Final Exam INFO 3130

Cody Frisby

12/11/2017

R Code

```
## Read the data from the provided excel file
path <- "~/Documents/school/info3130/data/finalexamdata.xlsx"
sheets <- excel_sheets(path)
df1 <- read_excel(path, sheets[1])
## df1 has an extra row that shouldn't be there
df1 <- df1[-380, ]
df2 <- read_excel(path, sheets[2])
df3 <- read_excel(path, sheets[3])
df4 <- read_excel(path, sheets[4])
df5 <- read_excel(path, sheets[5])
df6 <- read_excel(path, sheets[6])
df7 <- read_excel(path, sheets[7])
#########################################################
# Here's a function I wrote to do some repetitive work.
## Split the data into a test and training set
CrossVal <- function(df, p = 2/3) { # takes an R dataframe
  n <- dim(df)[1] # number of rows in n
  s <- sample(1:n, size = n * p) # simple random sample of size n*p
  train <- df[s, ] # training dataset
  test <- df[-s, ] # cross validation training set.
  return(list(train = train, test = test))
}
```