hw2markdown

```
getwd()
```

```
## [1] "/Users/aldrin/Desktop/econhw2"
```

```
setwd("~/Desktop/econhw2")
```

R Markdown

Question 2.2.1.a

```
data <- read.csv("college_mobility_rates.csv")
mylist <- list(data$name)
print(lengths(mylist))
```

```
## [1] 2202
```

Question 2.2.1.b There are 53 variables.

```
names(data)
```

```
##  [1] "super_opeid"        "name"               "type"
##  [4] "tier"               "tier_name"          "state"
##  [7] "czname"             "par_median"         "par_rank"
## [10] "par_q1"             "par_q2"             "par_q3"
## [13] "par_q4"             "par_q5"             "par_top1pc"
## [16] "k_rank"             "k_median"           "k_q1"
## [19] "k_q2"               "k_q3"               "k_q4"
## [22] "k_q5"               "k_top1pc"           "kq1_cond_parq1"
## [25] "kq2_cond_parq1"     "kq3_cond_parq1"     "kq4_cond_parq1"
## [28] "kq5_cond_parq1"     "kq1_cond_parq2"     "kq2_cond_parq2"
## [31] "kq3_cond_parq2"     "kq4_cond_parq2"     "kq5_cond_parq2"
## [34] "kq1_cond_parq3"     "kq2_cond_parq3"     "kq3_cond_parq3"
## [37] "kq4_cond_parq3"     "kq5_cond_parq3"     "kq1_cond_parq4"
## [40] "kq2_cond_parq4"     "kq3_cond_parq4"     "kq4_cond_parq4"
## [43] "kq5_cond_parq4"     "kq1_cond_parq5"     "kq2_cond_parq5"
## [46] "kq3_cond_parq5"     "kq4_cond_parq5"     "kq5_cond_parq5"
## [49] "ktop1pc_cond_parq1" "ktop1pc_cond_parq2" "ktop1pc_cond_parq3"
## [52] "ktop1pc_cond_parq4" "ktop1pc_cond_parq5"
```

```
length(names(data))
```

```
## [1] 53
```

Question 2.2.2.a

```
sorted1 <- data[order(data$par_q1),]
#These are the colleges with lowest par_q1
print(sorted1[1:10,c("name","par_q1")])
```

```
##                               name     par_q1
## 2076 Washington And Lee University 0.01118963
## 1924 University Of Mary Washington 0.01267039
## 425      College Of William & Mary 0.01346071
## 510                Davidson College 0.01381698
## 1965      University Of Notre Dame 0.01437956
## 388                  Colby College 0.01485769
## 759         Hampden Sydney College 0.01643258
## 1034    Loyola University Maryland 0.01688441
## 1768              Stonehill College 0.01722926
## 1976        University Of Richmond 0.01736581
```

```
#These are the colleges with highest par_q1
print(sorted1[2193:2202,c("name","par_q1")])
```

```
##                                       name    par_q1
## 1162     Mississippi Valley State University 0.4545876
## 233           CUNY, Hostos Community College 0.4582601
## 860  International Career Development Center 0.4606968
## 1201              Moultrie Technical College 0.4642638
## 179                         Boricua College 0.4665152
## 1689              Southern Careers Institute 0.4710565
## 2001      University Of Texas At Brownsville 0.4734971
## 674                Franklin Career Institute 0.4997946
## 1673                     South Texas College 0.5236070
## 1864              United Talmudical Seminary 0.6097748
```

