Assignment3_stat123

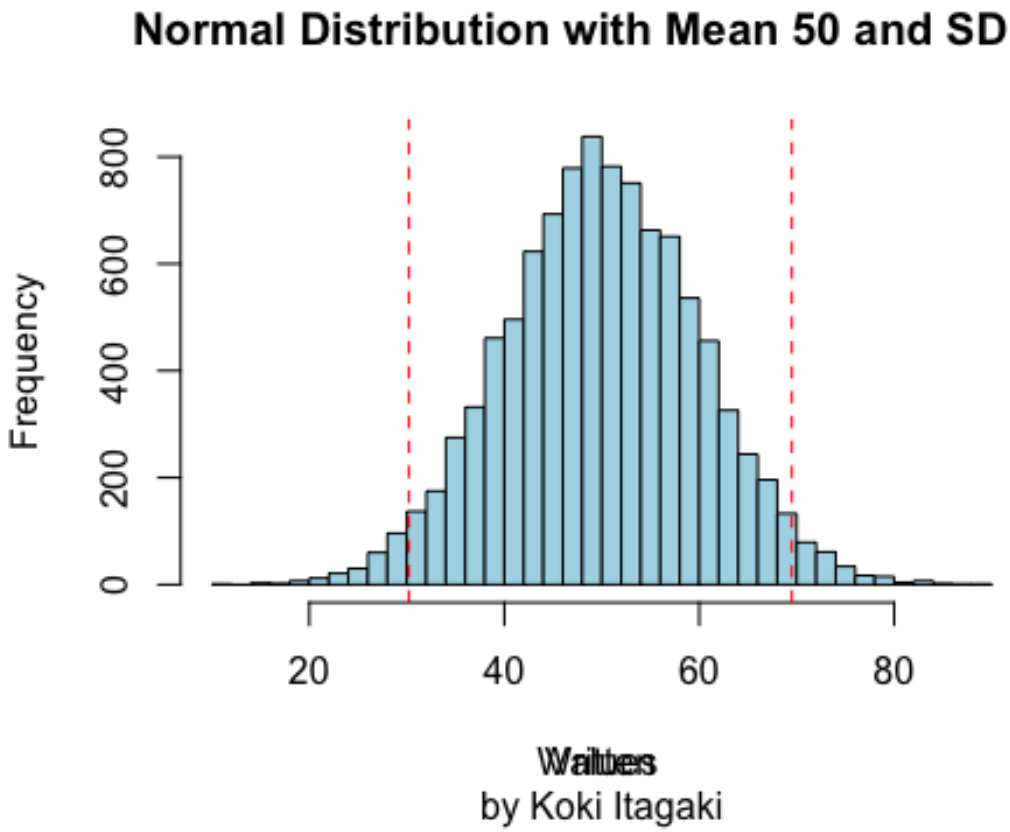
Koki Itagaki

2023-03-09

#Q1 The commands generate a normal distribution with a mean of 50 and a standard deviation of 10. Create a histogram to show the range of values that covers the middle 95% of this distribution, using the title of “Normal Distribution with Mean 50 and SD 10”, light blue color, and breaks equal to 30.

```
set.seed(123)
data <- rnorm(10000, mean = 50, sd = 10)
hist(data, main ="Normal Distribution with Mean 50 and SD 10", sub = "Written
by Koki Itagaki",xlab = "Values", ylab = "Frequency",breaks = 30,
col = "lightblue")

q <- quantile(data, c(0.025, 0.975))
abline(v = q, lty = 2, col = "red")
```



2. Download the data sets sampledata.csv with 100 numeric variables, name it mydata and save it to whatever directory you are using for this question.

