```
Stat 145, Tue 31-Aug-2021 -- Tue 31-Aug-2021
Biostatistics
Spring 2021



------------------------------
Tuesday, August 31st 2021
------------------------------
Topic:: Data
Read::  Lock5 1.2
HW::   PS00

How multivariate data is arranged
 - look at iris, AllCountries
 - variables vs. values
     Socrative quiz: STAT 143: Value or Variable
     Room SCOFIELD188
 - rows for cases/subjects/units; columns for variables
     Socrative quiz: Lock5-1.1, questions 1-4
 - sample vs population
     statistical inference
     simple random samples
     bias
       Socrative quiz: Lock5-1.1, question 5
       Socrative quiz: Lock5-1.2

New R commands:
 help():
   help(nrow)   gives help about a command
   help(faithful)   explains dataset; not always available
 read.csv():
   ssurv <- read.csv("http://scofield.site/teaching/data/csv/ssurv.csv")
 nrow(): nrow(ssurv)
 filter():  filter(ssurv, speedtickets >= 3)
 subset():  subset(ssurv, select=c(oncampus, speedtickets))
 sort(~cds, data=ssurv)  does not work, though it feels like it should
     with(ssurv, sort(cds))    does what the above appears to intend
     sort(ssurv$cds)       this also works
 tally():
   tally(~class, data=ssurv)
   tally(~randomnum, data=ssurv)    quantitative var., but meaningful result
```

```
  tally(~gpa, data=ssurv)     quantitative var., cluttery result
gf_bar():  gf_bar(~class, data=ssurv)
gf_histogram(): gf_histogram(~speedtickets, data=ssurv)
```