The data contains the Total Sales of Appliance Units in the Philippines from Jan 2000 to Dec 2009 in PhilMonthlyData.csv

Training dataset = Jan 2000 – Dec 2007; and

Test dataset = Jan 2008 – Dec 2009.

R code

```r
library(dplyr)
library(fpp2)
library(ggplot2)
library(nortest)
#read data
ph_monthly <-
read.csv("C:/Users/Administrator/Documents/PhilMonthlyData.csv")
sales_ts <- ts(na.omit(ph_monthly$sale_app), frequency = 12, start =
c(2000,1))
#train and test dataset
train1 <- subset(sales_ts, start = 1, end = 12*8)
test1 <- subset(sales_ts, start = length(sales_ts) - 12*2+1)
```

```r
fit1 <- ets(train1)
summary(fit1)
```

The R code above gives the output:

```
ETS(M,A,M)

Call:
 ets(y = train1)

  Smoothing parameters:
    alpha = 0.5734
    beta  = 0.0139
    gamma = 2e-04

  Initial states:
    l = 597464.3559
    b = 1084.4734
    s = 1.2361 1.1346 1.0226 0.9175 0.9177 0.9557
           1.0759 1.1248 1.013 0.9696 0.8006 0.832

  sigma:  0.0714

     AIC     AICc      BIC
2463.934 2471.780 2507.528

Training set error measures:
                    ME     RMSE   MAE       MPE     MAPE      MASE        ACF1
Training set -3819.613 33707.2 26105 -1.060642 5.279745 0.3384195 -0.01535319
```

Using the ets() function, the best performing model based on the AICc of the training dataset is the multiplicative Holt-Winter's method with multiplicative errors or ETS(M, A, M) whose set of state space system of equations is given by:

$$y_t = (l_{t-1} + b_{t-1})\, s_{t-m} (1 + \varepsilon_t)$$

$$l_t = (l_{t-1} + b_{t-1})(1 + 0.5734\varepsilon_t)$$

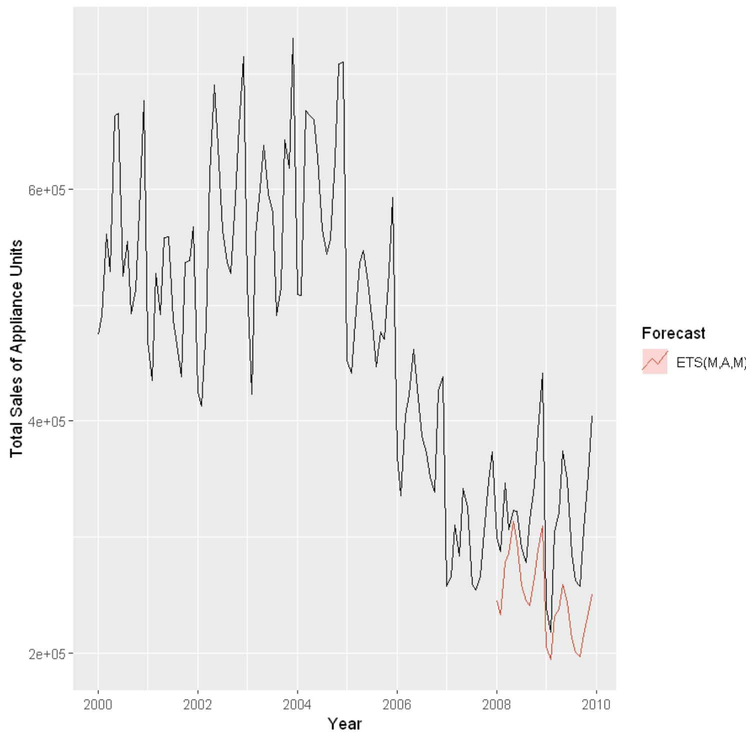
$$b_t = b_{t-1} + 0.0139(l_{t-1} + b_{t-1})\varepsilon_t$$

$$s_t = s_{t-m}(1 + 0.0002\varepsilon_t)$$

where the parameter estimates are $\hat{\alpha} = 0.5734$, $\hat{\beta} = 0.0139$, $and\ \hat{\gamma} = 0.0002$, and the initial estimates are $l_0 = 597464.3559$, $b_t = 1084.4734$, $s_t =$ (1.2361, 1.1346, 1.0226, 0.9175, 0.9177, 0.9557, 1.0759, 1.1248, 1.013, 0.9696, 0.8006, 0.832). The $\gamma\ and\ \beta$ are close to 0, so it is to be expected that the seasonal component is smooth and that the trend component is not changing over time. The value of α is closer to 1 than to 0 indicating that more weight is given to recent observations, and the level component is slightly rough.

