Case Study #2: Forecasting Walmart's Revenue

Case Solutions

The data set for case study #2 represents quarterly revenues (in $million) in Walmart from the first quarter of 2005 through the first quarter of 2022 (*673_case2.csv*). This quarterly data is collected from *www.macrotrends.net/stocks/charts/WMT/walmart/revenue*. The goal is to forecast Walmart's quarterly revenue in the four quarters (Q1-Q4) of 2023 and 2024.

Questions

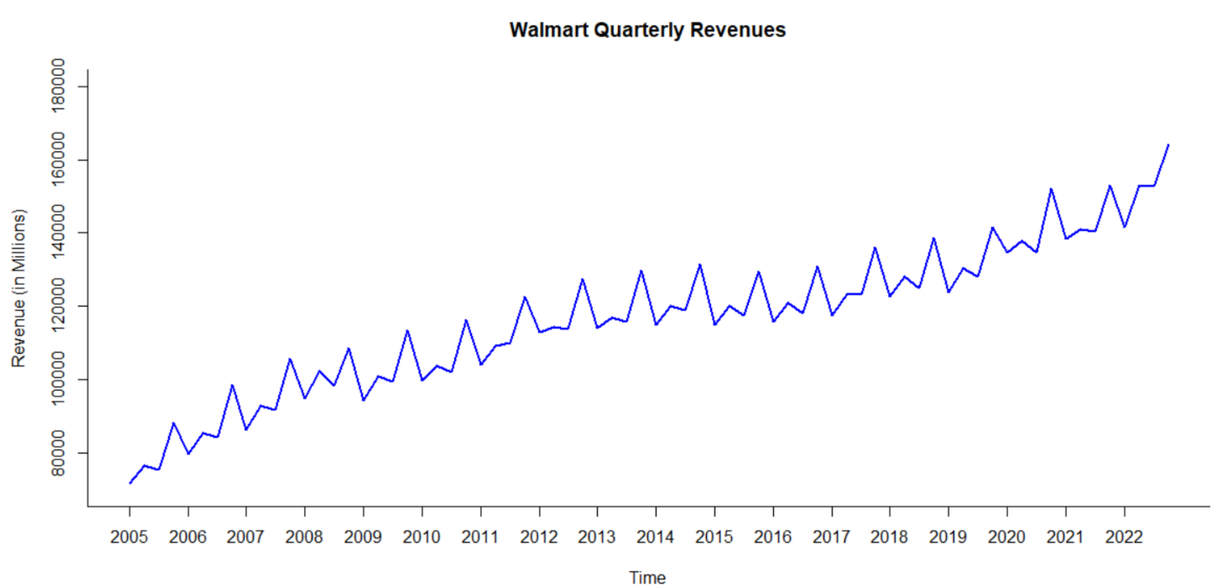
1. **Plot the data and visualize time series components.**

1a. Create time series data set in R using the *ts()* function.

(The answers for this question can vary). We loaded into R, using the *read.csv()* function, the *673_case2.csv* data, and then created a time series data *revenue.ts* for Walmart's revenue using the ts() function from Q1 of 2005 through Q4 of 2022, with the frequency of 4 for quarterly data.

1b. Apply the *plot()* function to create a data plot with the historical data, provide it in your report, and explain what time series components can be visualized in this plot.

The plot of the of Walmart's revenue is shown below.



The plot shows potentially an upward non-linear (but also possibly linear) trend pattern. The plot also shows an additive seasonal pattern with the high revenue at the end of each year (Q4) and low revenue at the beginning of the year (approximately in Q1-Q2).

2. Apply five regression models using data partition.

Consider the following 5 regression-based models:

i. Regression model with linear trend
ii. Regression mode with quadratic trend
iii. Regression model with seasonality
iv. Regression model with linear trend and seasonality
v. Regression model with quadratic trend and seasonality.

2a. Develop data partition with the validation partition of 16 periods and the rest for the training partition.