```
set.seed(123)
mydata<- read.csv("/Users/itagakikouki/stat123/assignment3/sampledata.csv")

head(mydata)
```

```
##   X        X1        X2        X3        X4        X5       X6        X7
## 1 1 10.353165 11.246402  8.734532 10.224021  7.624291 9.594864  8.246859
## 2 2 11.436858  8.889904 11.844420 10.252278 10.085732 6.283675 12.421713
## 3 3  9.376116  9.847429 10.470260  9.700535 13.931214 9.757506 10.450268
## 4 4 12.710452 10.142229  8.699626 15.906552  9.582598 8.832959 10.876766
## 5 5 12.216360 10.074403 10.276487 10.954854  9.371108 7.068743  8.748403
## 6 6 12.056541  8.121844  9.746651  9.014022  8.680201 9.284938  9.969255
##          X8        X9       X10      X11       X12       X13       X14
## 1 10.122774  6.324963  9.037005 11.10702 10.471290  7.754075  5.815037
## 2  9.059512 15.250142 11.375589 10.13844 11.695332  9.520430  7.308356
## 3  6.977037 12.581651 10.976542 10.11821 10.068704 11.870623  9.392137
## 4 12.010535  9.333234  8.776299  8.31860 10.572371  9.533327  7.547380
## 5  9.035347  7.196679 10.928936 13.97744  9.603742  7.142955  6.552112
## 6 11.967442 13.685087  8.523867 13.64634 10.740660 11.893536 11.053279
##         X15       X16       X17       X18       X19      X20       X21
## 1 10.717503 10.907561 11.725619 12.264639 10.374095 13.79291 11.331133
```

```
## 2 11.850201  7.613318  9.679526  8.591416 13.793210 13.11931 11.985109
## 3 11.723977  9.646584 12.299709 10.139946  7.411961 10.97708  8.741936
## 4 11.211317  8.945622  9.624904 11.252967  8.930406 11.83010  6.719885
## 5 11.141444 12.583906  8.899751  8.967542 12.544281 10.36405 11.905543
## 6  7.711188 10.125393 11.792976 10.795670  7.695228 10.42247 11.984309
##         X22       X23       X24       X25       X26       X27       X28
## 1  9.790344  9.536986  9.239807  7.333545  9.449976 10.167507  8.310950
## 2  7.925528  8.127395 10.532521 10.268498 10.081694 14.374780  7.317465
## 3 11.881832 12.597102 10.802889 10.245241 10.235828 10.271030  8.864447
## 4 10.333023 11.106808  9.284770  7.404019 11.542562  8.258747 10.912984
## 5  6.558825  9.618592  8.242858  8.823469  8.464498 11.273266  8.221997
## 6 15.109425  8.003208 12.164444 10.971113 13.492850  5.999478 13.307765
##         X29       X30      X31       X32       X33       X34       X35
## 1 11.530102 10.161966 12.94311 10.596824  7.585470  4.468858  5.258602
## 2 11.184359  8.509448 11.14295 12.761768  8.175802  9.889200 12.178901
## 3  9.625953  9.148622 12.35236 11.905947 11.464857  9.690780 14.336223
## 4  9.009819  9.404370 10.13220 11.198459 10.410736 10.940265 10.075461
## 5  8.255211  9.730142 10.27867 13.970744  8.048338 12.249828  7.396805
## 6 11.654445  8.109379 13.55904  8.977603  5.109642 10.646140  9.669790
##         X36       X37       X38       X39       X40      X41       X42
## 1 12.118949  6.055851 10.462612 11.253287  9.604470 12.14812 11.988954
## 2 10.896306  6.435955 12.306634 11.689357 11.049657 10.48228  9.728771
## 3 12.267130  9.641855  8.489463  8.334252  7.003536 10.20935  8.342157
## 4  9.198834 10.817856 12.104468  9.052364 10.283094 11.28058 12.550745
## 5 10.265327  5.873353  9.670285 12.474431 10.296376 11.04587 11.637294
## 6  8.798843  8.164297 10.432105 13.648419  9.631612 10.85445  8.345753
##         X43       X44       X45       X46       X47       X48       X49
## 1  8.406148 12.364601  7.540592  9.867355  6.499486  7.973895 10.918302
## 2 10.573763 11.668280 10.158869  7.648090 11.287619 14.745091 11.767352
## 3 13.030389 12.620543 10.820426 10.295454 12.130683  8.909452 10.957665
## 4  7.259144 12.983320  8.384981 13.332883  8.429785  9.231571  7.336664
## 5  8.698656  8.327978 13.148115 11.338862  8.125283  9.578749 10.155500
## 6 12.857847  9.636928 10.601212  9.936401 10.700537 11.294522 12.915806
##         X50       X51       X52       X53       X54       X55       X56
## 1  9.046154  8.729372 12.224644  7.503507  8.242156 10.497857 10.654315
## 2 13.980654 12.284227 10.565644  9.172907  9.704614 10.451013  9.349013
## 3 10.304985 10.393650  9.503623 12.138279 10.895544 11.244538 10.627914
## 4  8.449591  8.265459  9.013695  8.597805 11.682322  8.720456 15.922070
## 5 10.896211  7.614882  6.981698 12.179579  6.758287  9.164894  9.593103
## 6 10.896749  9.650549  8.225448 10.744416  9.752940 12.814967  9.804589
##         X57       X58       X59       X60       X61       X62       X63
## 1 11.616781  8.972562 11.875917  8.796707  8.252482 12.410058 12.396276
## 2  9.143754  9.730134  9.327363  8.450145  7.666617  8.504043  8.674451
## 3 14.277069 10.548518  8.456704  8.790065 10.412393  8.121105  8.542378
## 4  6.354094 12.050382  8.874999  7.544598 11.152939  4.442488  7.329739
## 5  9.750462 12.258476  8.472039  8.119747  8.088582 10.313286 10.089516
## 6  6.910074 11.556100 11.697855 10.745171 13.810482  8.888217  7.740175
##         X64       X65       X66       X67       X68       X69       X70
## 1 10.359664 11.819065  8.228900  9.247741  9.536222  6.352703 10.147951
## 2 10.128446  7.049142  8.767323 12.357671  6.226217 11.204706 10.913352
```

