Homework 3

Zabih Yousuf

## Problem 1

data1 <- read.csv("Homework 3 Data 1.csv", header = TRUE)  
data2 <- read.csv("Homework 3 Data 2.csv", header = TRUE)  
data3 <- read.csv("Homework 3 Data 3.csv", header = TRUE)  
data4 <- read.csv("Homework 3 Data 4.csv", header = TRUE)  
data5 <- read.csv("Homework 3 Data 5.csv", header = TRUE)  
data6 <- read.csv("Homework 3 Data 6.csv", header = TRUE)  
data7 <- read.csv("Homework 3 Data 7.csv", header = TRUE)  
data8 <- read.csv("Homework 3 Data 8.csv", header = TRUE)

## Problem 2

### Part (A)

int\_rev <- rbind(data1, data3)  
  
int\_rev

## Year Quarter Revenue  
## 1 2010 Q4 3617  
## 2 2011 Q1 12279  
## 3 2011 Q2 18896  
## 4 2011 Q3 22687  
## 5 2011 Q4 28988  
## 6 2012 Q1 43425  
## 7 2012 Q2 64973  
## 8 2012 Q3 77744  
## 9 2012 Q4 101400  
## 10 2013 Q1 142019  
## 11 2013 Q2 165902  
## 12 2013 Q3 183051  
## 13 2013 Q4 221418  
## 14 2014 Q1 267118  
## 15 2014 Q2 307461  
## 16 2014 Q3 345685  
## 17 2014 Q4 387797  
## 18 2015 Q1 415397  
## 19 2015 Q2 454763  
## 20 2015 Q3 516870  
## 21 2015 Q4 566405  
## 22 2016 Q1 651748  
## 23 2016 Q2 758201  
## 24 2016 Q3 853480  
## 25 2016 Q4 947666  
## 26 2017 Q1 1046199  
## 27 2017 Q2 1165228  
## 28 2017 Q3 1327435  
## 29 2017 Q4 1550329

### Part (B)

names(data2) <- names(data4)  
us\_rev <- rbind(data2, data4)  
  
us\_rev

## Year Quarter Revenue  
## 1 2011 Q4 476334  
## 2 2012 Q1 506665  
## 3 2012 Q2 532705  
## 4 2012 Q3 556027  
## 5 2012 Q4 589471  
## 6 2013 Q1 638649  
## 7 2013 Q2 671089  
## 8 2013 Q3 701083  
## 9 2013 Q4 740554  
## 10 2014 Q1 798617  
## 11 2014 Q2 838225  
## 12 2014 Q3 877150  
## 13 2014 Q4 917442  
## 14 2015 Q1 984532  
## 15 2015 Q2 1025913  
## 16 2015 Q3 1063961  
## 17 2015 Q4 1105933  
## 18 2016 Q1 1161241  
## 19 2016 Q2 1208271  
## 20 2016 Q3 1304333  
## 21 2016 Q4 1403462  
## 22 2017 Q1 1470042  
## 23 2017 Q2 1505499  
## 24 2017 Q3 1547210  
## 25 2017 Q4 1630274

### Part (C)

names(data7) <- names(data5)  
int\_mem <- rbind(data5, data7)  
  
int\_mem

## Year Quarter Membership  
## 1 2010 Q3 133  
## 2 2010 Q4 509  
## 3 2011 Q1 803  
## 4 2011 Q2 967  
## 5 2011 Q3 1480  
## 6 2011 Q4 1858  
## 7 2012 Q1 3065  
## 8 2012 Q2 3624  
## 9 2012 Q3 4311  
## 10 2012 Q4 6121  
## 11 2013 Q1 7142  
## 12 2013 Q2 7747  
## 13 2013 Q3 9188  
## 14 2013 Q4 10930  
## 15 2014 Q1 12683  
## 16 2014 Q2 13801  
## 17 2014 Q3 15843  
## 18 2014 Q4 18277  
## 19 2015 Q1 20877  
## 20 2015 Q2 23251  
## 21 2015 Q3 25987  
## 22 2015 Q4 30024  
## 23 2016 Q1 34533  
## 24 2016 Q2 36048  
## 25 2016 Q3 39246  
## 26 2016 Q4 44365  
## 27 2017 Q1 47894  
## 28 2017 Q2 52031  
## 29 2017 Q3 56476  
## 30 2017 Q4 62832

