Homework 8

STAT 462 (Fall 2020)

UNAUTHORIZED DISTRIBUTION AND/OR UPLOADING OF THIS DOCUMENT IS STRICTLY PROHIBITED.

Clearly label your answers to each question and each sub-question. Your answers MUST be uploaded to Canvas as a <HWx_Yourfirtname.nb.html> file by the deadline.

The dataset *HollywoodMovies2011.csv* includes information on movies that came out of Hollywood in 2011. We want to build a model to preict *profitability*, which is the percent of the budget recovered in profits. The dataset contains five **explanatory** variables and the details are as follows.

Variable Name	Description
RottenTomatoes	Meata rating of critical reviews, from the Rotten Tomatoes website
AudienceScore	Average audience score, from the Rotten Tomatoes website
TheatersOpenWeek	Number of theaters showing the moview on opening weekend
BOAverageOpenWeek	Average box office revenue per heater opening week-end, in dollars
DomesticGross	Gross revenus in the US by the end of 2011, in millions of dollars

---

(1). Read the data into R and fit the full first order regression model. Write down the estimated regression line. This model will be referred to as model 1 hereafter. Provide any R outputs you might have used.

(2). By looking at the test results for the partial slopes (at 10% level), identify any variables you would like to drop from model 1. Provide reasons for your choice(s). You do not have to write down any steps for hypothesis testing here.

(3). Re-ft the model by eliminating the variables you decided to drop in part (2). Write down the estimated regression line. This model will be referred to as model 2 hereafter. Provide any R outputs you might have used.

(4). Use adjusted R^2 and forward model selection approach to fit the **best** first order model for this data. Summarize your R outputs in a table as below. your table should clearly illustrates the step-by-step approach for forward model selection. Once you identify the **best** model using adjusted R^2, fit the model and write down the estimated regression line.

Variable(s)	R^2	Adj. R^2
RottenTomatoes		
AudienceScore		
TheatersOpenWeek		
BOAverageOpenWeek		
DomesticGross		
DomesticGross, RottenTomatoes		
DomesticGross, AudienceScore		

Variable(s)	R^2	Adj. R^2
DomesticGross, TheatersOpenWeek		
DomesticGross, BOAverageOpenWeek		
DomesticGross, RottenTomatoes, AudienceScore		
DomesticGross, RottenTomatoes, TheatersOpenWeek		
DomesticGross, RottenTomatoes, BOAverageOpenWeek		
DomesticGross, RottenTomatoes, AudienceScore, TheatersOpenWeek		
DomesticGross, RottenTomatoes, AudienceScore, BOAverageOpenWeek		