The training period (*train.ts*) and validation period (*valid.ts*) time series partitions of 56 and 16 records, respectively, are listed below.

```
train.ts
       Qtr1   Qtr2   Qtr3   Qtr4
2005  71680  76697  75397  88327
2006  79676  85430  84467  98795
2007  86410  92999  91865 105749
2008  94940 102342  98345 108627
2009  94242 100876  99373 113594
2010  99811 103726 101952 116360
2011 104189 109366 110226 122728
2012 113010 114282 113800 127559
2013 114070 116830 115688 129706
2014 114960 120125 119001 131565
2015 114826 120229 117408 129667
2016 115904 120854 118179 130936
2017 117542 123355 123179 136267
2018 122690 128028 124894 138793
```

```
valid.ts
       Qtr1   Qtr2   Qtr3   Qtr4
2019 123925 130377 127991 141671
2020 134622 137742 134708 152079
2021 138310 141048 140525 152871
2022 141569 152859 152813 164048
```

2b. b) Use the *tslm()* function for the training partition to develop each of the 5 regression models from the above list. Apply the *summary()* function to identify the model structure and parameters for each regression model, show them in your report, and also present the respective model equation and define its predictors. Briefly explain if the model is a good fit, statistically significant, and thus may be applied for forecasting. Use each model to forecast revenues for the validation period using the *forecast()* function, and present this forecast in your report.

We developed the 5 regression models identified in question 2.

1. Regression Model with Linear Trend

The model using the *summary()* function is presented below:

```
Call:
tslm(formula = train.ts ~ trend)

Residuals:
     Min       1Q   Median       3Q      Max 
-13110.7  -5193.0     75.3   3564.5  14923.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 83892.49    1885.24   44.50   <2e-16 ***
trend         898.22      57.54   15.61   <2e-16 ***
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6960 on 54 degrees of freedom
Multiple R-squared:  0.8186, Adjusted R-squared:  0.8152
F-statistic: 243.7 on 1 and 54 DF,  p-value: < 2.2e-16
```

The *regression model with linear trend* contains a single independent variable: period index (*t*). The model's equation is:

$$y_t = 83892.49 + 898.22\ t$$

According to the model summary, the *regression model with linear trend* is statistically significant. It has a high R-squared of 0.8186 and adjusted R_squared of 0.8152, which represents a good fit for the training data. The intercept and coefficient for the trend (t) variable are statistically significant (p-values are much lower than 0.05 or 0.01). Therefore, this model <u>may be used for time series forecasting</u>.

The forecast for the training period is the following:

```
        Point Forecast     Lo 0     Hi 0
2019 Q1       135090.9 135090.9 135090.9
2019 Q2       135989.2 135989.2 135989.2
2019 Q3       136887.4 136887.4 136887.4
2019 Q4       137785.6 137785.6 137785.6
2020 Q1       138683.8 138683.8 138683.8
2020 Q2       139582.0 139582.0 139582.0
2020 Q3       140480.2 140480.2 140480.2
2020 Q4       141378.5 141378.5 141378.5
2021 Q1       142276.7 142276.7 142276.7
2021 Q2       143174.9 143174.9 143174.9
2021 Q3       144073.1 144073.1 144073.1
2021 Q4       144971.3 144971.3 144971.3
2022 Q1       145869.6 145869.6 145869.6
2022 Q2       146767.8 146767.8 146767.8
2022 Q3       147666.0 147666.0 147666.0
2022 Q4       148564.2 148564.2 148564.2
```

2. Regression Model with Quadratic Trend

The model using the *summary()* function is presented below:

```
Call:
tslm(formula = train.ts ~ trend + I(trend^2))

Residuals:
   Min     1Q Median     3Q    Max 
 -8242  -3866  -1776   5259  12107 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 75536.325   2497.172  30.249  < 2e-16 ***
trend        1762.649    202.132   8.720 8.05e-12 ***
I(trend^2)    -15.165      3.437  -4.412 5.06e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6008 on 53 degrees of freedom
Multiple R-squared:  0.8673, Adjusted R-squared:  0.8623
F-statistic: 173.2 on 2 and 53 DF,  p-value: < 2.2e-16
```

The *regression model with quadratic trend* contains two independent variables: period index (*t*) and squared period index (t^2). The model's equation is:

$$y_t = 75536.33 + 1762.65\ t - 15.17\ t^2$$

According to the model summary, the *regression model with quadratic trend* is statistically significant. It has a high R-squared of only 0.8673 (adj. R_squared is 0.8623), which is a good fit for the training data.

The coefficients for the trend (t) and quadratic trend (t^2) variables are statistically significant (p-values for both coefficients are much lower than 0.05 or 0.01). This model may also be used for time series forecasting.

