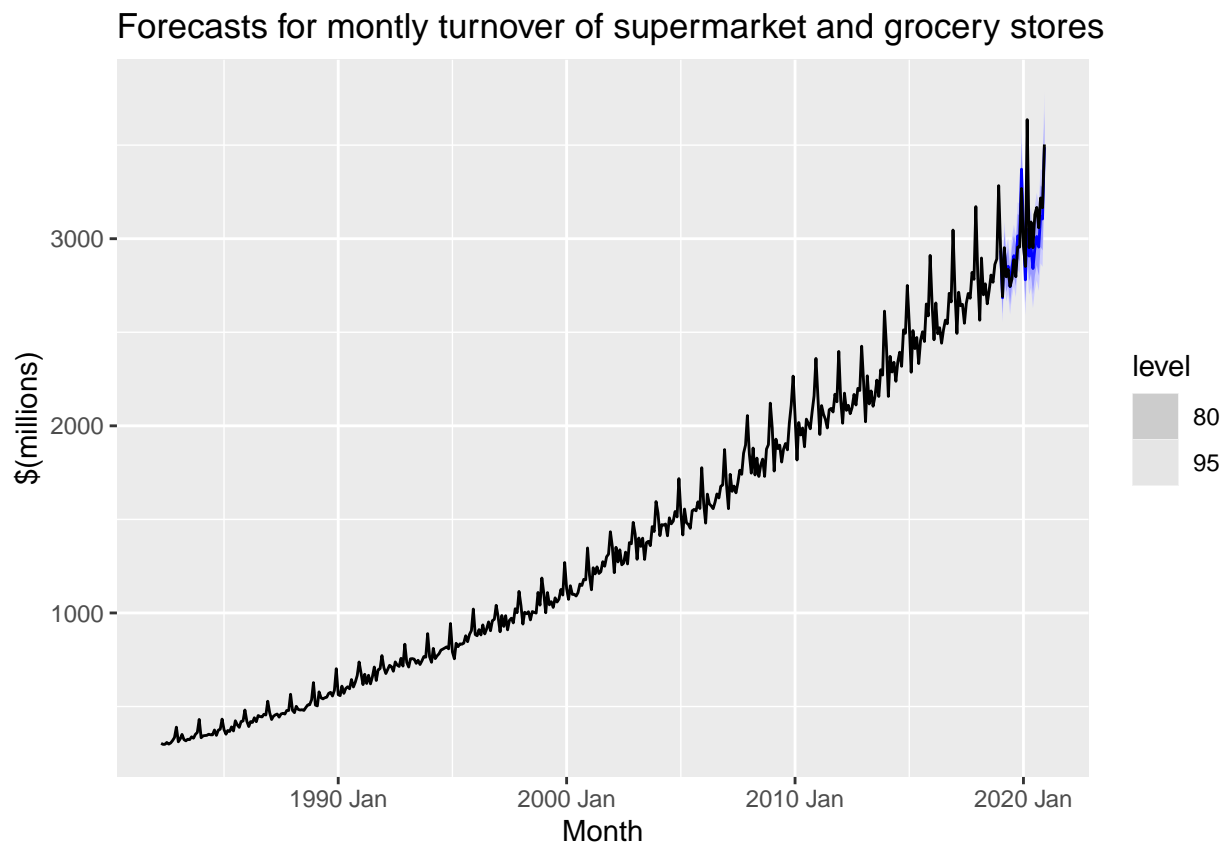


```r
fit3 <- train |>
  model(
    ETS_forbidden = ETS(y ~ error("A") + trend("Ad") + season("M"))
    )

fc2 <-  fit3 |>
  forecast(h="2 years")

fc2 |>
  autoplot(
    my_series
    )+
```

```
labs(y= "$(millions)",
     title= "Forecasts for montly turnover of supermarket and grocery stores")
```

## Forecasts for montly turnover of supermarket and grocery stores



#Both plots display the point forecasts and 80% and 95% prediction intervals for the turnover of supermarket and grocery stores in New South Wales, using two different models: ETS(M,A,M) and ETS(A,Ad,M). The point forecasts for both models appear reasonable, and the intervals are relatively narrow. However, upon closer inspection, the first plot, which represents the ETS(M,A,M) model, closely aligns with the original sample plot, suggesting that this model may provide more accurate forecasts than the ETS(A,Ad,M) model.