### Part (D)

names(data8) <- names(data6)  
data8[20,2] <- 'Q4'  
  
us\_mem <- rbind(data6, data8)  
us\_mem

## Year Quarter Membership  
## 1 2011 Q3 21448  
## 2 2011 Q4 21671  
## 3 2012 Q1 23410  
## 4 2012 Q2 23938  
## 5 2012 Q3 25101  
## 6 2012 Q4 27146  
## 7 2013 Q1 29174  
## 8 2013 Q2 29807  
## 9 2013 Q3 31092  
## 10 2013 Q4 33420  
## 11 2014 Q1 35674  
## 12 2014 Q2 36244  
## 13 2014 Q3 37219  
## 14 2014 Q4 39114  
## 15 2015 Q1 41397  
## 16 2015 Q2 42300  
## 17 2015 Q3 43181  
## 18 2015 Q4 44738  
## 19 2016 Q1 46967  
## 20 2016 Q2 47129  
## 21 2016 Q3 47497  
## 22 2016 Q4 49431  
## 23 2017 Q1 50854  
## 24 2017 Q2 51921  
## 25 2017 Q3 52772  
## 26 2017 Q4 54750

## Problem 3

### Part (A)

rev\_wide <- merge(int\_rev,us\_rev,by=c("Year", "Quarter"), all = TRUE)  
names(rev\_wide) <- c("Year", "Quarter", "Int.revenue", "US.revenue")  
  
rev\_wide

## Year Quarter Int.revenue US.revenue  
## 1 2010 Q4 3617 NA  
## 2 2011 Q1 12279 NA  
## 3 2011 Q2 18896 NA  
## 4 2011 Q3 22687 NA  
## 5 2011 Q4 28988 476334  
## 6 2012 Q1 43425 506665  
## 7 2012 Q2 64973 532705  
## 8 2012 Q3 77744 556027  
## 9 2012 Q4 101400 589471  
## 10 2013 Q1 142019 638649  
## 11 2013 Q2 165902 671089  
## 12 2013 Q3 183051 701083  
## 13 2013 Q4 221418 740554  
## 14 2014 Q1 267118 798617  
## 15 2014 Q2 307461 838225  
## 16 2014 Q3 345685 877150  
## 17 2014 Q4 387797 917442  
## 18 2015 Q1 415397 984532  
## 19 2015 Q2 454763 1025913  
## 20 2015 Q3 516870 1063961  
## 21 2015 Q4 566405 1105933  
## 22 2016 Q1 651748 1161241  
## 23 2016 Q2 758201 1208271  
## 24 2016 Q3 853480 1304333  
## 25 2016 Q4 947666 1403462  
## 26 2017 Q1 1046199 1470042  
## 27 2017 Q2 1165228 1505499  
## 28 2017 Q3 1327435 1547210  
## 29 2017 Q4 1550329 1630274

### Part (B)

mem\_wide <- merge(int\_mem, us\_mem, by=c("Year", "Quarter"), all = TRUE)  
names(mem\_wide) <- c("Year", "Quarter", "Int.membership", "US.membership")  
  
mem\_wide

## Year Quarter Int.membership US.membership  
## 1 2010 Q3 133 NA  
## 2 2010 Q4 509 NA  
## 3 2011 Q1 803 NA  
## 4 2011 Q2 967 NA  
## 5 2011 Q3 1480 21448  
## 6 2011 Q4 1858 21671  
## 7 2012 Q1 3065 23410  
## 8 2012 Q2 3624 23938  
## 9 2012 Q3 4311 25101  
## 10 2012 Q4 6121 27146  
## 11 2013 Q1 7142 29174  
## 12 2013 Q2 7747 29807  
## 13 2013 Q3 9188 31092  
## 14 2013 Q4 10930 33420  
## 15 2014 Q1 12683 35674  
## 16 2014 Q2 13801 36244  
## 17 2014 Q3 15843 37219  
## 18 2014 Q4 18277 39114  
## 19 2015 Q1 20877 41397  
## 20 2015 Q2 23251 42300  
## 21 2015 Q3 25987 43181  
## 22 2015 Q4 30024 44738  
## 23 2016 Q1 34533 46967  
## 24 2016 Q2 36048 47129  
## 25 2016 Q3 39246 47497  
## 26 2016 Q4 44365 49431  
## 27 2017 Q1 47894 50854  
## 28 2017 Q2 52031 51921  
## 29 2017 Q3 56476 52772  
## 30 2017 Q4 62832 54750