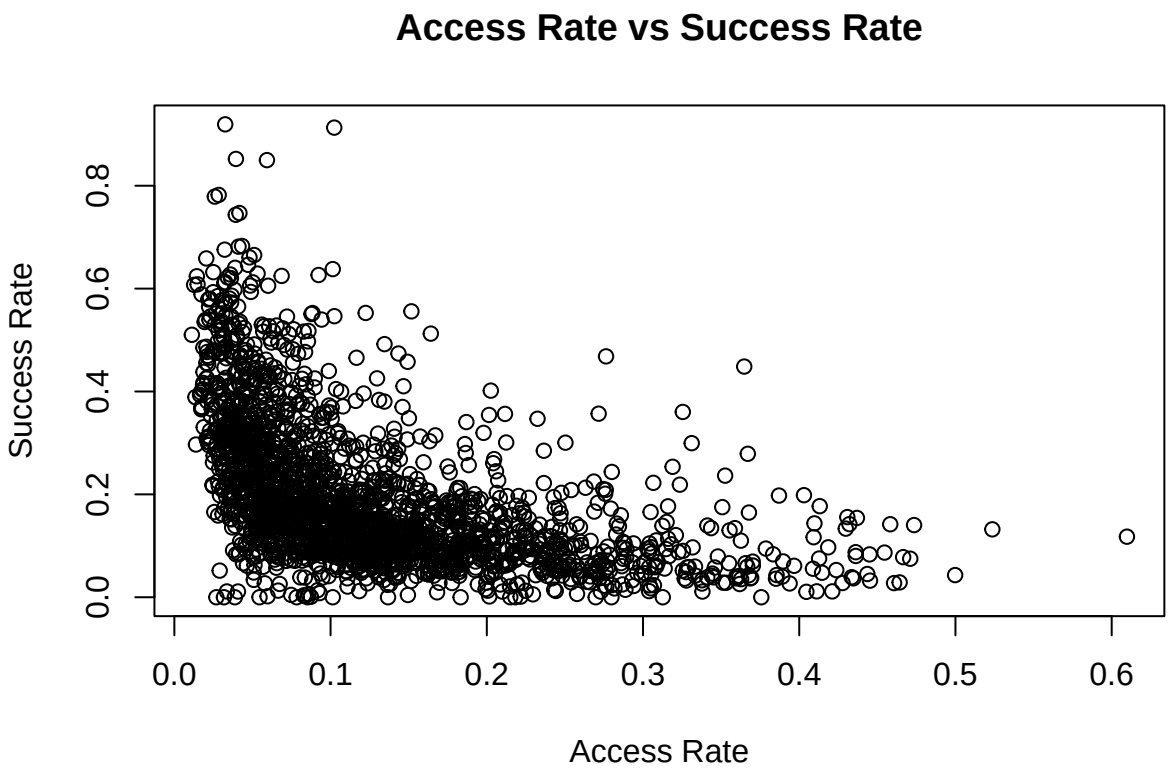
Question 2.2.2.b Yale has second lowest success rate among the four.

```
data[data$name %in% c("Cornell University", "Princeton University",
                      "Yale University", "Quinnipiac University"), c("name","kq5_cond_parq1")]
```

```
##                       name kq5_cond_parq1
## 475     Cornell University      0.5935437
## 1452  Princeton University      0.6586524
## 1462 Quinnipiac University      0.4852847
## 2191       Yale University      0.5730308
```

Question 2.2.2.c The higher the access rate, the lower the success rate, implying that access rate is negatively correlated with success rate. There are very few colleges who do well on both access and success rates. Most do very poorly on both metrics.

```
plot(data$par_q1, data$kq5_cond_parq1, xlab = "Access Rate", ylab = "Success Rate",
     main = "Access Rate vs Success Rate")
```



Question 2.2.2.d Min, Max, Mean, Median all shown below.

```
data$mr_kq5_pq1 <- data$par_q1*data$kq5_cond_parq1
summary(data$mr_kq5_pq1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.01031 0.01512 0.01827 0.02184 0.16358
```