The forecast for the training period is the following:

```
        Point Forecast     Lo 0     Hi 0
2019 Q1       126734.8 126734.8 126734.8
2019 Q2       126753.4 126753.4 126753.4
2019 Q3       126741.7 126741.7 126741.7
2019 Q4       126699.6 126699.6 126699.6
2020 Q1       126627.3 126627.3 126627.3
2020 Q2       126524.6 126524.6 126524.6
2020 Q3       126391.5 126391.5 126391.5
2020 Q4       126228.2 126228.2 126228.2
2021 Q1       126034.5 126034.5 126034.5
2021 Q2       125810.5 125810.5 125810.5
2021 Q3       125556.1 125556.1 125556.1
2021 Q4       125271.4 125271.4 125271.4
2022 Q1       124956.4 124956.4 124956.4
2022 Q2       124611.0 124611.0 124611.0
2022 Q3       124235.4 124235.4 124235.4
2022 Q4       123829.4 123829.4 123829.4
```

3. Regression Model with Seasonality

The model using the *summary()* function is presented below:

```
Call:
tslm(formula = train.ts ~ season)

Residuals:
   Min     1Q Median     3Q    Max 
-31578  -8489   4793  11720  19804 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   103139       4094  25.190  < 2e-16 ***
season2         5085       5790   0.878  0.38389    
season3         3559       5790   0.615  0.54149    
season4        16766       5790   2.896  0.00552 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15320 on 52 degrees of freedom
Multiple R-squared:  0.1536, Adjusted R-squared:  0.1048 
F-statistic: 3.146 on 3 and 52 DF,  p-value: 0.03275
```

The *regression model with seasonality* contains 3 independent seasonal dummy variables for Q2 (season2 – D_2), Q3 (season3 – D_3) and Q4 (season4 – D_4). The model's equation is:

$$y_t = 103139 + 5085\, D_2 + 3559\, D_3 + 16766\, D_4$$

The model's summary shows a very low R-squared of 0.1536 and even lower adjusted R_squared of 0.1048, and two regression coefficients for the seasonal variables are statistically insignificant. However, the coefficient for D_4 will be statistically significant for p-value < 0.05). Overall, this regression model is not a good fit, and thus cannot be applied for time series forecasting.

4. Regression Model with Linear Trend and Seasonality

The model using the *summary()* function is presented below:

```
Call:
tslm(formula = train.ts ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-8721.9 -3092.6   687.4  3165.2  8112.5 
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 79403.89    1506.15  52.720  < 2e-16 ***
trend         879.09      35.81  24.549  < 2e-16 ***
season2      4205.84    1633.58   2.575    0.013 *  
season3      1800.68    1634.76   1.101    0.276    
season4     14128.66    1636.72   8.632 1.51e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4321 on 51 degrees of freedom
Multiple R-squared:  0.934,  Adjusted R-squared:  0.9288 
F-statistic: 180.3 on 4 and 51 DF,  p-value: < 2.2e-16
```

The regression model contains 4 independent variables: trend index (t) and 3 seasonal dummy variables for Q2 (season2 – D_2), Q3 (season3 – D_3) and Q4 (season4 – D_4). The model's equation is:

$$y_t = 79403.89 + 879.09\, t + 4205.84\, D_2 + 1800.68\, D_3 + 14128.66\, D_4$$

The model's summary shows a very high R-squared of 0.934 and adj. R-squared of 0.9288, which is a very good fir for the training data, statistically, and regression coefficients for D_2 and D_4 are statistically significant (p-value <0.01 or 0.05). Although, the regression coefficient for D_3 is statistically insignificant for p-value > 0.05, overall, this regression model is a very good fit and statistically significant, and thus can be applied for time series forecasting.

The forecast for the training period is the following:

```
        Point Forecast     Lo 0     Hi 0
2019 Q1       129511.9 129511.9 129511.9
2019 Q2       134596.9 134596.9 134596.9
2019 Q3       133070.8 133070.8 133070.8
2019 Q4       146277.9 146277.9 146277.9
2020 Q1       133028.3 133028.3 133028.3
2020 Q2       138113.2 138113.2 138113.2
2020 Q3       136587.2 136587.2 136587.2
2020 Q4       149794.2 149794.2 149794.2
2021 Q1       136544.7 136544.7 136544.7
2021 Q2       141629.6 141629.6 141629.6
2021 Q3       140103.5 140103.5 140103.5
2021 Q4       153310.6 153310.6 153310.6
2022 Q1       140061.0 140061.0 140061.0
2022 Q2       145145.9 145145.9 145145.9
2022 Q3       143619.9 143619.9 143619.9
2022 Q4       156826.9 156826.9 156826.9
```