Based on this model, the plot of the forecasted value of sale_app for the test data added into the plot of the full dataset can be obtained by using the R code:

```
fcast1 <- forecast(fit1, h = 24)
autoplot(sales_ts) + autolayer(fcast1, series="ETS(M,A,M)", PI=FALSE)
+ xlab("Year") + ylab("Total Sales of Appliance Units") +
guides(colour=guide_legend(title="Forecast"))
```



From the plot above, it seems that the forecasted values from ETS(M, A, M) model sometimes greatly deviate from the actual values of the test dataset. Take for example, the sales of appliance units in December 2008 and 2009 both exceed 40000, while the forecasted values for these respective months are around 30000 and 25000. Thus, the ETS(M, A, M) model might not be a good fit for the test dataset, but it seems to capture the seasonal pattern.

R code:

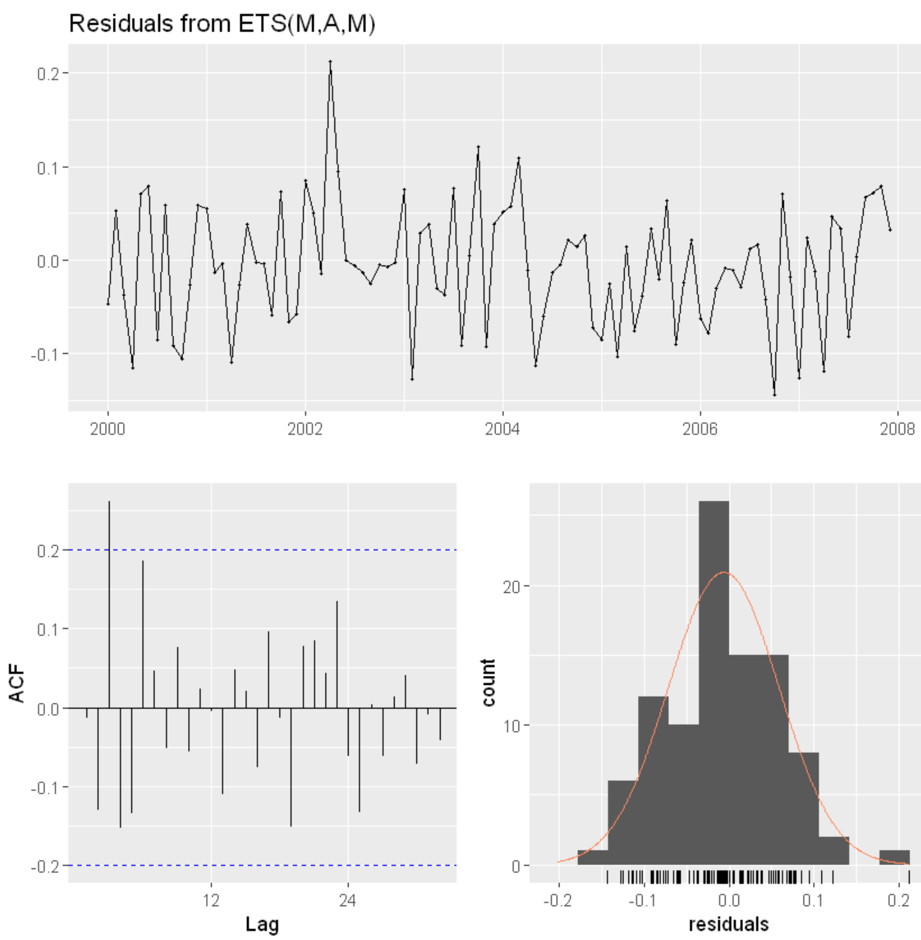
```
(acc1 <- data.frame(accuracy(fcast1, test1)) %>% select(RMSE,MAE,
MAPE, MASE) %>% slice(2))
```