```
## 3 12.204197  9.037229  6.136217 10.652554 10.349069  8.472348 10.641109
## 4  6.638437  9.256400  6.839837  8.059895 11.909612  8.555874  9.482450
## 5  8.661424 10.194384 13.818475 10.412335  6.542613  9.632963  6.873860
## 6  9.004504  9.174873  5.166197  9.946845 12.220979 10.645811  7.298632
##         X71       X72       X73       X74       X75       X76       X77
## 1  6.275873  8.996379  5.809302 14.415275  7.252148 10.454983  6.764762
## 2  7.374602 15.823182  9.025264  6.656874  8.510868 11.343271 11.792683
## 3  9.570314  7.510254 13.222437 15.805248 10.205701 10.577600  9.176723
## 4 10.676949 11.010357 10.591856  7.942463 10.979675 12.360937  9.778166
## 5 10.286598 11.060039  9.589705 11.324309  8.574804  9.345460  7.749151
## 6 12.492055  7.662230 11.582389 11.395684 10.479052  9.717727 10.461672
##         X78       X79       X80       X81       X82       X83       X84
## 1  7.430603 10.463782 10.261311  8.183003 12.742443  9.010159 14.609614
## 2  7.535016  8.926836 10.485026  9.778291  7.725968  6.744816  9.649889
## 3 12.678233  9.367299 11.730008  8.065484 10.076085  7.844371  9.059496
## 4  9.270345 12.442927  8.576436 11.576627  7.109164  7.285485  8.415209
## 5 10.935537 12.040561  8.288863  8.727212  9.734642  9.675220 11.629345
## 6  9.391531  9.489748 11.584679  6.660051 10.569495 10.757783  8.200979
##         X85       X86       X87       X88       X89       X90       X91
## 1 10.389524 12.231845  8.591748  7.507866 10.854535 10.583506  7.979616
## 2  8.906829  8.366771  8.882704 10.024477 10.901538  8.534206  9.727534
## 3 10.541930 11.415799 12.261308  7.850642 10.323975  4.593684 11.346825
## 4 11.741792  8.195738 10.408802  8.804237 10.423320 10.865386 12.500771
## 5 10.175239  9.244368  8.862222  8.388238  8.556853 10.417531 12.277727
## 6  8.428789 10.453316 10.195435 11.224810 10.449550  9.861774 11.500738
##         X92       X93       X94       X95       X96       X97       X98
## 1  9.552379  8.931482  7.659649 10.285336 11.425415 12.680207 13.271138
## 2 12.159421 10.792899 13.472575 11.277980  8.821800  5.655376 12.419709
## 3 10.843115 13.801263 10.728009 11.261597  9.837331  8.999988 11.545261
## 4 13.086591 10.310071  8.093910 11.534466  8.323641  8.747441 12.860878
## 5  9.954324 12.342984 12.225523  7.994352 10.159747 11.401719 11.417528
## 6  9.448319 11.204877  9.358066 11.841902  9.413235 10.716352  9.379502
##        X99      X100
## 1 5.592469  7.085130
## 2 9.695320 10.573643
## 3 8.962925 10.706037
## 4 6.708526  8.690901
## 5 5.970788 12.367359
## 6 5.286078  9.788107
```

```
dim(mydata)
```

```
## [1]  50 101
```

```
#(a) Use a for loop to calculate the mean and standard deviation for each
variable in mydata.
```

```
means <- numeric(ncol(mydata))
sds <- numeric(ncol(mydata))
```

```r
for (i in 1:ncol(mydata)) {
  means[i] <- mean(mydata[,i])
  sds[i] <- sd(mydata[,i])
  print(paste("Mean for X ", i," is ", means[i]))
  print(paste0("Standard deviations for X",i," is ", sds[i]))

}
```

```
## [1] "Mean for X  1  is  25.5"
## [1] "Standard deviations for X1 is 14.5773797371133"
## [1] "Mean for X  2  is  10.6929409517177"
## [1] "Standard deviations for X2 is 2.19780623647846"
## [1] "Mean for X  3  is  9.92505174802658"
## [1] "Standard deviations for X3 is 2.20674143485657"
## [1] "Mean for X  4  is  10.3150304417091"
## [1] "Standard deviations for X4 is 1.7635044150341"
## [1] "Mean for X  5  is  10.9647123807944"
## [1] "Standard deviations for X5 is 1.96792208782613"
## [1] "Mean for X  6  is  10.2338888517815"
## [1] "Standard deviations for X6 is 1.99822331767603"
## [1] "Mean for X  7  is  9.87063597849014"
## [1] "Standard deviations for X7 is 2.08246742855641"
## [1] "Mean for X  8  is  9.8053712308362"
## [1] "Standard deviations for X8 is 2.0141955660897"
## [1] "Mean for X  9  is  9.64241379928302"
## [1] "Standard deviations for X9 is 1.76481590486114"
## [1] "Mean for X  10  is  10.0222819669712"
## [1] "Standard deviations for X10 is 2.37247514949163"
## [1] "Mean for X  11  is  10.2130383255438"
## [1] "Standard deviations for X11 is 1.75768129047564"
## [1] "Mean for X  12  is  10.3076880014402"
## [1] "Standard deviations for X12 is 1.87266795963238"
## [1] "Mean for X  13  is  9.8146155897913"
## [1] "Standard deviations for X13 is 2.0385366032797"