5. Regression Model with Quadratic Trend and Seasonality

The model using the *summary()* function is presented below:

```
Call:
tslm(formula = train.ts ~ trend + I(trend^2) + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-3583.3 -1950.1   232.7  1443.6  5664.9 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 71042.26    1042.04  68.176  < 2e-16 ***
trend        1745.66      74.67  23.379  < 2e-16 ***
I(trend^2)    -15.20       1.27 -11.974 2.67e-16 ***
season2      4175.43     838.90   4.977 8.04e-06 ***
season3      1770.27     839.51   2.109     0.04 *  
season4     14128.66     840.51  16.810  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2219 on 50 degrees of freedom
Multiple R-squared:  0.9829, Adjusted R-squared:  0.9812 
F-statistic: 575.7 on 5 and 50 DF,  p-value: < 2.2e-16
```

The *regression model with quadratic trend and seasonality* contains 5 independent variables: trend index (t), squared trend index (t^2), and 3 seasonal dummy variables for Q2 (season2 – D_2), Q3 (season3 – D_3) and Q4 (season4 – D_4). The model's equation is:

$$y_t = 71042.26 + 1745.66\ t - 15.20\ t^2 + 4175{,}43\ D_2 + 1770.27\ D_3 + 14128.66\ D_4$$

The model summary shows a very high R-squared of 0.9829 and adj. R-squared of 0.9812, and regression coefficients for D_2 and D_4 being statistically significant (p-value < 0.01). The regression coefficient for D_3 is statistically significant for p-value of 0.05. Overall, this regression model is a very good fit and thus can be applied for time series forecasting.

The forecast for the training period is the following:

```
        Point Forecast     Lo 0     Hi 0
2019 Q1       121150.3 121150.3 121150.3
2019 Q2       125323.1 125323.1 125323.1
2019 Q3       122884.8 122884.8 122884.8
2019 Q4       135179.7 135179.7 135179.7
2020 Q1       120957.2 120957.2 120957.2
2020 Q2       125008.3 125008.3 125008.3
2020 Q3       122448.4 122448.4 122448.4
2020 Q4       134621.7 134621.7 134621.7
2021 Q1       120277.5 120277.5 120277.5
2021 Q2       124207.0 124207.0 124207.0
2021 Q3       121525.5 121525.5 121525.5
2021 Q4       133577.1 133577.1 133577.1
2022 Q1       119111.3 119111.3 119111.3
2022 Q2       122919.2 122919.2 122919.2
2022 Q3       120116.1 120116.1 120116.1
2022 Q4       132046.1 132046.1 132046.1
```

2c. Apply the *accuracy()* function to compare performance measure of the 5 forecasts you developed in 2b. Present the accuracy measures in your report, compare them, and, using MAPE and RMSE, identify the two most accurate regression models for forecasting.

The accuracy measures for the 5 regression models are presented below:

```
> Regression Model with Linear Trend
              ME     RMSE      MAE    MPE  MAPE  ACF1 Theil's U
Test set -130.199 7224.999 6281.152 -0.444 4.411 0.017     0.721

> Regression Model with Quadratic Trend
              ME    RMSE      MAE    MPE   MAPE ACF1 Theil's U
Test set 15884.52 19446.6 16235.75 10.695 10.979 0.43     2.061

> Regression Model with Seasonality
              ME     RMSE      MAE    MPE   MAPE  ACF1 Theil's U
Test set 32205.66 33186.73 32205.66 22.519 22.519 0.764     3.517

> Regression Model with Linear Trend and Seasonality
             ME     RMSE    MAE   MPE  MAPE  ACF1 Theil's U
Test set 558.466 4399.761 3404.1 0.216 2.383 0.667     0.448

> Regression Model with Quadratic Trend and Seasonality
             ME     RMSE      MAE    MPE   MAPE  ACF1 Theil's U
Test set 16612.79 18942.93 16612.79 11.383 11.383 0.775     1.992
```

Based on the lowest values of MAPE and RMSE accuracy measures for the validation period, the best model is *regression model with linear trend and seasonality* (MAPE and RMSE for the validation period forecast are 2.38% and 4399.76, respectively). This model is selected as the first most accurate model for forecasting. The *regression model with linear trend* has second lowest MAPE of 4.41% and RMSE of 7225.00, and is selected as the second best model for forecasting. The *regression model with quadratic trend and seasonality* has the MAPE of 11.38% and RMSE of 16612.79, which are close to those received by the *regression model with the quadratic trend*. However, the *regression model with quadratic trend* *and seasonality* better reflects the existing trend and seasonality components in the historical data, and thus we select the latter model as the third best model to be considered further for the entire data set.

