Chapter 5
Sampling: Surveys and How to Ask Questions

Section 5.3

Simple random sampling is easy to perform in R. In this example we will randomly choose 10 students weekly television watching amounts from the dataset UCDavis1.RData. The first step will be to find the number of students \( n \) from which we are sampling. A vector of ID labels, 1 through \( n \), will be created from which 10 numbers will be sampled without replacement. These 10 randomly selected ID labels will then be used to extract the TV watching amounts from the corresponding 10 students. The R command used most often for random sampling is `sample()`. You give the function a vector of values from which to sample and then can give options that specify the number of values to sample (`size`), and whether or not to sample with or without replacement (`replace`, default is `F` which is sampling without replacement). If something more complicated than simple random sampling is required, you can also specify a vector of unequal sampling probabilities using the `prob` option. In simple random sampling, the sampling probabilities are all equal. The below screen copy demonstrates simple random sampling. Note that you may not be able to exactly reproduce the results below because the `sample()` function is random.

```
> load("C:/RData/UCDavis1.RData")
> names( ucdavis1 )
[1] "m" "o" "g" "a" "m" "a" "g" "a" "m" "o"
> attach( ucdavis1 )
> length( t )  # find number of students
[1] 173
>
> IDsample <- sample( 1:173, size=10 )  # sample from numbers 1 through 173
> IDsample  
[1] 135 15 164 28 170 161 116 46 107 188

> t[ IDsample ]  # their TV amounts
[1] 4.0 10.0 3.8 30.0 2.0 10.0 24.0 10.0 20.0 10.0
```