### Part (C)

stream\_wide <- merge(rev\_wide,mem\_wide, all = TRUE)[,c(1,2,6,4,5,3)]  
names(stream\_wide) <- c("Year","Quarter", "US.membership", "US.revenue", "Int.membership", "Int.revenue")  
  
stream\_wide

## Year Quarter US.membership US.revenue Int.membership Int.revenue  
## 1 2010 Q3 NA NA 133 NA  
## 2 2010 Q4 NA NA 509 3617  
## 3 2011 Q1 NA NA 803 12279  
## 4 2011 Q2 NA NA 967 18896  
## 5 2011 Q3 21448 NA 1480 22687  
## 6 2011 Q4 21671 476334 1858 28988  
## 7 2012 Q1 23410 506665 3065 43425  
## 8 2012 Q2 23938 532705 3624 64973  
## 9 2012 Q3 25101 556027 4311 77744  
## 10 2012 Q4 27146 589471 6121 101400  
## 11 2013 Q1 29174 638649 7142 142019  
## 12 2013 Q2 29807 671089 7747 165902  
## 13 2013 Q3 31092 701083 9188 183051  
## 14 2013 Q4 33420 740554 10930 221418  
## 15 2014 Q1 35674 798617 12683 267118  
## 16 2014 Q2 36244 838225 13801 307461  
## 17 2014 Q3 37219 877150 15843 345685  
## 18 2014 Q4 39114 917442 18277 387797  
## 19 2015 Q1 41397 984532 20877 415397  
## 20 2015 Q2 42300 1025913 23251 454763  
## 21 2015 Q3 43181 1063961 25987 516870  
## 22 2015 Q4 44738 1105933 30024 566405  
## 23 2016 Q1 46967 1161241 34533 651748  
## 24 2016 Q2 47129 1208271 36048 758201  
## 25 2016 Q3 47497 1304333 39246 853480  
## 26 2016 Q4 49431 1403462 44365 947666  
## 27 2017 Q1 50854 1470042 47894 1046199  
## 28 2017 Q2 51921 1505499 52031 1165228  
## 29 2017 Q3 52772 1547210 56476 1327435  
## 30 2017 Q4 54750 1630274 62832 1550329

## Problem 4

### Part (A)

int\_mem$Service <- rep('Int', length(int\_mem[1]))  
int\_mem

## Year Quarter Membership Service  
## 1 2010 Q3 133 Int  
## 2 2010 Q4 509 Int  
## 3 2011 Q1 803 Int  
## 4 2011 Q2 967 Int  
## 5 2011 Q3 1480 Int  
## 6 2011 Q4 1858 Int  
## 7 2012 Q1 3065 Int  
## 8 2012 Q2 3624 Int  
## 9 2012 Q3 4311 Int  
## 10 2012 Q4 6121 Int  
## 11 2013 Q1 7142 Int  
## 12 2013 Q2 7747 Int  
## 13 2013 Q3 9188 Int  
## 14 2013 Q4 10930 Int  
## 15 2014 Q1 12683 Int  
## 16 2014 Q2 13801 Int  
## 17 2014 Q3 15843 Int  
## 18 2014 Q4 18277 Int  
## 19 2015 Q1 20877 Int  
## 20 2015 Q2 23251 Int  
## 21 2015 Q3 25987 Int  
## 22 2015 Q4 30024 Int  
## 23 2016 Q1 34533 Int  
## 24 2016 Q2 36048 Int  
## 25 2016 Q3 39246 Int  
## 26 2016 Q4 44365 Int  
## 27 2017 Q1 47894 Int  
## 28 2017 Q2 52031 Int  
## 29 2017 Q3 56476 Int  
## 30 2017 Q4 62832 Int