3. Employ the entire data set to make time series forecast.

3a. Apply the three most accurate regression models identified in 2c to make the forecast in the four quarters (Q1-Q4) of 2023 and 2024. For that, use the entire data set to develop the regression model using the *tslm()* function. Apply the *summary()* function to identify the model structure and parameters, show them in your report, and also present the respective model equation and define its predictors. Briefly explain if the model is a good fit, statistically significant, and thus may be applied for forecasting. Use each model to forecast Walmart's revenue in Q1-Q4 of 2023 and 2024 using the *forecast()* function, and present this forecast in your report.

1. Regression Model with Linear Trend and Seasonality

For forecasting the Walmart's revenue in 2023-2024, we apply for the entire data set the *regression model with linear trend and seasonality*, which was identified to be the first most accurate model in question *2c.* The *regression model with linear trend and seasonality* for the entire data set is presented below:

```
Call:
tslm(formula = revenue.ts ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-8399.4 -3469.4   742.8  3020.8  8475.4 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 78776.90    1332.46  59.121  < 2e-16 ***
trend         895.86      24.54  36.503  < 2e-16 ***
season2      4370.20    1440.77   3.033  0.00344 ** 
season3      1954.67    1441.39   1.356  0.17962    
season4     14366.09    1442.44   9.960 7.55e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4322 on 67 degrees of freedom
Multiple R-squared:  0.9569, Adjusted R-squared:  0.9544 
F-statistic: 372.3 on 4 and 67 DF,  p-value: < 2.2e-16
```

This *regression model with linear trend and seasonality* contains 4 independent variables: trend index (t) and 3 seasonal dummy variables for Q2 (season2 – D_2), Q3 (season3 – D_3) and Q4 (season4 – D_4). The regression equation is:

$$y_t = 78776.90 + 895.86\, t + 4370.20\, D_2 + 1954.67\, D_3 + 14366.09\, D_4$$

The model's summary shows a high R-squared of 0.9569 and adj. R-squared of 0.9544, and all regression coefficients being statistically significant (p-value <0.01 or 0.05) except for the regression coefficient for D_3. Overall, this regression model is a very good fit and can be used for forecasting Walmart's revenue in 2023-2024. The appropriate forecast is shown below:

```
        Point Forecast     Lo 0     Hi 0
2023 Q1       144174.7 144174.7 144174.7
2023 Q2       149440.7 149440.7 149440.7
2023 Q3       147921.1 147921.1 147921.1
2023 Q4       161228.3 161228.3 161228.3
2024 Q1       147758.1 147758.1 147758.1
2024 Q2       153024.2 153024.2 153024.2
2024 Q3       151504.5 151504.5 151504.5
2024 Q4       164811.8 164811.8 164811.8
```

2. Regression Model with Linear Trend

For forecasting the Walmart's revenue in 2023-2024, we apply for the entire data set the *regression model with linear trend*, which was identified to be the second best model in question 2c. The *regression model with linear trend* for the entire data set is presented below:

```
Call:
tslm(formula = revenue.ts ~ trend)

Residuals:
     Min       1Q   Median       3Q      Max 
-12747.5  -5174.8   -885.5   3998.7  15178.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  83519.9     1671.6   49.96   <2e-16 ***
trend          907.6       39.8   22.81   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7018 on 70 degrees of freedom
Multiple R-squared:  0.8814, Adjusted R-squared:  0.8797 
F-statistic: 520.1 on 1 and 70 DF,  p-value: < 2.2e-16
```

This *regression model with linear trend* contains one independent variables: trend index (t). The regression equation is:

$$y_t = 83519.9 + 907.6\ t$$

The model's summary shows a high R-squared of 0.8814 and adj. R-squared of 0.8797, and all regression coefficients being statistically significant (p-value <0.01 or 0.05). Thus, it may be also good for time series forecasting. The appropriate forecast is shown below:

```
        Point Forecast     Lo 0     Hi 0
2023 Q1       149777.2 149777.2 149777.2
2023 Q2       150684.8 150684.8 150684.8
2023 Q3       151592.4 151592.4 151592.4
2023 Q4       152500.1 152500.1 152500.1
2024 Q1       153407.7 153407.7 153407.7
2024 Q2       154315.3 154315.3 154315.3
2024 Q3       155223.0 155223.0 155223.0
2024 Q4       156130.6 156130.6 156130.6
```