## [1] "Mean for X  14  is  9.89365970535917"
## [1] "Standard deviations for X14 is 2.05957778413919"
## [1] "Mean for X  15  is  9.74916952308151"
## [1] "Standard deviations for X15 is 2.34817564843427"
## [1] "Mean for X  16  is  10.0968511503409"
## [1] "Standard deviations for X16 is 2.24219753398267"
## [1] "Mean for X  17  is  9.89377928434736"
## [1] "Standard deviations for X17 is 2.12034611508984"
## [1] "Mean for X  18  is  10.7011418953943"
## [1] "Standard deviations for X18 is 2.01843583629102"
## [1] "Mean for X  19  is  10.2619519605416"
## [1] "Standard deviations for X19 is 1.98809502868869"
## [1] "Mean for X  20  is  10.0752987725278"
## [1] "Standard deviations for X20 is 1.79471308302779"
## [1] "Mean for X  21  is  10.0490978511563"
```

```
## [1] "Standard deviations for X21 is 1.7560207552847"
## [1] "Mean for X  22  is  9.71722600477967"
## [1] "Standard deviations for X22 is 1.72821600640201"
## [1] "Mean for X  23  is  10.2928796735334"
## [1] "Standard deviations for X23 is 1.98465974377182"
## [1] "Mean for X  24  is  10.1330179797552"
## [1] "Standard deviations for X24 is 2.00127741003955"
## [1] "Mean for X  25  is  9.91043926456588"
## [1] "Standard deviations for X25 is 1.95002389648213"
## [1] "Mean for X  26  is  9.73942605893218"
## [1] "Standard deviations for X26 is 1.97204388752838"
## [1] "Mean for X  27  is  10.1997220287585"
## [1] "Standard deviations for X27 is 1.6375399091384"
## [1] "Mean for X  28  is  10.5555358619908"
## [1] "Standard deviations for X28 is 2.10721099583196"
## [1] "Mean for X  29  is  10.3634373871079"
## [1] "Standard deviations for X29 is 1.91382508313481"
## [1] "Mean for X  30  is  10.2060640776202"
## [1] "Standard deviations for X30 is 1.93087799535108"
## [1] "Mean for X  31  is  9.81610581953023"
## [1] "Standard deviations for X31 is 1.82267620425264"
## [1] "Mean for X  32  is  10.6625625577957"
## [1] "Standard deviations for X32 is 1.87806476780078"
## [1] "Mean for X  33  is  10.3343648917034"
## [1] "Standard deviations for X33 is 1.94776895968434"
## [1] "Mean for X  34  is  9.61255124335969"
## [1] "Standard deviations for X34 is 2.19205947693261"
## [1] "Mean for X  35  is  10.0919982654201"
## [1] "Standard deviations for X35 is 1.79826355014383"
## [1] "Mean for X  36  is  10.139857066439"
## [1] "Standard deviations for X36 is 2.37487482463593"
## [1] "Mean for X  37  is  10.0935179211966"
## [1] "Standard deviations for X37 is 2.11868033476247"
## [1] "Mean for X  38  is  9.60741511112754"
## [1] "Standard deviations for X38 is 2.19862421232863"
## [1] "Mean for X  39  is  10.0037056534431"
## [1] "Standard deviations for X39 is 1.86851003446559"
## [1] "Mean for X  40  is  10.0562183193546"
## [1] "Standard deviations for X40 is 2.28995371908233"
## [1] "Mean for X  41  is  9.84692517837302"
## [1] "Standard deviations for X41 is 1.93012324406212"
## [1] "Mean for X  42  is  9.90924647232305"
## [1] "Standard deviations for X42 is 2.09674871431367"
## [1] "Mean for X  43  is  9.75047036843201"
## [1] "Standard deviations for X43 is 2.16686835701963"
## [1] "Mean for X  44  is  9.75477342879049"
## [1] "Standard deviations for X44 is 1.89291434206125"
## [1] "Mean for X  45  is  10.0770174974138"
## [1] "Standard deviations for X45 is 2.01524600457948"
## [1] "Mean for X  46  is  10.050344316637"
```

```
## [1] "Standard deviations for X46 is 1.47236543491801"
## [1] "Mean for X  47  is  10.1882171744323"
## [1] "Standard deviations for X47 is 2.24607074104989"
## [1] "Mean for X  48  is  9.80278287349694"
## [1] "Standard deviations for X48 is 2.0313411128133"
## [1] "Mean for X  49  is  9.55261793255721"
## [1] "Standard deviations for X49 is 2.00691068732325"
## [1] "Mean for X  50  is  9.91745394789796"
## [1] "Standard deviations for X50 is 2.14445126395144"
## [1] "Mean for X  51  is  10.1698428864992"
## [1] "Standard deviations for X51 is 1.96416913603995"
## [1] "Mean for X  52  is  9.85602738614628"
## [1] "Standard deviations for X52 is 1.77910027071599"
## [1] "Mean for X  53  is  9.29944213230399"
## [1] "Standard deviations for X53 is 2.06983514555851"
## [1] "Mean for X  54  is  10.0456986299851"
## [1] "Standard deviations for X54 is 1.87441866971326"
## [1] "Mean for X  55  is  10.1843465940327"
## [1] "Standard deviations for X55 is 2.03631152011921"
## [1] "Mean for X  56  is  9.88597013348946"
## [1] "Standard deviations for X56 is 1.71752527294391"
## [1] "Mean for X  57  is  9.85273827976136"
## [1] "Standard deviations for X57 is 2.18970134029823"
## [1] "Mean for X  58  is  9.83853322627137"
## [1] "Standard deviations for X58 is 2.06639082115625"
## [1] "Mean for X  59  is  9.83221977315724"
