BAN 5753, Time Series (ARIMA, ARIMAX) Exercise (10 Points)

You must do it alone (it is not a group activity)

Use the weekly sales data (Saledata.sas7bdat) posted on canvas for this exercise. It has 84 weeks of sales data from a company. The variables are as follows:

The variables in the two data sets are shown below:

**Date** Week ending date

**PrintMedia** adv expense in print in thousands of $

**TVRadio** adv expense in TV and Radio in thousands of $

**Internet** adv expense in Internet/Web/Social Media in thousands of $

**SalesAmount** Weekly sales in thousands of $

1. Explore the target variable, SalesAmount via appropriate plots. What do you see in the plot? Does it look stationary? Can you test if this a stationary series? What AR or MA process do you think describe its ACF, PACF, and IACF? (1 point)
2. Use training (first 78 observation) and validation (last 6 observation) then run following models on the target variable. Write a concise report by comparing and contrasting each of the following models (create an appropriate summary table) on following criteria : Number of parameters in the model, Whether these parameters are statistically significant, whether the error autocorrelations after fitting the model is significant (up to 24 lags), is the error normal, Fit statistics (AIC, BIC on training data) and Accuracy statistics (MAPE and RMSE on validation data). Based on this analysis, which is your best model? Why? (3 points)
   1. AR(1) model
   2. MA(1) model
   3. AR(3) model
   4. MA(3) model
   5. ARMA(1,1) model
   6. ARMA(3,3) Model
3. Use PrintMedia, TvRadio and Internet as X variables to predict the target variable. Use training (first 78 observation) and validation (last 6 observation) then run following models. Write a concise report by comparing and contrasting each of the following models (create an appropriate summary table) on following criteria : Number of parameters in the model, Whether these parameters are statistically significant, whether the error autocorrelations after fitting the model is significant (up to 24 lags), is the error normal, Fit statistics (AIC, BIC on training data) and Accuracy statistics (MAPE and RMSE on validation data). Based on this analysis, which is your best model? Why? (3 points)
   1. ARMAX with p=1
   2. ARMAX with p=1 and q=1
4. Use your choice of open source (R or Python) to compare and contrast the best ARIMA model from #2 and ARMAX model from #3 and write a short report of this comparison. (3 points)

Deliverables:

As you complete the exercise, create a short report in Microsoft Word and in this report answer the questions in the exercise description. Copy and paste supporting documents/diagrams/screenshots as needed to justify your answer. Make sure you print your name, section number, student ID# on the report and turn-in the report as communicated by your instructor. *For R/Python part of this assignment, also turn-in your codes in an appendix of your report.*