Question 2.2.2.e

```
sorted2 <- data[order(data$mr_kq5_pq1),]
#Here are the highest 10
print(sorted2[2193:2202,c("name","mr_kq5_pq1")])
```

```
##                                                                    name
## 1253 New York City College Of Technology Of The City University Of New
## 1755                        State University Of New York At Stony Brook
## 1384                                                    Pace University
## 1044                                                   MCPHS University
## 227                              CUNY John Jay College Of Criminal Justice
## 258                               California State University, Los Angeles
## 229                                                    CUNY Lehman College
## 355                                  City College Of New York - CUNY
## 222                                   CUNY Bernard M. Baruch College
## 2044                        Vaughn College Of Aeronautics And Technology
```

```
##      mr_kq5_pq1
## 1253 0.08334075
## 1755 0.08412748
## 1384 0.08432647
## 1044 0.09343507
## 227  0.09691438
## 258  0.09918455
## 229  0.10235137
## 355  0.11723747
## 222  0.12938586
## 2044 0.16357975
```

```
# Here are the lowest 10
print(sorted2[1:10,c("name","mr_kq5_pq1")])
```

```
##                                                      name mr_kq5_pq1
## 129                    Bel - Rea Institute Of Animal Technology          0
## 277                              Capri Institute Of Hair Design          0
## 373          Cleveland Institute Of Dental - Medical Assistants          0
## 612                          Empire Beauty School of Portland, ME          0
## 645                             Florida College Of Natural Health          0
## 910                                   Kansas Wesleyan University          0
## 965                                            Landmark College          0
## 975                            Latter Day Saints Business College          0
## 1096                               McNally Smith College Of Music          0
## 1134 Midred Elley College And Austin's School Of Spa Technology          0
```

Question 2.2.2.f Yale is second highest in mobility compared to the four.

```
sorted2[sorted2$name %in% c("Cornell University", "Princeton University",
                            "Yale University", "Quinnipiac University"), c("name", "mr_kq5_pq1")]
```

```
##                        name  mr_kq5_pq1
## 1462 Quinnipiac University 0.009007483
## 1452  Princeton University 0.013457949
## 2191       Yale University 0.020817068
## 475     Cornell University 0.029070235
```

Question 2.3.1

```
table1 <- as.data.frame(table(data$type))
rownames(table1) = c("public", "private non-profit", "for-profit")
colnames(table1) = c("Var1","Freq")
print(table1)
```

```
##                     Var1 Freq
## public                 1 1190
## private non-profit     2  819
## for-profit             3  190
```

```
table1 <- table1[-c(3,4)]
```

Question 2.3.2

```
table1$mobility <- 1
table1[1,3] <- mean(data[data$type == 1,]$mr_kq5_pq1, na.rm = TRUE)
table1[2,3] <- mean(data[data$type == 2,]$mr_kq5_pq1, na.rm = TRUE)
table1[3,3] <- mean(data[data$type == 3,]$mr_kq5_pq1, na.rm = TRUE)
print(table1)
```

```
##                     Var1 Freq   mobility
## public                 1 1190 0.01916873
## private non-profit     2  819 0.01724121
## for-profit             3  190 0.01715562
```

Question 2.3.3

```
print(sorted2[2193:2202,c("name","tier_name")])
```

```
##                                                                    name
## 1253 New York City College Of Technology Of The City University Of New
## 1755                        State University Of New York At Stony Brook
## 1384                                                  Pace University
## 1044                                                MCPHS University
## 227                           CUNY John Jay College Of Criminal Justice
## 258                            California State University, Los Angeles
## 229                                             CUNY Lehman College
## 355                                    City College Of New York - CUNY
## 222                                     CUNY Bernard M. Baruch College
## 2044                         Vaughn College Of Aeronautics And Technology
##                                    tier_name
## 1253                          Selective public
## 1755                   Highly selective public
## 1384                         Selective private
## 1044 Nonselective four-year private not-for-profit
## 227                           Selective public
## 258                           Selective public
## 229                           Selective public
## 355                           Selective public
## 222                           Selective public
## 2044 Nonselective four-year private not-for-profit
```

Question 2.3.4 From 2.3.2, we found that public colleges have the highest mobility rate among the three "types". In addition, we find in 2.3.3 that Selective public colleges dominate the 10 colleges with the highest mobility. Therefore I think that public "type" colleges have the most contribution to mobility.