

Appendix

Sampling Procedures

```
> # List of the department
> department <- c("AFST", "AMST", "ARBC", "ARTH", "ASST", "ASTR",
+ "BIOL", "CHEM", "CHIN", "CAMS", "CLAS", "CGSC", "CS", "CCST",
+ "DANC", "ECON", "EDUC", "ENGL", "ENTS", "EUST", "FREN", "GEOL",
+ "GERM", "GRK", "HIST", "IDSC", "JAPN", "LATN", "LTAM", "LING",
+ "LCST", "MATH", "MUSC", "NEUR", "PHIL", "PE", "PHYS", "POSC",
+ "PSYC", "RELG", "RUSS", "SOAN", "SPAN", "ARTS", "THEA", "WGST")
>
> mi <- 2 # size of ssus
> n <- 20 # size of psus
>
> set.seed(50)
> clus <- sample(department, n, replace = F) # SRS of the departments with size = 20
> clus # department sampled
[1] "MUSC" "EUST" "CHIN" "WGST" "GEOL" "AMST" "LATN" "IDSC" "RUSS" "ARTH"
[11] "DANC" "CAMS" "SOAN" "ARBC" "ARTS" "FREN" "PSYC" "CLAS" "SPAN" "ASST"
>
> set.seed(45)
> danc <- sample(10, mi, replace = F) # SRS of the classes with size = 2
> danc # row number of the sampled classes the list of classes
[1] 7 3
> danc.data <- data.frame(size = c(20, 25), reg = c(15, 11), dept = "danc",
+ Mi = 10)
>
> clas <- sample(4, mi, replace = F) # SRS of the classes with size = 2
> clas # row number of the sampled classes the list of classes
[1] 1 2
> clas.data <- data.frame(size = c(25, 15), reg = c(23, 15), dept = "clas",
+ Mi = 4)
>
> cams <- sample(15, mi, replace = F) # SRS of the classes with size = 2
> cams # row number of the sampled classes the list of classes
[1] 6 5
> cams.data <- data.frame(size = c(25, 0), reg = c(28, 11), dept = "cams",
+ Mi = 15)
>
> geol <- sample(11, mi, replace = F) # SRS of the classes with size = 2
> geol # row number of the sampled classes the list of classes
[1] 3 6
> geol.data <- data.frame(size = c(20, 18), reg = c(25, 17), dept = "geol",
+ Mi = 11)
>
> musc <- sample(220, mi, replace = F) # SRS of the classes with size = 2
> musc # row number of the sampled classes the list of classes
[1] 41 2
> musc.data <- data.frame(size = c(50, 50), reg = c(1, 0), dept = "musc",
+ Mi = 220)
>
> eust <- sample(2, mi, replace = F) # SRS of the classes with size = 2
```

```

> eust.data <- data.frame(size = c(15, 24), reg = c(15, 22), dept = "eust",
+   Mi = 2)
>
> chin <- sample(6, mi, replace = F) # SRS of the classes with size = 2
> chin # row number of the sampled classes the list of classes
[1] 2 3
> chin.data <- data.frame(size = c(16, 25), reg = c(2, 9), dept = "chin",
+   Mi = 6)
>
> wgst <- sample(8, mi, replace = F) # SRS of the classes with size = 2
> wgst # row number of the sampled classes the list of classes
[1] 4 3
> wgst.data <- data.frame(size = c(25, 30), reg = c(7, 10), dept = "wgst",
+   Mi = 8)
>
> amst <- sample(3, mi, replace = F) # SRS of the classes with size = 2
> amst # row number of the sampled classes the list of classes
[1] 3 1
> amst.data <- data.frame(size = c(25, 25), reg = c(15, 10), dept = "amst",
+   Mi = 3)
>
> latn <- sample(3, mi, replace = F) # SRS of the classes with size = 2
> latn # row number of the sampled classes the list of classes
[1] 2 1
> latn.data <- data.frame(size = c(25, 25), reg = c(14, 8), dept = "latn",
+   Mi = 3)
>
> idsc <- sample(14, mi, replace = F) # SRS of the classes with size = 2
> idsc # row number of the sampled classes the list of classes
[1] 2 5
> idsc.data <- data.frame(size = c(0, 15), reg = c(4, 8), dept = "idsc",
+   Mi = 14)
>
> russ <- sample(5, mi, replace = F) # SRS of the classes with size = 2
> russ # row number of the sampled classes the list of classes
[1] 3 5
> russ.data <- data.frame(size = c(20, 10), reg = c(7, 3), dept = "russ",
+   Mi = 5)
>
> arth <- sample(10, mi, replace = F) # SRS of the classes with size = 2
> arth # row number of the sampled classes the list of classes
[1] 9 1
> arth.data <- data.frame(size = c(25, 25), reg = c(22, 20), dept = "arth",
+   Mi = 10)
>
> soan <- sample(15, mi, replace = F) # SRS of the classes with size = 2
> soan # row number of the sampled classes the list of classes
[1] 15 11
> soan.data <- data.frame(size = c(25, 25), reg = c(6, 1), dept = "soan",
+   Mi = 15)
>
> arbc <- sample(3, mi, replace = F) # SRS of the classes with size = 2
> arbc # row number of the sampled classes the list of classes

