DATA 697 - Week 3 assignment

Adam Gersowitz

2/8/2020

Overview

Week 3 assignment will be working with various data sets to practice Data manipulation and processing

Question 1

Using the 173 majors listed in fivethirtyeight.com's College Majors dataset [https://fivethirtyeight.com/features/the-economic-guide-to-picking-a-college-major/], provide code that identifies the majors that contain either "DATA" or "STATISTICS"

```
library (readr)
library(RCurl)


x <- getURL("https://raw.githubusercontent.com/fivethirtyeight/data/master/college-majors/majors-list.cs
majors <- read.csv(text=x)
#print(fightsongs)
#head(majors)

majorssdf <- data.frame(majors)
data_stats_majors <- subset(majorssdf, grepl("DATA", Major) | grepl("STATISTICS", Major))

data_stats_majors
```

```
##    FOD1P                                     Major          Major_Category
## 44  6212 MANAGEMENT INFORMATION SYSTEMS AND STATISTICS                Business
## 52  2101      COMPUTER PROGRAMMING AND DATA PROCESSING Computers & Mathematics
## 59  3702               STATISTICS AND DECISION SCIENCE Computers & Mathematics
```

Question 2

Write code that transforms the data below:

[1] "bell pepper" "bilberry" "blackberry" "blood orange"

[5] "blueberry" "cantaloupe" "chili pepper" "cloudberry"

[9] "elderberry" "lime" "lychee" "mulberry"

[13] "olive" "salal berry"

Into a format like this:

c(“bell pepper”, “bilberry”, “blackberry”, “blood orange”, “blueberry”, “cantaloupe”, “chili pepper”, “cloudberry”, “elderberry”, “lime”, “lychee”, “mulberry”, “olive”, “salal berry”)

```
list1 = c("bell pepper", "bilberry", "blackberry","blood orange")
list2 = c("blueberry","cantalope","chili pepper","cloudberry")
list3 = c("elderberry","lime","lychee","mulberry")
list4 = c("olive","salal berry")

final<-c(list1,list2,list3,list4)
final
```

```
##  [1] "bell pepper"  "bilberry"     "blackberry"   "blood orange" "blueberry"
##  [6] "cantalope"    "chili pepper" "cloudberry"   "elderberry"   "lime"
## [11] "lychee"       "mulberry"     "olive"        "salal berry"
```

Question 3

Describe, in words, what these expressions will match: (.)\1\1

```
#(.)\1\1
#"(.)(.)\\2\\1"
#(..)\1
#"(.).\\1.\\1"
#"(.)(.)(.).*\\3\\2\\1"
```