Home Work 1

Antonio J Bayquen

February 10, 2016

```
#include the RCurl package to enable download of the database from the internet website
library(RCurl)
```

```
## Warning: package 'RCurl' was built under R version 3.2.3
```

```
## Loading required package: bitops
```

```
URL <- "https://archive.ics.uci.edu/ml/machine-learning-databases/bridges/bridges.data.version2"
x <- getURL(URL)
#output csv file to out2 dataframe without a header row
out2 <- read.csv(textConnection(x), header=FALSE)
View(out2)
#write dataset to a a csv file
write.csv(out2, file='pitt_briges.csv', row.names=FALSE)

#create subset of dataset - exclude some columns
newdata1 <- subset(out2,select=c(-V3,-V4,-V8,-V9,-V12))
View(newdata1)

#give a descriptive name to selected columns
library(plyr)
newdata2 <- rename(newdata1, c("V1"="Identifier", "V2"="River", "V5"="Purpose", "V6"="Length", "V7"="La
View(newdata2)

#spell out all the rivers that passes through Pittsburgh, PA
factor(newdata2$River)
```

```
##   [1] M A A A M A A M A A A M A A A M A A A M M A O M A M A A M A M O M O A
##  [36] M M A A M M M O M A M M A A M A M A M M M A M Y M A O O A M O A A A A
##  [71] A A A A O A M M A M A M O O M M A A A A O Y Y M M M M A O O M A M A O
## [106] M O A
## Levels: A M O Y
```

```
table(newdata2$River)
```

```
##
##  A  M  O  Y
## 49 41 15  3
```

```
newdata2$River <- as.character(newdata2$River)
newdata2$River[newdata2$River == "A"] <- "Allegheny"
newdata2$River[newdata2$River == "M"] <- "Monongahela"
newdata2$River[newdata2$River == "O"] <- "Ohio"
newdata2$River[newdata2$River == "Y"] <- "Youghiogheny"
```

```r
View(newdata2)

#Get the mean (excluding missing data) of the # of Lanes for all the bridges
newdata3 <- subset(newdata2, Lanes != "?", Lanes)
mean(as.numeric(newdata3$Lanes))
```

```
## [1] 3.293478
```

```r
#replace ARCH to more SPECIFIC ARCH-T for bridge architecture of identified bridge
rowid <- (newdata2$Identifier == "E28")
tmp <- as.character(newdata2$Type)
tmp[rowid == TRUE] <- "ARCH-T"
newdata2$Type <- factor(tmp)

#replace ARCH to more SPECIFIC TIED-A for bridge architecture of identified bridge
rowid <- (newdata2$Identifier %in% c("E91","E90","E84","E83","E73") )
tmp <- as.character(newdata2$Type)
tmp[rowid == TRUE] <- "TIED-A"
newdata2$Type <- factor(tmp)

#replace ARCH to more SPECIFIC NOT-TIED for bridge architecture of identified bridge
rowid <- (newdata2$Identifier %in% c("E97","E78","E77","E75","E66","E64","E43") )
tmp <- as.character(newdata2$Type)
tmp[rowid == TRUE] <- "NOT-TIED"
newdata2$Type <- factor(tmp)

#write transformed Pittsburgh-Bridges dataset to a a csv file
write.csv(newdata2, file='PittBrid.csv')
```