Homework 2

Zabih Yousuf

## Problem 1

### Part A

Netflix <- read.table("Homework 2 Data 1.txt", sep = "\t", fill = TRUE) #

### Part B

Netflix

## V1 V2 V3 V4 V5  
## 1 2012 Q1 23410 3065 869791  
## 2 2012 Q2 23938 NA 889163  
## 3 2012 Q3 25101 4311 905089  
## 4 2012 Q4 27146 6121 945239  
## 5 2013 Q1 NA 7142 1023961  
## 6 2013 Q2 29807 7747 1069372  
## 7 2013 Q3 31092 9188 1105999  
## 8 2013 Q4 33420 NA 1175230  
## 9 2014 Q1 35674 12683 1270089  
## 10 2014 Q2 36244 13801 1340407  
## 11 2014 Q3 NA 15843 1409432  
## 12 2014 Q4 39114 18277 1484728  
## 13 2015 Q1 41397 20877 1573129  
## 14 2015 Q2 NA 23251 NA  
## 15 2015 Q3 43181 25987 1738355  
## 16 2015 Q4 44738 30024 1823333  
## 17 2016 Q1 46967 34533 1957736  
## 18 2016 Q2 NA NA 2105204  
## 19 2016 Q3 47497 39246 2290188  
## 20 2016 Q4 49431 44365 2477541  
## 21 2017 Q1 50854 47894 2636635  
## 22 2017 Q2 51921 52031 2785464  
## 23 2017 Q3 NA 56476 2984859  
## 24 2017 Q4 NA NA NA

### Part c

names(Netflix) <- c("Year","Quarter","US Streaming","Int.Stream", "All.Rev")

### Part D

usStreaming <- is.na(Netflix[3])

### Part E

Netflix[usStreaming,][2:3]

## Quarter US Streaming  
## 5 Q1 NA  
## 11 Q3 NA  
## 14 Q2 NA  
## 18 Q2 NA  
## 23 Q3 NA  
## 24 Q4 NA

### Part F

table(Netflix[usStreaming,][1:2])

## Quarter  
## Year Q1 Q2 Q3 Q4   
## 2013 1 0 0 0  
## 2014 0 0 1 0  
## 2015 0 1 0 0  
## 2016 0 1 0 0  
## 2017 0 0 1 1

## Problem 2

### Part A

Netflix <- read.csv("Homework 2 Data 2.csv")  
Netflix

## Year Quarter Revenue Cumulative.Revenue Marketing Cumulative.Marketing  
## 1 2012 Q1 869791 869791 122288 122288  
## 2 2012 Q2 889163 1758954 108441 230729  
## 3 2012 Q3 905089 2664043 102913 333642  
## 4 2012 Q4 945239 3609282 105566 439208  
## 5 2013 Q1 1023961 1023961 119086 119086  
## 6 2013 Q2 1069372 2093333 114611 233697  
## 7 2013 Q3 1105999 3199332 108228 341925  
## 8 2013 Q4 1175230 4374562 128017 469942  
## 9 2014 Q1 1270089 1270089 137098 137098  
## 10 2014 Q2 1340407 2610496 120763 257861  
## 11 2014 Q3 1409432 4019928 145654 403515  
## 12 2014 Q4 1484728 5504656 203671 607186  
## 13 2015 Q1 1573129 1573129 194677 194677  
## 14 2015 Q2 1644694 3217823 197140 391817  
## 15 2015 Q3 1738355 4956178 208102 599919  
## 16 2015 Q4 1823333 6779511 224173 824092  
## 17 2016 Q1 1957736 1957736 208010 208010  
## 18 2016 Q2 2105204 4062940 216029 424039  
## 19 2016 Q3 2290188 6353128 282043 706082  
## 20 2016 Q4 2477541 8830669 284996 991078  
## 21 2017 Q1 2636635 2636635 271270 271270  
## 22 2017 Q2 2785464 5422099 274323 545593  
## 23 2017 Q3 2984859 8406958 312490 858083  
## 24 2017 Q4 3285755 11692713 419939 1278022