int\_rev$Service <- rep('Int', length(int\_rev[1]))  
int\_rev

## Year Quarter Revenue Service  
## 1 2010 Q4 3617 Int  
## 2 2011 Q1 12279 Int  
## 3 2011 Q2 18896 Int  
## 4 2011 Q3 22687 Int  
## 5 2011 Q4 28988 Int  
## 6 2012 Q1 43425 Int  
## 7 2012 Q2 64973 Int  
## 8 2012 Q3 77744 Int  
## 9 2012 Q4 101400 Int  
## 10 2013 Q1 142019 Int  
## 11 2013 Q2 165902 Int  
## 12 2013 Q3 183051 Int  
## 13 2013 Q4 221418 Int  
## 14 2014 Q1 267118 Int  
## 15 2014 Q2 307461 Int  
## 16 2014 Q3 345685 Int  
## 17 2014 Q4 387797 Int  
## 18 2015 Q1 415397 Int  
## 19 2015 Q2 454763 Int  
## 20 2015 Q3 516870 Int  
## 21 2015 Q4 566405 Int  
## 22 2016 Q1 651748 Int  
## 23 2016 Q2 758201 Int  
## 24 2016 Q3 853480 Int  
## 25 2016 Q4 947666 Int  
## 26 2017 Q1 1046199 Int  
## 27 2017 Q2 1165228 Int  
## 28 2017 Q3 1327435 Int  
## 29 2017 Q4 1550329 Int

### Part (B)

us\_rev$Service <- rep('US', length(us\_rev[1]))  
us\_rev

## Year Quarter Revenue Service  
## 1 2011 Q4 476334 US  
## 2 2012 Q1 506665 US  
## 3 2012 Q2 532705 US  
## 4 2012 Q3 556027 US  
## 5 2012 Q4 589471 US  
## 6 2013 Q1 638649 US  
## 7 2013 Q2 671089 US  
## 8 2013 Q3 701083 US  
## 9 2013 Q4 740554 US  
## 10 2014 Q1 798617 US  
## 11 2014 Q2 838225 US  
## 12 2014 Q3 877150 US  
## 13 2014 Q4 917442 US  
## 14 2015 Q1 984532 US  
## 15 2015 Q2 1025913 US  
## 16 2015 Q3 1063961 US  
## 17 2015 Q4 1105933 US  
## 18 2016 Q1 1161241 US  
## 19 2016 Q2 1208271 US  
## 20 2016 Q3 1304333 US  
## 21 2016 Q4 1403462 US  
## 22 2017 Q1 1470042 US  
## 23 2017 Q2 1505499 US  
## 24 2017 Q3 1547210 US  
## 25 2017 Q4 1630274 US

us\_mem$Service <- rep('US', length(us\_mem[1]))  
us\_mem

## Year Quarter Membership Service  
## 1 2011 Q3 21448 US  
## 2 2011 Q4 21671 US  
## 3 2012 Q1 23410 US  
## 4 2012 Q2 23938 US  
## 5 2012 Q3 25101 US  
## 6 2012 Q4 27146 US  
## 7 2013 Q1 29174 US  
## 8 2013 Q2 29807 US  
## 9 2013 Q3 31092 US  
## 10 2013 Q4 33420 US  
## 11 2014 Q1 35674 US  
## 12 2014 Q2 36244 US  
## 13 2014 Q3 37219 US  
## 14 2014 Q4 39114 US  
## 15 2015 Q1 41397 US  
## 16 2015 Q2 42300 US  
## 17 2015 Q3 43181 US  
## 18 2015 Q4 44738 US  
## 19 2016 Q1 46967 US  
## 20 2016 Q2 47129 US  
## 21 2016 Q3 47497 US  
## 22 2016 Q4 49431 US  
## 23 2017 Q1 50854 US  
## 24 2017 Q2 51921 US  
## 25 2017 Q3 52772 US  
## 26 2017 Q4 54750 US