```

```

[1] 2 1
> arbc.data <- data.frame(size = c(16, 25), reg = c(17, 15), dept = "arbc",
+   Mi = 3)
>
> arts <- sample(11, mi, replace = F) # SRS of the classes with size = 2
> arts # row number of the sampled classes the list of classes
[1] 5 4
> arts.data <- data.frame(size = c(11, 14), reg = c(14, 10), dept = "arts",
+   Mi = 11)
>
> fren <- sample(11, mi, replace = F) # SRS of the classes with size = 2
> fren # row number of the sampled classes the list of classes
[1] 6 9
> fren.data <- data.frame(size = c(15, 20), reg = c(9, 9), dept = "fren",
+   Mi = 11)
>
> psyc <- sample(17, mi, replace = F) # SRS of the classes with size = 2
> psyc # row number of the sampled classes the list of classes
[1] 8 9
> psyc.data <- data.frame(size = c(1, 30), reg = c(1, 17), dept = "psyc",
+   Mi = 17)
>
> span <- sample(22, mi, replace = F) # SRS of the classes with size = 2
> span # row number of the sampled classes the list of classes
[1] 22 7
> span.data <- data.frame(size = c(20, 16), reg = c(19, 16), dept = "span",
+   Mi = 22)
>
> asst <- sample(4, mi, replace = F) # SRS of the classes with size = 2
> asst # row number of the sampled classes the list of classes
[1] 1 2
> asst.data <- data.frame(size = c(25, 25), reg = c(5, 3), dept = "asst",
+   Mi = 4)
>
> # Combine all the data
> data.class <- rbind(asst.data, span.data, psyc.data, fren.data,
+   arts.data, arbc.data, soan.data, arth.data, russ.data, idsc.data,
+   latn.data, amst.data, wgst.data, chin.data, eust.data, musc.data,
+   geol.data, cams.data, clas.data, danc.data)

```

Data Analysis

```

> data.class$mi <- mi # mi = 2
> data.class$N <- 46
> data.class$n <- n # n = 20
> data.class$wts <- data.class$N/data.class$n * data.class$Mi/data.class$mi # sampling weights
> data.class$elem.id <- 1:nrow(data.class)
> data.class$less20 <- ifelse(data.class$reg < 20, 1, 0) # registered size < 20
> data.class %>% group_by(dept) %>% summarize(n = n()) # test
# A tibble: 20 x 2
  dept      n
<fct> <int>

```

```

1 asst      2
2 span      2
3 psyc      2
4 fren      2
5 arts      2
6 arbc      2
7 soan      2
8 arth      2
9 russ      2
10 idsc     2
11 latn     2
12 amst     2
13 wgst     2
14 chin     2
15 eust     2
16 musc     2
17 geol     2
18 cams     2
19 clas     2
20 danc     2
>
> # survey object
> class.clus <- svydesign(id = ~dept + elem.id, fpc = ~N + Mi,
+   weights = ~wts, data = data.class)
>
> # average number of seats available in the class
> (size.est <- svymean(~size, class.clus, deff = T))
      mean      SE  DEff
size 36.0140  6.2948 5.7857
> confint(size.est)
      2.5 %   97.5 %
size 23.67643 48.35149
>
> # average number of registered students in each class
> (reg.est <- svymean(~reg, class.clus, deff = T))
      mean      SE  DEff
reg  5.7170  2.4259 4.1611
> confint(reg.est)
      2.5 %   97.5 %
reg 0.9623972 10.47161
>
> # ratio of registered size to the number of available seats
> (ratio.est <- svyratio(~reg, ~size, class.clus))
Ratio estimator: svyratio.survey.design2(~reg, ~size, class.clus)
Ratios=
      size
reg 0.1587441
SEs=
      size
reg 0.09406127
> confint(ratio.est)
      2.5 %   97.5 %
reg/size -0.02561257 0.3431008