### Part B

remCumu <- Netflix[c(-4,-6)]

### Part C

remCumu[order(remCumu),][3:4]

## Revenue Marketing  
## NA NA NA  
## NA.1 NA NA  
## NA.2 NA NA  
## NA.3 NA NA  
## NA.4 NA NA  
## NA.5 NA NA  
## NA.6 NA NA  
## NA.7 NA NA  
## NA.8 NA NA  
## NA.9 NA NA  
## NA.10 NA NA  
## NA.11 NA NA  
## NA.12 NA NA  
## NA.13 NA NA  
## NA.14 NA NA  
## NA.15 NA NA  
## NA.16 NA NA  
## NA.17 NA NA  
## NA.18 NA NA  
## NA.19 NA NA  
## NA.20 NA NA  
## NA.21 NA NA  
## NA.22 NA NA  
## NA.23 NA NA  
## NA.24 NA NA  
## NA.25 NA NA  
## NA.26 NA NA  
## NA.27 NA NA  
## NA.28 NA NA  
## NA.29 NA NA  
## 1 869791 122288  
## 2 889163 108441  
## 3 905089 102913  
## 4 945239 105566  
## 5 1023961 119086  
## 6 1069372 114611  
## 7 1105999 108228  
## 8 1175230 128017  
## 9 1270089 137098  
## 10 1340407 120763  
## 11 1409432 145654  
## 12 1484728 203671  
## 13 1573129 194677  
## 14 1644694 197140  
## 15 1738355 208102  
## 16 1823333 224173  
## 17 1957736 208010  
## 18 2105204 216029  
## 19 2290188 282043  
## 20 2477541 284996  
## 21 2636635 271270  
## 22 2785464 274323  
## 23 2984859 312490  
## 24 3285755 419939  
## NA.30 NA NA  
## NA.31 NA NA  
## NA.32 NA NA  
## NA.33 NA NA  
## NA.34 NA NA  
## NA.35 NA NA  
## NA.36 NA NA  
## NA.37 NA NA  
## NA.38 NA NA  
## NA.39 NA NA  
## NA.40 NA NA  
## NA.41 NA NA  
## NA.42 NA NA  
## NA.43 NA NA  
## NA.44 NA NA  
## NA.45 NA NA  
## NA.46 NA NA  
## NA.47 NA NA  
## NA.48 NA NA  
## NA.49 NA NA  
## NA.50 NA NA  
## NA.51 NA NA  
## NA.52 NA NA  
## NA.53 NA NA  
## NA.54 NA NA  
## NA.55 NA NA  
## NA.56 NA NA  
## NA.57 NA NA  
## NA.58 NA NA  
## NA.59 NA NA  
## NA.60 NA NA  
## NA.61 NA NA  
## NA.62 NA NA  
## NA.63 NA NA  
## NA.64 NA NA  
## NA.65 NA NA  
## NA.66 NA NA  
## NA.67 NA NA  
## NA.68 NA NA  
## NA.69 NA NA  
## NA.70 NA NA  
## NA.71 NA NA

Yes there is a positive relationship between the two variables. There is an instance where the marketing values decrease but the revenue goes up however the majority of the values under the revenue variable increase when we sort the values from least to greatest.

### Part D

lastQuarter <- subset(Netflix, Quarter == 'Q4', select = c(1,4,6))

### Part E

write.table(lastQuarter, "Homework 2 Out 2.txt", row.names = FALSE, quote = FALSE)

### Part F

lastQuarter[order(lastQuarter$Cumulative.Marketing),]

## Year Cumulative.Revenue Cumulative.Marketing  
## 4 2012 3609282 439208  
## 8 2013 4374562 469942  
## 12 2014 5504656 607186  
## 16 2015 6779511 824092  
## 20 2016 8830669 991078  
## 24 2017 11692713 1278022

Here there is a clear positive relationship between the two variables. As marketing values increase so does revenue.

### Part G

We can conclude that the more money spent on marketing, the revenue for the year will also increase.