## [1] "Standard deviations for X59 is 2.21243490346657"
## [1] "Mean for X  60  is  9.67489191204435"
## [1] "Standard deviations for X60 is 1.75747697992272"
## [1] "Mean for X  61  is  9.9117342280433"
## [1] "Standard deviations for X61 is 1.99960735693599"
## [1] "Mean for X  62  is  9.66954311794616"
## [1] "Standard deviations for X62 is 1.95473649433698"
## [1] "Mean for X  63  is  9.7445091771929"
## [1] "Standard deviations for X63 is 1.84656938531263"
## [1] "Mean for X  64  is  10.2459689621982"
## [1] "Standard deviations for X64 is 2.02595484823864"
## [1] "Mean for X  65  is  9.93439552037724"
## [1] "Standard deviations for X65 is 1.69847243706218"
## [1] "Mean for X  66  is  10.177343259393"
## [1] "Standard deviations for X66 is 1.76527937512441"
## [1] "Mean for X  67  is  10.058951876421"
## [1] "Standard deviations for X67 is 2.28523913776151"
## [1] "Mean for X  68  is  9.85752624108566"
## [1] "Standard deviations for X68 is 2.20258183217853"
## [1] "Mean for X  69  is  9.85780549745033"
## [1] "Standard deviations for X69 is 2.08842309251934"
## [1] "Mean for X  70  is  9.600570502589"
## [1] "Standard deviations for X70 is 1.81690291233666"
## [1] "Mean for X  71  is  9.69638454324324"
```

```
## [1] "Standard deviations for X71 is 1.97517622385811"
## [1] "Mean for X  72  is  10.2330764236993"
## [1] "Standard deviations for X72 is 1.97104901017973"
## [1] "Mean for X  73  is  10.0273899232358"
## [1] "Standard deviations for X73 is 2.21717126504305"
## [1] "Mean for X  74  is  9.81283591274165"
## [1] "Standard deviations for X74 is 2.09409026712224"
## [1] "Mean for X  75  is  10.1915531028835"
## [1] "Standard deviations for X75 is 2.15188110525683"
## [1] "Mean for X  76  is  10.0544928470481"
## [1] "Standard deviations for X76 is 1.79189434890672"
## [1] "Mean for X  77  is  10.0673871213663"
## [1] "Standard deviations for X77 is 2.05238801428236"
## [1] "Mean for X  78  is  9.81779326085923"
## [1] "Standard deviations for X78 is 1.95355630313533"
## [1] "Mean for X  79  is  9.58930751522763"
## [1] "Standard deviations for X79 is 2.14423246753112"
## [1] "Mean for X  80  is  10.1884183185841"
## [1] "Standard deviations for X80 is 1.77456944097653"
## [1] "Mean for X  81  is  10.0068015977026"
## [1] "Standard deviations for X81 is 2.05989166147886"
## [1] "Mean for X  82  is  10.3014174429949"
## [1] "Standard deviations for X82 is 2.17742598732186"
## [1] "Mean for X  83  is  9.95389237256036"
## [1] "Standard deviations for X83 is 2.38851991257558"
## [1] "Mean for X  84  is  9.70886644026716"
## [1] "Standard deviations for X84 is 2.14491532877989"
## [1] "Mean for X  85  is  10.3455413640895"
## [1] "Standard deviations for X85 is 2.09339960301944"
## [1] "Mean for X  86  is  10.1957229170815"
## [1] "Standard deviations for X86 is 1.92525057839115"
## [1] "Mean for X  87  is  9.92191144874065"
## [1] "Standard deviations for X87 is 2.02073126281017"
## [1] "Mean for X  88  is  9.89795969511626"
## [1] "Standard deviations for X88 is 1.89634751172102"
## [1] "Mean for X  89  is  9.79102847356755"
## [1] "Standard deviations for X89 is 2.07407449077112"
## [1] "Mean for X  90  is  10.3441838260588"
## [1] "Standard deviations for X90 is 1.68072302169824"
## [1] "Mean for X  91  is  9.69027901406397"
## [1] "Standard deviations for X91 is 2.02233643959528"
## [1] "Mean for X  92  is  10.2946624181412"
## [1] "Standard deviations for X92 is 1.87416465639807"
## [1] "Mean for X  93  is  10.2927387388247"
## [1] "Standard deviations for X93 is 1.66099141660221"
## [1] "Mean for X  94  is  10.1637570880095"
## [1] "Standard deviations for X94 is 2.03983119379407"
## [1] "Mean for X  95  is  10.3392297699083"
## [1] "Standard deviations for X95 is 2.23479748066963"
## [1] "Mean for X  96  is  10.2445362245418"
```

```
## [1] "Standard deviations for X96 is 1.79568066700882"
## [1] "Mean for X  97  is  9.93485411309578"
## [1] "Standard deviations for X97 is 2.13229267621733"
## [1] "Mean for X  98  is  10.1673828265125"
## [1] "Standard deviations for X98 is 1.75755870718165"
## [1] "Mean for X  99  is  10.0414371825744"
## [1] "Standard deviations for X99 is 2.12059259195975"
## [1] "Mean for X  100  is  9.43835872213286"
## [1] "Standard deviations for X100 is 2.10353824321756"
## [1] "Mean for X  101  is  10.4168021199426"
## [1] "Standard deviations for X101 is 1.97922941417638"

#(b) Use an if statement to identify the variables with a mean greater than
#10.5 and store their names in a vector
new_vector <- character()

for (i in 1:ncol(mydata)) {
  if((means[i]) > 10.5){
    new_vecrtor<- c(new_vector, names(mydata)[i])
  }
}
print(new_vecrtor)

## [1] "X31"
```