### Part (C)

rev\_long <- rbind(int\_rev, us\_rev)[,c(1,2,4,3)]  
names(rev\_long)<- c("Year", "Quarter", "Service", "Revenue")  
  
rev\_long

## Year Quarter Service Revenue  
## 1 2010 Q4 Int 3617  
## 2 2011 Q1 Int 12279  
## 3 2011 Q2 Int 18896  
## 4 2011 Q3 Int 22687  
## 5 2011 Q4 Int 28988  
## 6 2012 Q1 Int 43425  
## 7 2012 Q2 Int 64973  
## 8 2012 Q3 Int 77744  
## 9 2012 Q4 Int 101400  
## 10 2013 Q1 Int 142019  
## 11 2013 Q2 Int 165902  
## 12 2013 Q3 Int 183051  
## 13 2013 Q4 Int 221418  
## 14 2014 Q1 Int 267118  
## 15 2014 Q2 Int 307461  
## 16 2014 Q3 Int 345685  
## 17 2014 Q4 Int 387797  
## 18 2015 Q1 Int 415397  
## 19 2015 Q2 Int 454763  
## 20 2015 Q3 Int 516870  
## 21 2015 Q4 Int 566405  
## 22 2016 Q1 Int 651748  
## 23 2016 Q2 Int 758201  
## 24 2016 Q3 Int 853480  
## 25 2016 Q4 Int 947666  
## 26 2017 Q1 Int 1046199  
## 27 2017 Q2 Int 1165228  
## 28 2017 Q3 Int 1327435  
## 29 2017 Q4 Int 1550329  
## 30 2011 Q4 US 476334  
## 31 2012 Q1 US 506665  
## 32 2012 Q2 US 532705  
## 33 2012 Q3 US 556027  
## 34 2012 Q4 US 589471  
## 35 2013 Q1 US 638649  
## 36 2013 Q2 US 671089  
## 37 2013 Q3 US 701083  
## 38 2013 Q4 US 740554  
## 39 2014 Q1 US 798617  
## 40 2014 Q2 US 838225  
## 41 2014 Q3 US 877150  
## 42 2014 Q4 US 917442  
## 43 2015 Q1 US 984532  
## 44 2015 Q2 US 1025913  
## 45 2015 Q3 US 1063961  
## 46 2015 Q4 US 1105933  
## 47 2016 Q1 US 1161241  
## 48 2016 Q2 US 1208271  
## 49 2016 Q3 US 1304333  
## 50 2016 Q4 US 1403462  
## 51 2017 Q1 US 1470042  
## 52 2017 Q2 US 1505499  
## 53 2017 Q3 US 1547210  
## 54 2017 Q4 US 1630274

### Part (D)

mem\_long <- rbind(int\_mem, us\_mem)[,c(1,2,4,3)]  
names(rev\_long) <- c("Year", "Quarter", "Service", "Membership")  
  
mem\_long

## Year Quarter Service Membership  
## 1 2010 Q3 Int 133  
## 2 2010 Q4 Int 509  
## 3 2011 Q1 Int 803  
## 4 2011 Q2 Int 967  
## 5 2011 Q3 Int 1480  
## 6 2011 Q4 Int 1858  
## 7 2012 Q1 Int 3065  
## 8 2012 Q2 Int 3624  
## 9 2012 Q3 Int 4311  
## 10 2012 Q4 Int 6121  
## 11 2013 Q1 Int 7142  
## 12 2013 Q2 Int 7747  
## 13 2013 Q3 Int 9188  
## 14 2013 Q4 Int 10930  
## 15 2014 Q1 Int 12683  
## 16 2014 Q2 Int 13801  
## 17 2014 Q3 Int 15843  
## 18 2014 Q4 Int 18277  
## 19 2015 Q1 Int 20877  
## 20 2015 Q2 Int 23251  
## 21 2015 Q3 Int 25987  
## 22 2015 Q4 Int 30024  
## 23 2016 Q1 Int 34533  
## 24 2016 Q2 Int 36048  
## 25 2016 Q3 Int 39246  
## 26 2016 Q4 Int 44365  
## 27 2017 Q1 Int 47894  
## 28 2017 Q2 Int 52031  
## 29 2017 Q3 Int 56476  
## 30 2017 Q4 Int 62832  
## 31 2011 Q3 US 21448  
## 32 2011 Q4 US 21671  
## 33 2012 Q1 US 23410  
## 34 2012 Q2 US 23938  
## 35 2012 Q3 US 25101  
## 36 2012 Q4 US 27146  
## 37 2013 Q1 US 29174  
## 38 2013 Q2 US 29807  
## 39 2013 Q3 US 31092  
## 40 2013 Q4 US 33420  
## 41 2014 Q1 US 35674  
## 42 2014 Q2 US 36244  
## 43 2014 Q3 US 37219  
## 44 2014 Q4 US 39114  
## 45 2015 Q1 US 41397  
## 46 2015 Q2 US 42300  
## 47 2015 Q3 US 43181  
## 48 2015 Q4 US 44738  
## 49 2016 Q1 US 46967  
## 50 2016 Q2 US 47129  
## 51 2016 Q3 US 47497  
## 52 2016 Q4 US 49431  
## 53 2017 Q1 US 50854  
## 54 2017 Q2 US 51921  
## 55 2017 Q3 US 52772  
## 56 2017 Q4 US 54750

