**Module 1 Assignment**

**QMB-6304 Analytical Methods for Business**

Write a simple R script to execute the following:

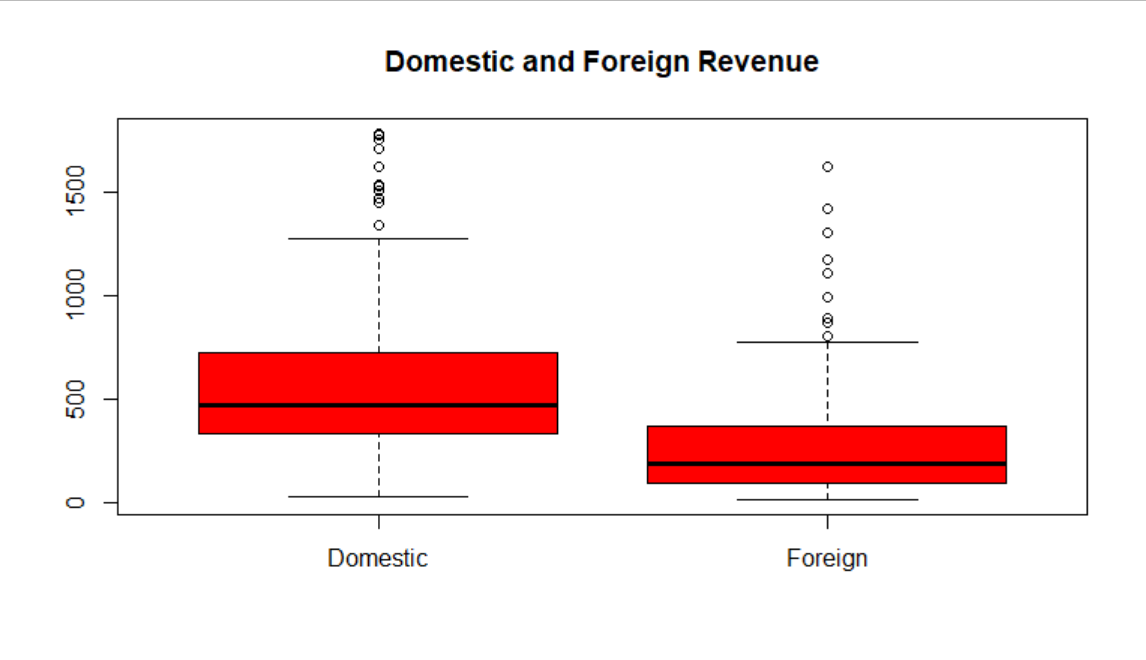
**Preprocessing**

1. Load the file “6304 Module 1 Assignment Data.xlsx” into R. This file contains information on 155 films made by [RKO Studios](https://en.wikipedia.org/wiki/RKO_Pictures) in Hollywood in the 1930s and 1940s. This will be your master data set.
2. Using the numerical portion of your U number as a random number seed, take a random sample of 45 films from the full data set using the method presented in class. This will be your primary data set
3. Create a new variable your primary data which will hold the total revenue generated by each film. This will of course be the addition of the domestic and foreign revenues variables.
4. Create a new variable in your primary data which will hold the total profit generated by each film. This will of course be calculated as total revenue minus production costs minus distribution costs.

**Analysis**

Using your primary data object in R calculate and report the following:

1. The structure of the data object using the str() command.
2. Mean, Median, Standard Deviation, Skewness, and Kurtosis of the profit variable. Based on these descriptive measurements do you think the distribution of this data is symmetric?
3. A boxplot of the profit variable. Color your boxplot green and give it an appropriate main title.
4. Quantiles of the domestic revenue variable. Show your quartiles running from the minimum to maximum values for the variable, incrementing by .10.
5. A simple histogram of the profit variable. Color your histogram blue. From this would you say a majority of RKO films were profitable or money losers?
6. The boxplot shown on the next page was created using the full (N=155) data set. Write the R code which will replicate this graphic with the full data. Show your code and the result of the code.



Your deliverable will be a single MS-Word file showing 1) the R script which executes the above instructions, 2) the results of the code, and 3) any interpretations required. The first line of your script file should be a “#” comment line showing your name as it appears in Canvas. Results should be presented in the order in which they are listed here. Deliverable due time will be announced in class and on Canvas. **This is an individual assignment to be completed and submitted by the time stated on Canvas. No collaboration of any sort is allowed on this assignment.**