3. Use the dataset mydata in question 2.

```
set.seed(123)
#(a) Calculate the mean for each variable.
sapply(mydata,mean)

##         X        X1        X2        X3        X4        X5        X6
X7
## 25.500000 10.692941  9.925052 10.315030 10.964712 10.233889  9.870636
9.805371
##        X8        X9       X10       X11       X12       X13       X14
X15
##  9.642414 10.022282 10.213038 10.307688  9.814616  9.893660  9.749170
10.096851
##       X16       X17       X18       X19       X20       X21       X22
X23
##  9.893779 10.701142 10.261952 10.075299 10.049098  9.717226 10.292880
10.133018
##       X24       X25       X26       X27       X28       X29       X30
X31
##  9.910439  9.739426 10.199722 10.555536 10.363437 10.206064  9.816106
10.662563
##       X32       X33       X34       X35       X36       X37       X38
X39
## 10.334365  9.612551 10.091998 10.139857 10.093518  9.607415 10.003706
10.056218
```

```
##       X40       X41       X42       X43       X44       X45       X46
X47
##  9.846925  9.909246  9.750470  9.754773 10.077017 10.050344 10.188217
9.802783
##       X48       X49       X50       X51       X52       X53       X54
X55
##  9.552618  9.917454 10.169843  9.856027  9.299442 10.045699 10.184347
9.885970
##       X56       X57       X58       X59       X60       X61       X62
X63
##  9.852738  9.838533  9.832220  9.674892  9.911734  9.669543  9.744509
10.245969
##       X64       X65       X66       X67       X68       X69       X70
X71
##  9.934396 10.177343 10.058952  9.857526  9.857805  9.600571  9.696385
10.233076
##       X72       X73       X74       X75       X76       X77       X78
X79
## 10.027390  9.812836 10.191553 10.054493 10.067387  9.817793  9.589308
10.188418
##       X80       X81       X82       X83       X84       X85       X86
X87
## 10.006802 10.301417  9.953892  9.708866 10.345541 10.195723  9.921911
9.897960
##       X88       X89       X90       X91       X92       X93       X94
X95
##  9.791028 10.344184  9.690279 10.294662 10.292739 10.163757 10.339230
10.244536
##       X96       X97       X98       X99      X100
##  9.934854 10.167383 10.041437  9.438359 10.416802

print(is.numeric(unlist(mydata)))

## [1] TRUE

#(b) Calculate the 90% confidence interval for each variable

alpha<-0.1
for(i in 1:ncol(mydata)){
  n <- length(mydata[[i]])
  mean_val <- mean(mydata[[i]])
  sd_val <- sd(mydata[[i]])
  t_val <- qt(1 - alpha / 2, df = n - 1)
  lower_ci <- mean_val - t_val * sd_val / sqrt(n)
  upper_ci <- mean_val + t_val * sd_val / sqrt(n)
  print(paste("90% of confidence interval for variable X",i," is
",round(lower_ci,   2), ", ", round(upper_ci, 2)))
}

## [1] "90% of confidence interval for variable X 1  is  22.04 ,  28.96"
## [1] "90% of confidence interval for variable X 2  is  10.17 ,  11.21"
```

```
## [1] "90% of confidence interval for variable X 3  is  9.4 ,  10.45"
## [1] "90% of confidence interval for variable X 4  is  9.9 ,  10.73"
## [1] "90% of confidence interval for variable X 5  is  10.5 ,  11.43"
## [1] "90% of confidence interval for variable X 6  is  9.76 ,  10.71"
## [1] "90% of confidence interval for variable X 7  is  9.38 ,  10.36"
## [1] "90% of confidence interval for variable X 8  is  9.33 ,  10.28"
## [1] "90% of confidence interval for variable X 9  is  9.22 ,  10.06"
## [1] "90% of confidence interval for variable X 10  is  9.46 ,  10.58"
## [1] "90% of confidence interval for variable X 11  is  9.8 ,  10.63"
## [1] "90% of confidence interval for variable X 12  is  9.86 ,  10.75"
## [1] "90% of confidence interval for variable X 13  is  9.33 ,  10.3"
## [1] "90% of confidence interval for variable X 14  is  9.41 ,  10.38"
## [1] "90% of confidence interval for variable X 15  is  9.19 ,  10.31"
## [1] "90% of confidence interval for variable X 16  is  9.57 ,  10.63"
## [1] "90% of confidence interval for variable X 17  is  9.39 ,  10.4"
## [1] "90% of confidence interval for variable X 18  is  10.22 ,  11.18"
## [1] "90% of confidence interval for variable X 19  is  9.79 ,  10.73"
## [1] "90% of confidence interval for variable X 20  is  9.65 ,  10.5"
## [1] "90% of confidence