The R code above gives:

RMSE	MAE	MAPE	MASE
79393.65	70063.98	21.27324	0.9082941

Above is a table of accuracy measures of ETS(M, A, M) with respect to the testing dataset from Jan 2008 to Dec 2009. The MASE of the selected model is ~0.9083, which means that this model slightly performs better than the naïve forecast model on the same dataset. The mean absolute percent error (MAPE) between the sales predicted by the ETS(M, A, M) model and the actual sales is ~21.27%, which seems to be high. The model's mean absolute error (MAE) is 70063.98, which means that, on average, the sales forecast is expected to be 70063.98 sales units away from the actual sales of appliance units.

```
checkresiduals(fit1)
nortest::ad.test(residuals(fit1))
```



Test	Test Statistic		*p*-value
Ljung-Box test	Q*	24.445	2.017e-05
Anderson-Darling normality test	A	0.37848	0.4002

From the ETS(M, A, M) residual diagnostics histogram, the mean of the residuals seems to be centered at 0, and the distribution of the histogram does not seem to be skewed. To formally test for the normality of the residuals, an Anderson-Darling test was conducted. Since the *p*-value is greater than .05, then there is enough evidence to conclude that there is no significant departure from normality. The time plot shows that the variation of the residuals stays the same across the historical data, which means that it can be assumed that the variance of the residuals is constant.

There is only one significant spike in the ACF, but to formally test for autocorrelation, a Ljung-Box Test was conducted. The table above shows the computed test statistics and *p*-value of the Ljung-Box test. Since the test has a *p*-value less than .05, it can be concluded that the residuals do not resemble white noise. In other words, the model has some significant autocorrelation in the residuals, which suggests that the model can still be used for forecasting, but the prediction intervals may not be accurate. Possible remedy is to consider other models or other ETS models which might pass all of these tests, but doing this might result to a higher AICc, which means that the ETS model produced is not the best performing model based on the AICc of the training dataset. Another remedy is to perform a Box-Cox transformation which might lead to an ETS model with residuals satisfying the properties of stable residual variance and residuals not distinguishable from a white noise.

The data used for this section contains the Quarterly Personal Consumption Expenditure, in Million Pesos from Q1 1981 to Q4 2008) in PhilQuarterData.csv. Split the data into training and test data set:

Training dataset = Q1 1981 – Q4 2005; and

Test dataset = Q1 2006 – Q4 2008.

R code

```
#read data
ph_quarterly <-
read.csv("C:/Users/Administrator/Documents/PhilQuarterData.csv")
pce_ts <- ts(na.omit(ph_quarterly$pce), frequency = 4, start =
c(1981,1))
#train and test dataset
train2 <- subset(pce_ts, start = 1, end = 25*4)
test2 <- subset(pce_ts, start = length(pce_ts) - 3*4+1)
```

```
fit2 <- ets(train2)
summary(fit2)
```

The R code above gives the output:

```
ETS(A,A,A)

Call:
 ets(y = train2)

  Smoothing parameters:
    alpha = 0.3154
    beta  = 0.1138
    gamma = 0.5858

  Initial states:
    l = 101255.6841
    b = 548.7882
    s = 14442.71 -2683.739 -980.3079 -10778.66

  sigma:  1767.53

     AIC     AICc      BIC
1965.647 1967.647 1989.093

Training set error measures:
                   ME     RMSE      MAE       MPE      MAPE      MASE      ACF1
Training set 212.7333 1695.355 1343.578 0.1213425 0.9932114 0.2330713 0.1411876
```

Using the ets() function, the best performing model based on the AICc of the training dataset is the ETS(A, A, A) whose set of state space system of equations is given by:

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$$

$$l_t = l_{t-1} + b_{t-1} + 0.3154\varepsilon_t$$

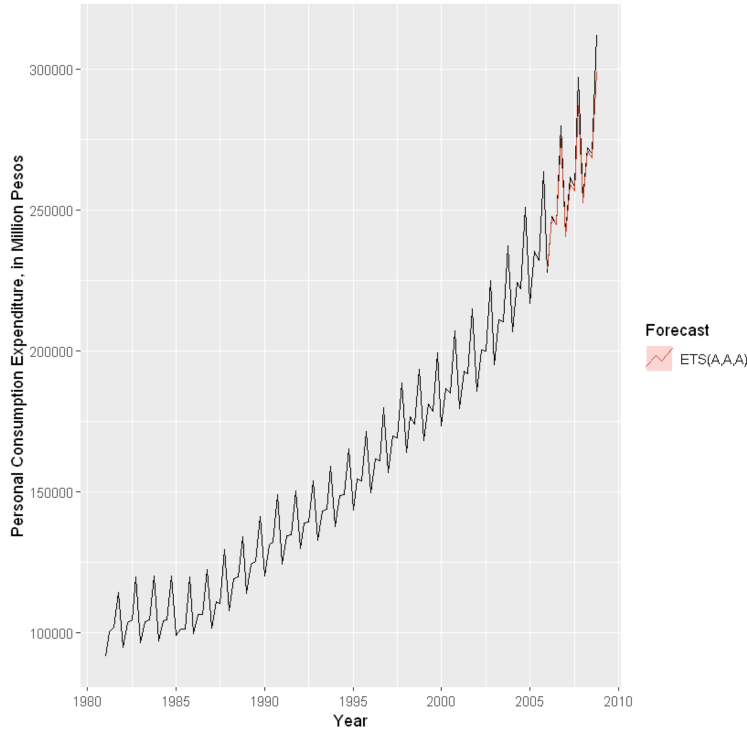
$$b_t = b_{t-1} + 0.1138\varepsilon_t$$

$$s_t = s_{t-m} + 0.5858\varepsilon_t$$

where the parameter estimates are $\widehat{\alpha} = 0.3154,\ \widehat{\beta} = 0.1138,\ and\ \widehat{\gamma} = 0.5858$, and the initial estimates are $l_0 = 101255.6841,\ b_t = 548.7882, s_t =$ (14442.71, -2683.739, -980.3079, -10778.66). The β is close to 0, so it is to be expected that the trend component is not changing over time. The value of α is closer to 0 than to 1 indicating that less weight is given to recent observations, and the changes in the level are slightly smooth. The value of γ is close to 1-α, which means that the seasonal component is rough.

Based on this model, the plot of the forecasted value of pce for the test data added into the plot of the full dataset can be obtained by using the R code:

```
fcast2 <- forecast(fit2, h = 12)
autoplot(pce_ts) +
  autolayer(fcast2, series="ETS(A,A,A)", PI=FALSE) +
  xlab("Year") + ylab("Personal Consumption Expenditure, in Million
Pesos") +
  guides(colour=guide_legend(title="Forecast"))
```



From the plot above, it seems that the forecast values of the ETS(A, A, A) model are a close match to the test data which contains the quarterly personal consumption expenditure, in million pesos, from Q1 2006 to Q4 2008. The model seems to capture the quarterly seasonal pattern and the increasing trend at the end of the data.