```

```

>
> # proportion of classes that have registered enrollments less
> # than 20 students
> (less20.est <- svymean(~less20, class.clus, deff = T))
      mean      SE  DEff
less20 0.934010 0.040574 1.0898
> confint(less20.est)
      2.5 %    97.5 %
less20 0.8544861 1.013534

```

Summary Stats

```

> stargazer(data.class, type = "text")
=====
Statistic N    Mean  St. Dev.  Min  Pctl(25)  Pctl(75)  Max
-----
size      40 21.150  9.991    0    15.8     25     50
reg       40 11.600  7.228    0     6.8    16.2    28
Mi        40 19.700 46.854    2     4     14.2   220
mi        40 2.000  0.000    2     2     2     2
N         40 46.000  0.000   46    46     46    46
n         40 20.000  0.000   20    20     20    20
wts       40 22.655 53.882  2.300  4.600   16.387 253.000
elem.id   40 20.500 11.690    1    10.8    30.2    40
less20    40 0.850  0.362    0     1     1     1
=====

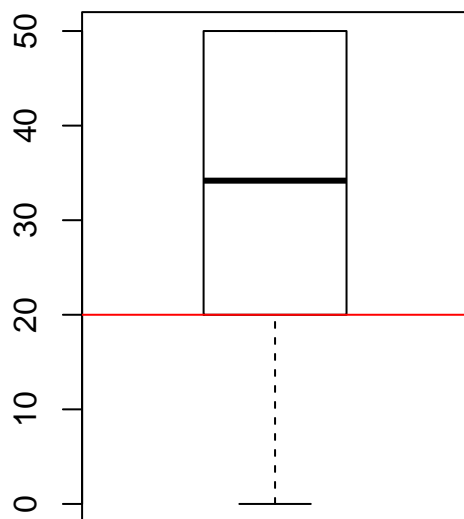
> data.class %>% group_by(dept) %>% summarize(Size = mean(size),
+   Registered = mean(reg))
# A tibble: 20 x 3
  dept    Size Registered
  <fct> <dbl>      <dbl>
1 asst   25         4
2 span   18       17.5
3 psyc   15.5        9
4 fren   17.5        9
5 arts   12.5       12
6 arbc   20.5       16
7 soan   25        3.5
8 arth   25       21
9 russ   15         5
10 idsc   7.5        6
11 latn   25       11
12 amst   25      12.5
13 wgst   27.5       8.5
14 chin   20.5       5.5
15 eust   19.5      18.5
16 musc   50        0.5
17 geol   19       21
18 cams   12.5     19.5
19 clas   20       19
20 danc   22.5      13

```

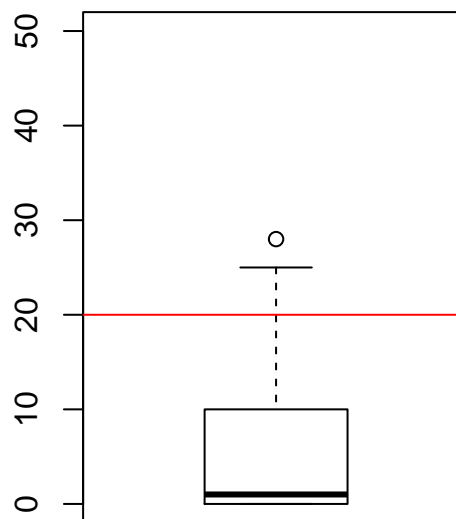
Graphing the Data

```
> par(mfrow = c(1, 2))
> svyboxplot(size ~ 1, class.clus, main = "Available Seats")
> abline(h = 20, col = "red")
> svyboxplot(reg ~ 1, class.clus, main = "Registered Class Size",
+   ylim = c(0, 50))
> abline(h = 20, col = "red")
```

Available Seats



Registered Class Size



```
> par(mfrow = c(1, 1))
>
> g1 <- ggplot(data.class, aes(x = size, y = reg)) + geom_point() +
+   geom_smooth() + labs(x = "Available Seats", y = "Registered Class Size")
>
> cor(data.class$size, data.class$reg)
[1] -0.03713851
>
> # Omit the outliers
> data.class2 <- data.class %>% filter(dept != "musc")
>
> g2 <- ggplot(data.class2, aes(x = size, y = reg)) + geom_point() +
+   geom_smooth() + labs(x = "Available Seats", y = "Registered Class Size")
>
> cor(data.class2$size, data.class2$reg)
[1] 0.2919652
>
> grid.arrange(g1, g2, ncol = 2)
`geom_smooth()` using method = 'loess' and formula 'y ~ x'
`geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