interval for variable X 21  is  9.63 ,  10.47"
## [1] "90% of confidence interval for variable X 22  is  9.31 ,  10.13"
## [1] "90% of confidence interval for variable X 23  is  9.82 ,  10.76"
## [1] "90% of confidence interval for variable X 24  is  9.66 ,  10.61"
## [1] "90% of confidence interval for variable X 25  is  9.45 ,  10.37"
## [1] "90% of confidence interval for variable X 26  is  9.27 ,  10.21"
## [1] "90% of confidence interval for variable X 27  is  9.81 ,  10.59"
## [1] "90% of confidence interval for variable X 28  is  10.06 ,  11.06"
## [1] "90% of confidence interval for variable X 29  is  9.91 ,  10.82"
## [1] "90% of confidence interval for variable X 30  is  9.75 ,  10.66"
## [1] "90% of confidence interval for variable X 31  is  9.38 ,  10.25"
## [1] "90% of confidence interval for variable X 32  is  10.22 ,  11.11"
## [1] "90% of confidence interval for variable X 33  is  9.87 ,  10.8"
## [1] "90% of confidence interval for variable X 34  is  9.09 ,  10.13"
## [1] "90% of confidence interval for variable X 35  is  9.67 ,  10.52"
## [1] "90% of confidence interval for variable X 36  is  9.58 ,  10.7"
## [1] "90% of confidence interval for variable X 37  is  9.59 ,  10.6"
## [1] "90% of confidence interval for variable X 38  is  9.09 ,  10.13"
## [1] "90% of confidence interval for variable X 39  is  9.56 ,  10.45"
## [1] "90% of confidence interval for variable X 40  is  9.51 ,  10.6"
## [1] "90% of confidence interval for variable X 41  is  9.39 ,  10.3"
## [1] "90% of confidence interval for variable X 42  is  9.41 ,  10.41"
## [1] "90% of confidence interval for variable X 43  is  9.24 ,  10.26"
## [1] "90% of confidence interval for variable X 44  is  9.31 ,  10.2"
## [1] "90% of confidence interval for variable X 45  is  9.6 ,  10.55"
## [1] "90% of confidence interval for variable X 46  is  9.7 ,  10.4"
## [1] "90% of confidence interval for variable X 47  is  9.66 ,  10.72"
## [1] "90% of confidence interval for variable X 48  is  9.32 ,  10.28"
## [1] "90% of confidence interval for variable X 49  is  9.08 ,  10.03"
## [1] "90% of confidence interval for variable X 50  is  9.41 ,  10.43"
## [1] "90% of confidence interval for variable X 51  is  9.7 ,  10.64"
## [1] "90% of confidence interval for variable X 52  is  9.43 ,  10.28"
```

```
## [1] "90% of confidence interval for variable X 53  is  8.81 ,  9.79"
## [1] "90% of confidence interval for variable X 54  is  9.6 ,  10.49"
## [1] "90% of confidence interval for variable X 55  is  9.7 ,  10.67"
## [1] "90% of confidence interval for variable X 56  is  9.48 ,  10.29"
## [1] "90% of confidence interval for variable X 57  is  9.33 ,  10.37"
## [1] "90% of confidence interval for variable X 58  is  9.35 ,  10.33"
## [1] "90% of confidence interval for variable X 59  is  9.31 ,  10.36"
## [1] "90% of confidence interval for variable X 60  is  9.26 ,  10.09"
## [1] "90% of confidence interval for variable X 61  is  9.44 ,  10.39"
## [1] "90% of confidence interval for variable X 62  is  9.21 ,  10.13"
## [1] "90% of confidence interval for variable X 63  is  9.31 ,  10.18"
## [1] "90% of confidence interval for variable X 64  is  9.77 ,  10.73"
## [1] "90% of confidence interval for variable X 65  is  9.53 ,  10.34"
## [1] "90% of confidence interval for variable X 66  is  9.76 ,  10.6"
## [1] "90% of confidence interval for variable X 67  is  9.52 ,  10.6"
## [1] "90% of confidence interval for variable X 68  is  9.34 ,  10.38"
## [1] "90% of confidence interval for variable X 69  is  9.36 ,  10.35"
## [1] "90% of confidence interval for variable X 70  is  9.17 ,  10.03"
## [1] "90% of confidence interval for variable X 71  is  9.23 ,  10.16"
## [1] "90% of confidence interval for variable X 72  is  9.77 ,  10.7"
## [1] "90% of confidence interval for variable X 73  is  9.5 ,  10.55"
## [1] "90% of confidence interval for variable X 74  is  9.32 ,  10.31"
## [1] "90% of confidence interval for variable X 75  is  9.68 ,  10.7"
## [1] "90% of confidence interval for variable X 76  is  9.63 ,  10.48"
## [1] "90% of confidence interval for variable X 77  is  9.58 ,  10.55"
