

## Introduction to Data Science: Stat-lab 7

### INSTRUCTIONS:

Type each one of the R commands in this unit on the space provided below.

You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##   Mean  :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
##   Max.  :25.0    Max.    :120.00
```

After you are done click the **Knit PDF** or **Knit HTML** button. A document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

Go to File -> Save. Name your file.

Upload the file to the Moodle site by using the link provided.

R commands:

```
USStatePops=read.table(file="poptable.txt",sep=",")
sample(USStatePops$V2, size = 16, replace=TRUE)
```

```
## [1] 601723 6392017 5988927 6346105 1826341 3751351 563626 897934 6346105
## [10] 4339367 5773552 2700551 1826341 5303925 601723 1052567
```

```
mean(sample(USStatePops$V2,size=16,replace=TRUE))
```

```
## [1] 3864762
```

```
replicate(4, mean(sample(USStatePops$V2,size=16,replace=TRUE)),simplify = TRUE)
```

```
## [1] 8278994 9622115 4953612 4365354
```

```
mean(replicate(4, mean(sample(USStatePops$V2,size=16,replace=TRUE)),simplify = TRUE))
```

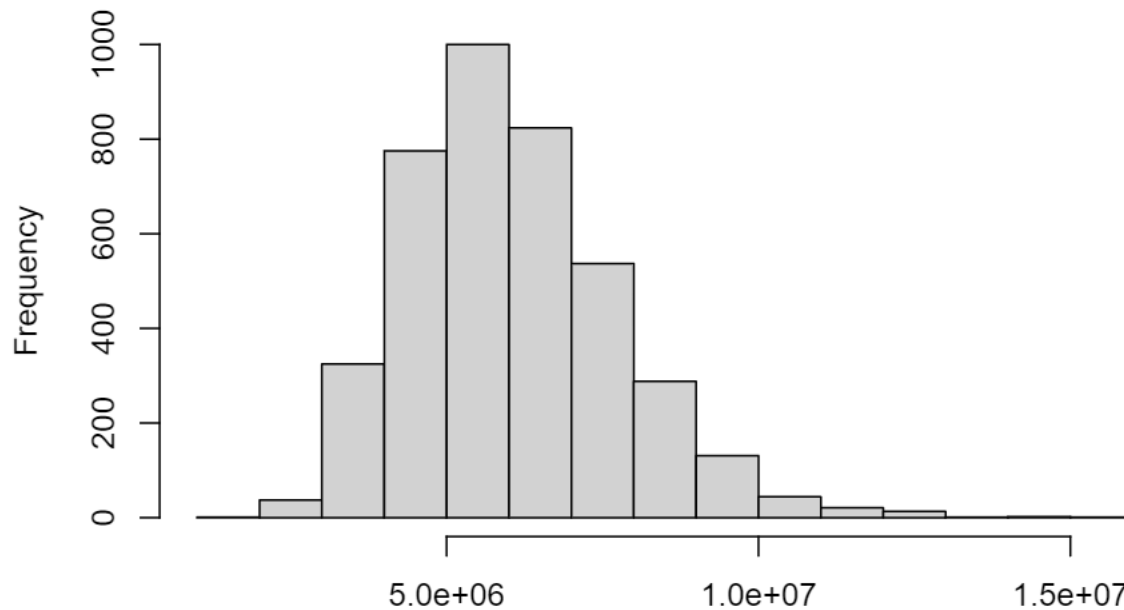
```
## [1] 6914275
```

```
mean(replicate(4000, mean(sample(USStatePops$V2,size=16,replace=TRUE)),simplify = TRUE))
```

```
## [1] 6005063
```

```
hist(replicate(4000, mean(sample(USStatePops$V2,size=16,replace=TRUE)),simplify = TRUE))
```

```
replicate(4000, mean(sample(USStatePops$V2, size = 16, replace = TRUE
```



```
replicate(4000, mean(sample(USStatePops$V2, size = 16, replace = TRUE)), simplify = 1
```

```
mean(replicate(100, mean(sample(USStatePops$V2, size=51, replace=TRUE)), simplify = TRUE))
```

```
## [1] 6066843
```

```
mean(replicate(100, mean(sample(USStatePops$V2, size=120, replace=TRUE)), simplify = TRUE))
```

```
## [1] 6022161
```

```
SampleMeans <- replicate(10000, mean(sample(USStatePops$V2, size = 5, replace=TRUE)), simplify=TRUE)
length(SampleMeans)
```

```
## [1] 10000
```

```
mean(SampleMeans)
```

```
## [1] 6100590
```

```
summary(SampleMeans)
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
##  694116  3871699  5401418  6100590  7641086 24321266
```

```
quantile(SampleMeans, probe=c(0.25,0.50,0.75))
```

```
##      0%      25%      50%      75%     100%
##  694116  3871699  5401418  7641086 24321266
```

```
quantile(SampleMeans, probs=c(0.025,0.975))
```

```
##      2.5%     97.5%
## 1981529 13537626
```

```
MysterySample <- c(3706690, 159358, 106405, 55519, 53883)
mean(MysterySample)
```

```
## [1] 816371
```

```
quantile(SampleMeans, probs=c(0.005,0.995))
```

```
##      0.5%      99.5%
```

```
## 1407120 16753481
```

```
sd(SampleMeans)
```

```
## [1] 3060716
```

```
sd(USStatePops$V2/sqrt(5))
```

```
## [1] 3051779
```

```
StdError<-sd(USStatePops$V2/sqrt(5))
```

```
CutPoints975<-mean(USStatePops$V2)+(2 * StdError)
```

```
CutPoints975
```

```
## [1] 12157391
```