SSC 442

Group 2

In-class Assignment 1/23

bank <- read.csv("http://msudataanalytics.github.io/SSC442/Labs/data/bank.csv")

null\_bank\_model = lm(balance ~ 1, data = bank)

full\_bank\_model = lm(balance ~ job+age+marital+education+default+housing+loan+contact+day+month+duration+campaign+previous, data = bank)

anova(null\_bank\_model, full\_bank\_model)

is.factor(bank$job)

is.factor(bank$age)

is.factor(bank$marital)

is.factor(bank$education)

is.factor(bank$default)

is.factor(bank$housing)

is.factor(bank$loan)

is.factor(bank$contact)

is.factor(bank$day)

is.factor(bank$month)

is.factor(bank$duration)

is.factor(bank$campaign)

is.factor(bank$previous)

null\_bank\_model2 = lm(balance ~ 1, data = bank)

full\_bank\_model2 = lm(balance ~ job+marital+education+default+housing+loan+contact+month, data = bank)

anova(null\_bank\_model, full\_bank\_model)

When we first ran the regression, we found that not all of the variables had an impact on the balance. After running the F-test, we found that age, day, duration, campaign, and previous did not have an impact on the balance.