3. Regression Model with Quadratic Trend and Seasonality

For forecasting Walmart's revenue in 2023-2024, we apply for the entire data set the *regression model with quadratic trend and seasonality*, which was identified to be the third best model in question 2c. The *regression model with quadratic trend and seasonality* for the entire data set is presented below:

```
Call:
tslm(formula = revenue.ts ~ trend + I(trend^2) + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-6808.2 -3879.7   702.9  2273.6 10413.8 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 76481.135   1746.317  43.796  < 2e-16 ***
trend        1082.210     97.259  11.127  < 2e-16 ***
I(trend^2)     -2.553      1.291  -1.977  0.05219 .  
season2      4365.090   1410.466   3.095  0.00289 ** 
season3      1949.564   1411.080   1.382  0.17175    
season4     14366.087   1412.100  10.174 3.77e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4231 on 66 degrees of freedom
Multiple R-squared:  0.9594, Adjusted R-squared:  0.9563 
```

```
F-statistic: 311.6 on 5 and 66 DF, p-value: < 2.2e-16
```

This *regression model with quadratic trend and seasonality* contains 5 independent variables: trend index (t), squared trend index (t^2), and 3 seasonal dummy variables for Q2 (season2 – D_2), Q3 (season3 – D_3) and Q4 (season4 – D_4). The regression equation is:

$$y_t = 76481.14 + 1082.21\ t - 2.55\ t^2 + 4365.09\ D_2 + 1949.56\ D_3 + 14366.09\ D_4$$

The model's summary shows a very high R-squared of 0.9594 and adj. R-squared of 0.9563, and all regression coefficients being statistically significant (p-value < 0.01 or 0.05) with exception for the marginally insignificant coefficient (p-value=0.0522) for t^2, and for regression coefficient for D_3. Overall, this regression model is a very good fit for the historical data set, and thus can be used for forecasting Walmart's revenue in 2023-2024. The appropriate forecast is shown below:

```
        Point Forecast     Lo 0     Hi 0
2023 Q1       141878.9 141878.9 141878.9
2023 Q2       146951.0 146951.0 146951.0
2023 Q3       145237.3 145237.3 145237.3
2023 Q4       158350.5 158350.5 158350.5
2024 Q1       144676.1 144676.1 144676.1
2024 Q2       149727.7 149727.7 149727.7
2024 Q3       147993.6 147993.6 147993.6
2024 Q4       161086.5 161086.5 161086.5
```

3b. Apply the *accuracy()* function to compare the performance measures of the regression models developed in 3a with those for naïve and seasonal naïve forecasts. Present the accuracy measures in your report, compare them, and identify, using MAPE and RMSE, which forecast is most accurate to forecast Walmart's quarterly revenue in Q1-Q4 of 2023-2024.

The accuracy measures for the three forecasts specified in question 3b are listed below along with the accuracy for the respective naïve and seasonal naïve forecast:

```
> Regression Model with Linear Trend and Seasonality
         ME    RMSE      MAE    MPE  MAPE  ACF1 Theil's U
Test set  0 4168.91 3526.714 -0.183 3.168 0.856     0.453

> Regression Model with Linear Trend
         ME     RMSE      MAE    MPE  MAPE  ACF1 Theil's U
Test set  0 6920.164 5659.574 -0.416 4.957 -0.01     0.721

> Regression Model with Quadratic Trend and Seasonality
         ME    RMSE      MAE    MPE  MAPE  ACF1 Theil's U
Test set  0 4050.66 3358.498 -0.144 2.935 0.846     0.417

> Naïve Model
               ME     RMSE      MAE   MPE  MAPE   ACF1 Theil's U
Test set 1300.958 9700.409 8140.732 0.813 7.018 -0.707         1

> Seasonal Naïve Model
               ME     RMSE      MAE   MPE  MAPE ACF1 Theil's U
Test set 4399.824 5599.183 4570.088 3.834 3.985  0.7     0.583
```

Based on the MAPE of 2.94% and RMSE of 4050.66, the *regression model with quadratic trend and seasonality* is most accurate model. Therefore, the *regression model with quadratic trend and seasonality* should be used in forecasting Walmart's revenue in Q1-Q4 of 2023-2024.