### Part (E)

stream\_long <- merge(rev\_long, mem\_long, by=c("Year", "Quarter", "Service"), all = TRUE)  
names(stream\_long) <- c("Year", "Quarter", "Service", "Membership", "Revenue")  
  
stream\_long

## Year Quarter Service Membership Revenue  
## 1 2010 Q3 Int NA 133  
## 2 2010 Q4 Int 3617 509  
## 3 2011 Q1 Int 12279 803  
## 4 2011 Q2 Int 18896 967  
## 5 2011 Q3 Int 22687 1480  
## 6 2011 Q3 US NA 21448  
## 7 2011 Q4 Int 28988 1858  
## 8 2011 Q4 US 476334 21671  
## 9 2012 Q1 Int 43425 3065  
## 10 2012 Q1 US 506665 23410  
## 11 2012 Q2 Int 64973 3624  
## 12 2012 Q2 US 532705 23938  
## 13 2012 Q3 Int 77744 4311  
## 14 2012 Q3 US 556027 25101  
## 15 2012 Q4 Int 101400 6121  
## 16 2012 Q4 US 589471 27146  
## 17 2013 Q1 Int 142019 7142  
## 18 2013 Q1 US 638649 29174  
## 19 2013 Q2 Int 165902 7747  
## 20 2013 Q2 US 671089 29807  
## 21 2013 Q3 Int 183051 9188  
## 22 2013 Q3 US 701083 31092  
## 23 2013 Q4 Int 221418 10930  
## 24 2013 Q4 US 740554 33420  
## 25 2014 Q1 Int 267118 12683  
## 26 2014 Q1 US 798617 35674  
## 27 2014 Q2 Int 307461 13801  
## 28 2014 Q2 US 838225 36244  
## 29 2014 Q3 Int 345685 15843  
## 30 2014 Q3 US 877150 37219  
## 31 2014 Q4 Int 387797 18277  
## 32 2014 Q4 US 917442 39114  
## 33 2015 Q1 Int 415397 20877  
## 34 2015 Q1 US 984532 41397  
## 35 2015 Q2 Int 454763 23251  
## 36 2015 Q2 US 1025913 42300  
## 37 2015 Q3 Int 516870 25987  
## 38 2015 Q3 US 1063961 43181  
## 39 2015 Q4 Int 566405 30024  
## 40 2015 Q4 US 1105933 44738  
## 41 2016 Q1 Int 651748 34533  
## 42 2016 Q1 US 1161241 46967  
## 43 2016 Q2 Int 758201 36048  
## 44 2016 Q2 US 1208271 47129  
## 45 2016 Q3 Int 853480 39246  
## 46 2016 Q3 US 1304333 47497  
## 47 2016 Q4 Int 947666 44365  
## 48 2016 Q4 US 1403462 49431  
## 49 2017 Q1 Int 1046199 47894  
## 50 2017 Q1 US 1470042 50854  
## 51 2017 Q2 Int 1165228 52031  
## 52 2017 Q2 US 1505499 51921  
## 53 2017 Q3 Int 1327435 56476  
## 54 2017 Q3 US 1547210 52772  
## 55 2017 Q4 Int 1550329 62832  
## 56 2017 Q4 US 1630274 54750

## Problem 5

In my opinion the wide format is easier to work with and read simply because it is easier to find information and understand what the data means. The main difference between the two is the number of columns and rows. The wide format has more columns however the rows on the long format are more than the ones on the wide format.