James Compagno – Due at 6PM

Quiz 2

Please include the relevant R code and output with your answers. Convert the file to a PDF and name it using your first and last name. Upload the PDF to the clearly marked subfolder under Assignments in Canvas by the specified deadline. The quiz is worth 10 points. Good luck!
The internet is NOT allowed for any purpose other than to download and upload the quiz.

Answers in blue

1. **(2 points)** A researcher wants to analyze the impact of government subsidies and firm size on production output for manufacturing firms in the automotive industry. She is aware that increasing government subsidies generally boosts production output, but the effect may diminish at higher levels due to inefficiencies. Using data on government subsidies (in millions of dollars), firm size (in number of employees), and production output (in units) for 60 firms over a period of one year, the following model was estimated:

$$\widehat{\text{Output}} = 118620 + 7740\text{Subsidy} - 0.42\text{Subsidy}^2 + 180\text{Size}.$$

1) What is the change in output resulting from $1 million increase in the subsidy, holding size constant?

```
> 7740-(2*0.42)
[1] 7739.16
```

Holding size constant the partial impact (change in output) is predicted to be 7739.16

2) Calculate the level of subsidy that maximizes production output.

   i) The level of subsidy that maximizes production output is given by $\frac{-7740}{2(-0.42)} = 9{,}214.29.$

2. **(8 points)** Explaining movie revenue involves several key factors. A higher budget often leads to better production quality and more extensive marketing, significantly impacting revenue. Positive ratings boost viewership and long-term success, while negative ratings deter audiences. Additionally, certain movie genres generally perform better at the box office. Refer to *Data Quiz 2* for a sample of 400 movies.

James Compagno – Due at 6PM

a. Estimate a regression model for domestic revenue (domrev) using budget, Rotten Tomatoes scores (tomatometer), and whether the movie is of the action/horror genre (acthorr). Both domestic revenue and budget are expressed in millions of U.S. dollars. Due to their skewed distributions, transform these two variables into natural logs before estimation. Use the estimated model to predict the domestic revenue (in $ millions) of an action/horror movie with a budget of $60 million and a Rotten Tomatoes score of 70. Report the regression output and the prediction.

Columns: tomatometer domrev acthorr budget

ln(domrev) = Btomatometer + B acthorr + ln(budget)

```
#2A
#ln(domrev) = Btomatometer + B acthorr + ln(budget)
Model1 <- lm(log(domrev) ~ tomatometer + acthorr + log(budget), data = myData)
summary(Model1)
```

```
Call:
lm(formula = log(domrev) ~ tomatometer + acthorr + log(budget),
data = myData)

Residuals:
Min 1Q Median 3Q Max
-6.5502 -0.3989 0.0735 0.5277 2.0101

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.740217 0.215493 3.435 0.000655 ***
tomatometer 0.009430 0.001901 4.961 0.00000104 ***
acthorr 0.345925 0.104655 3.305 0.001035 **
log(budget) 0.667157 0.048769 13.680 < 0.0000000000000002 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9757 on 396 degrees of freedom
Multiple R-squared: 0.4079, Adjusted R-squared: 0.4034
F-statistic: 90.92 on 3 and 396 DF, p-value: < 0.00000000000000022
```

```
preexpo <- predict(Model1, data.frame(acthorr = 1, budget=60, tomatometer = 70))
```

```
exp(preexpo)
> predict(Model1, data.frame(acthorr = 1, budget=60, tomatometer = 70))
1
4.477796
> preexpo <- predict(Model1, data.frame(acthorr = 1, budget=60, tomatometer = 70))
> exp(preexpo)
1
88.04045
```

b. You believe that the impact of ln(budget) on ln(revenue) depends on whether the movie belongs to the action/horror genre. Extend the model in part (a) to reflect this belief, and then estimate the extended model to predict the domestic revenue (in $ millions) of an action/horror movie with a budget of $60 million and a Rotten Tomatoes score of 70. Report the regression output of the extended model and the prediction.

ln(domrev) = Btomatometer + Bacthorr + Bln(budget) + B(acthorr * ln(budget))

```
Model2 <- lm(log(domrev) ~ tomatometer + acthorr + log(budget) + acthorr * log(budget), data = myData)
summary(Model2)
preexpo2 <- predict(Model2, data.frame(acthorr = 1, budget=60, tomatometer = 70))
exp(preexpo2)
```

```
Call:
lm(formula = log(domrev) ~ tomatometer + acthorr + log(budget) +
acthorr * log(budget), data = myData)

Residuals:
Min 1Q Median 3Q Max
-6.5260 -0.4005 0.0852 0.5117 2.2658

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.28768 0.31904 -0.902 0.368
tomatometer 0.01148 0.00192 5.977 0.00000000508 ***
acthorr 1.89254 0.37448 5.054 0.00000066359 ***
log(budget) 0.94418 0.08025 11.765 < 0.0000000000000002 ***
```

```
acthorr:log(budget) -0.43643 0.10164 -4.294 0.00002213005 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9549 on 395 degrees of freedom
Multiple R-squared: 0.4343, Adjusted R-squared: 0.4285
F-statistic: 75.8 on 4 and 395 DF, p-value: < 0.00000000000000022
> preexpo2 <- predict(Model2, data.frame(acthorr = 1, budget=60, tomatometer = 70))
> exp(preexpo2)
1
88.87907
```

c. Conduct the appropriate test at the 5% significance level to assess whether the variables action/horror and its interaction with ln(budget) jointly contribute to explaining ln(domestic revenue). Clearly state the hypotheses, calculate and report the appropriate test statistic and *p*-value, and your statistical inference.

H0: Bacthorr = B(acthorr * ln(budget)) = 0

HA: Bacthorr * Bln(budget) != 0

```
anova(Model1)
anova(Model2)
Ftest <- ((377.02-360.21)/1)/(360.21/395); Ftest
pf(Ftest,2,395,lower.tail=FALSE)
anova(Model1, Model2)

> Ftest <- ((377.02-360.21)/1)/(360.21/395); Ftest
[1] 18.43355
> pf(Ftest,2,395,lower.tail=FALSE)
[1] 0.00000002219609
> anova(Model1, Model2)
Analysis of Variance Table

Model 1: log(domrev) ~ tomatometer + acthorr + log(budget)
Model 2: log(domrev) ~ tomatometer + acthorr + log(budget) + acthorr *
```

James Compagno – Due at 6PM

log(budget)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 396 377.02

2 395 360.21 1 16.813 18.437 0.00002213 ***

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

At a 5% significance level, action/horror and its interaction with ln(budget) jointly contribute to ln(domestic revenue).