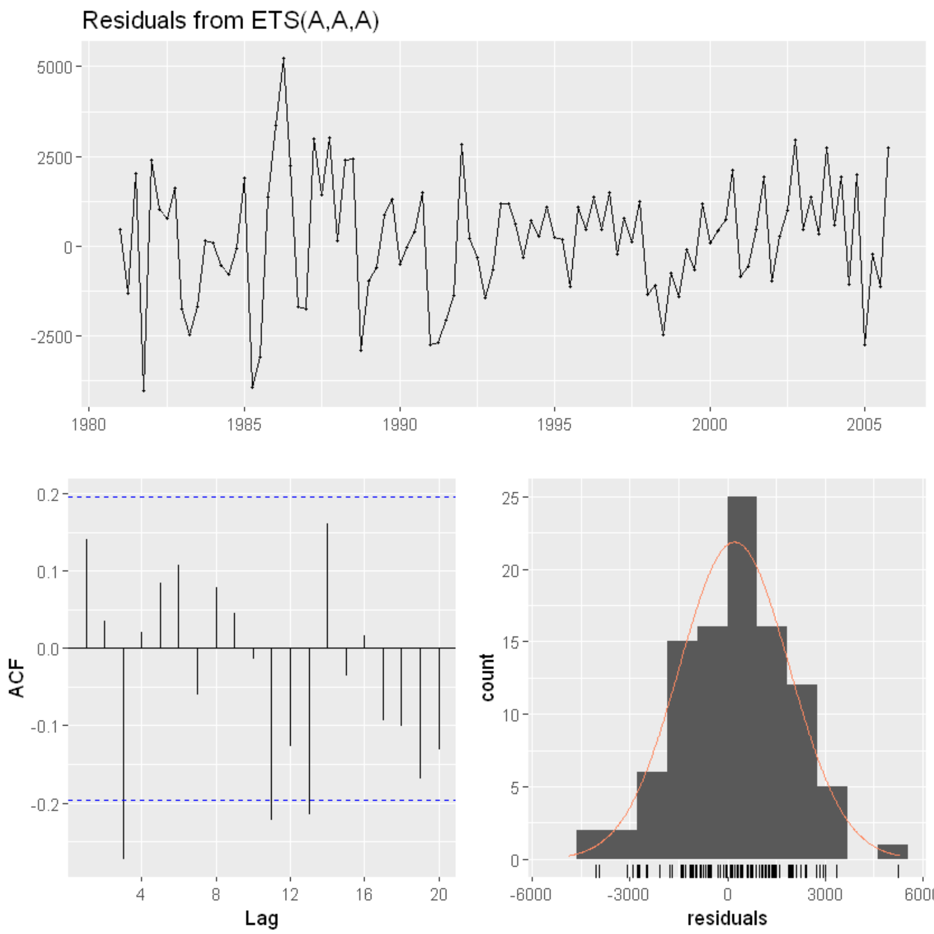
```
(acc2 <- data.frame(accuracy(fcast2, test2)) %>% select(RMSE,MAE,
MAPE, MASE) %>% slice(2))
```

The R code above gives:

RMSE	MAE	MAPE	MASE
5061.439	3178.63	1.1066	0.5514

Above is a table of accuracy measures of ETS(A, A, A) with respect to the testing dataset from Q1 2006 to Q4 2008. The MASE of the selected model is ~0.5514, which means that this model performs better than the naïve forecast model on the same dataset. The mean absolute percent error (MAPE) between the expenditure predicted by the ETS(A, A, A) model and the actual expenditure is ~1.11%, which seems to be small and good enough. The model's mean absolute error (MAE) is 3178.63, which means that, on average, the personal consumption expenditure forecast is expected to be 3178.63 million pesos away from the actual personal consumption expenditure value.

```r
checkresiduals(fit2)
nortest::ad.test(residuals(fit2))
```



Test	Test Statistic		*p*-value
Ljung-Box test	Q*	19.018	0.0002711
Anderson-Darling normality test	A	0.22155	0.8264

From the ETS(A, A, A) residual diagnostics histogram, the mean of the residuals seems to be centered at 0, and the distribution of the histogram does not seem to be skewed. To formally test for the normality of the residuals, an Anderson-Darling test was conducted. Since the *p*-value is greater than .05, then there is enough evidence to conclude that there is no significant departure from normality. The time plot shows that the variation of the residuals does not seem to be the same across the historical data, which means that it cannot be assumed that the variance of the residuals is constant. Box-Cox transformation on the training dataset might be performed to stabilize the variation of the residuals across the time plot.

There are three significant spikes in the ACF, but to formally test for autocorrelation, a Ljung-Box Test was conducted. The table above shows the computed test statistics and *p*-value of the Ljung-Box test. Since the test has a *p*-value less than .05, it can be concluded that the residuals do not resemble white noise. In other words, the model has some significant autocorrelation in the residuals, which suggests that the model can still be used for forecasting, but the prediction intervals may not be accurate. Possible remedy is to consider other models or other ETS models which might pass all of these tests, but doing this might result to a higher AICc, which means that the ETS model produced is not the best performing model based on the AICc of the training dataset. Another remedy is to perform a Box-Cox transformation which might lead to an ETS model with residuals satisfying the properties of stable residual variance and residuals not distinguishable from a white noise.