## Question 7

```
my_series |>
  model(ETS(y ~ error("M") + trend("A") + season("M"))) |>
  forecast(h="2 years") |>
  autoplot(my_series)
```
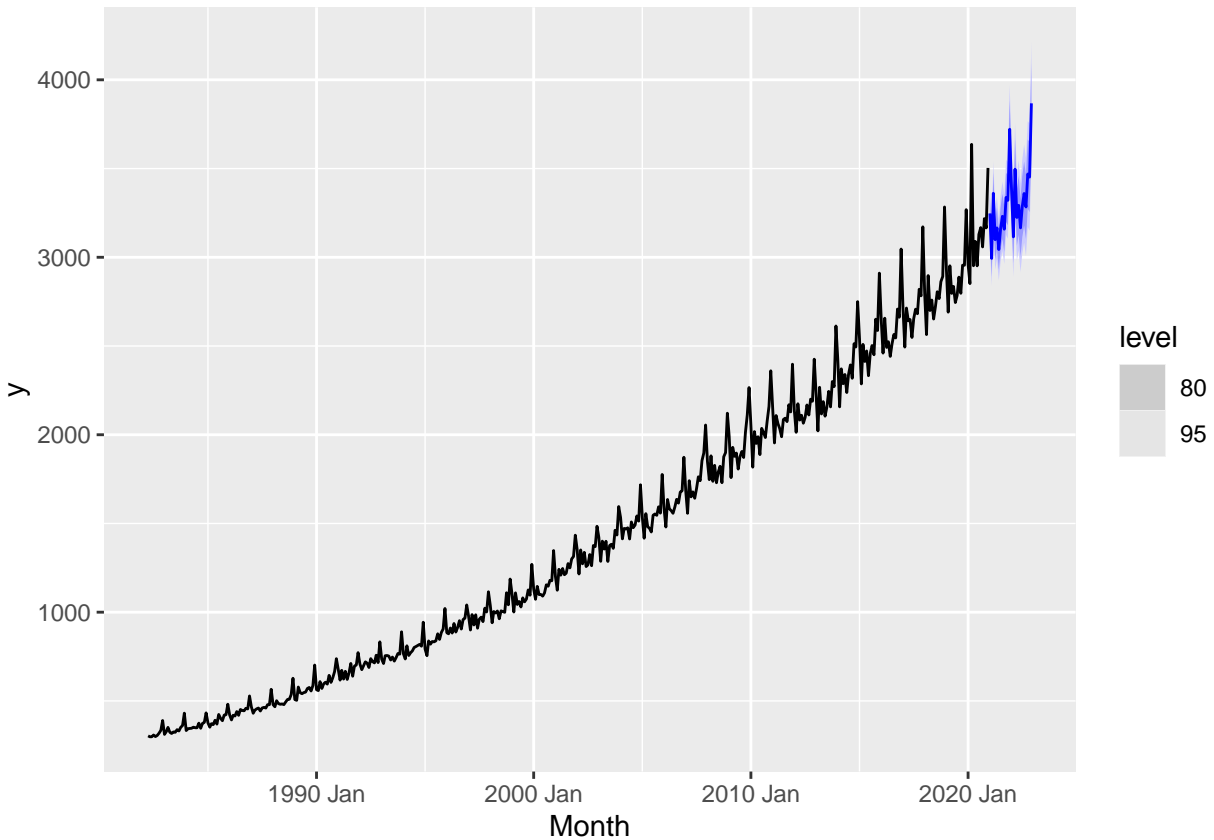


#The plot displays the predicted turnover for supermarket and grocery stores in New South Wales using an ETS(M,A,M) model, including 80% and 95% prediction intervals. The model's point forecasts appear reasonable, and the narrow intervals indicate that the model performs well in capturing the trends and seasonal patterns in the data. Furthermore, the ETS(M,A,M) model produces the minimum AIC, which suggests that it is the best model to use for accurate predictions.