## [1] "90% of confidence interval for variable X 78  is  9.35 ,  10.28"
## [1] "90% of confidence interval for variable X 79  is  9.08 ,  10.1"
## [1] "90% of confidence interval for variable X 80  is  9.77 ,  10.61"
## [1] "90% of confidence interval for variable X 81  is  9.52 ,  10.5"
## [1] "90% of confidence interval for variable X 82  is  9.79 ,  10.82"
## [1] "90% of confidence interval for variable X 83  is  9.39 ,  10.52"
## [1] "90% of confidence interval for variable X 84  is  9.2 ,  10.22"
## [1] "90% of confidence interval for variable X 85  is  9.85 ,  10.84"
## [1] "90% of confidence interval for variable X 86  is  9.74 ,  10.65"
## [1] "90% of confidence interval for variable X 87  is  9.44 ,  10.4"
## [1] "90% of confidence interval for variable X 88  is  9.45 ,  10.35"
## [1] "90% of confidence interval for variable X 89  is  9.3 ,  10.28"
## [1] "90% of confidence interval for variable X 90  is  9.95 ,  10.74"
## [1] "90% of confidence interval for variable X 91  is  9.21 ,  10.17"
## [1] "90% of confidence interval for variable X 92  is  9.85 ,  10.74"
## [1] "90% of confidence interval for variable X 93  is  9.9 ,  10.69"
## [1] "90% of confidence interval for variable X 94  is  9.68 ,  10.65"
## [1] "90% of confidence interval for variable X 95  is  9.81 ,  10.87"
## [1] "90% of confidence interval for variable X 96  is  9.82 ,  10.67"
## [1] "90% of confidence interval for variable X 97  is  9.43 ,  10.44"
## [1] "90% of confidence interval for variable X 98  is  9.75 ,  10.58"
## [1] "90% of confidence interval for variable X 99  is  9.54 ,  10.54"
## [1] "90% of confidence interval for variable X 100  is  8.94 ,  9.94"
## [1] "90% of confidence interval for variable X 101  is  9.95 ,  10.89"
```

```r
#(c) Print the results using a matrix with three columns named
#("Mean", "Lower_CI", and "Upper_CI").
result_matrix <- matrix(NA, nrow = ncol(mydata), ncol = 3,
                        dimnames = list(names(mydata), c("Mean", "Lower_CI",
"Upper_CI")))

for (i in 1:ncol(mydata)) {
  mean_val <- mean(mydata[[i]])
  n <- length(mydata[[i]])
  stder <- sd(mydata[[i]]) / sqrt(n)
  alpha <- 0.1
  t_val <- qt(1 - alpha/2 , df = n - 1)
  lower_ci <- mean_val - t_val * stder/sqrt(n)
  upper_ci <- mean_val + t_val * stder/sqrt(n)
  result_matrix[i, 1] <- mean_val
  result_matrix[i, 2] <- lower_ci
  result_matrix[i, 3] <- upper_ci
}

print(result_matrix)
```

```
##           Mean  Lower_CI  Upper_CI
## X    25.500000 25.011206 25.988794
## X1   10.692941 10.619246 10.766636
## X2    9.925052  9.851057  9.999046
## X3   10.315030 10.255898 10.374163
## X4   10.964712 10.898726 11.030699
## X5   10.233889 10.166886 10.300891
## X6    9.870636  9.800809  9.940463
## X7    9.805371  9.737833  9.872909
## X8    9.642414  9.583238  9.701590
## X9   10.022282  9.942730 10.101833
## X10  10.213038 10.154101 10.271975
## X11  10.307688 10.244896 10.370480
## X12   9.814616  9.746261  9.882970
## X13   9.893660  9.824600  9.962719
## X14   9.749170  9.670433  9.827906
## X15  10.096851 10.021668 10.172034
## X16   9.893779  9.822682  9.964877
## X17  10.701142 10.633462 10.768822
## X18  10.261952 10.195289 10.328615
## X19  10.075299 10.015120 10.135477
## X20  10.049098  9.990217 10.107979
## X21   9.717226  9.659277  9.775175
## X22  10.292880 10.226332 10.359427
## X23  10.133018 10.065913 10.200123
## X24   9.910439  9.845053  9.975826
## X25   9.739426  9.673301  9.805551
## X26  10.199722 10.144814 10.254630
## X27  10.555536 10.484879 10.626193
```

```
## X28  10.363437 10.299265 10.427610
## X29  10.206064 10.141320 10.270808
## X30   9.816106  9.754990  9.877222
## X31  10.662563 10.599589 10.725536
## X32  10.334365 10.269054 10.399676
## X33   9.612551  9.539049  9.686053
## X34  10.091998 10.031701 10.152296
## X35  10.139857 10.060225 10.219489
## X36  10.093518 10.022476 10.164559
## X37   9.607415  9.533693  9.681137
## X38  10.003706  9.941053 10.066359
## X39  10.056218  9.979434 10.133003
## X40   9.846925  9.782206  9.911644
## X41   9.909246  9.838940  9.979553
## X42   9.750470  9.677813  9.823128
## X43   9.754773  9.691302  9.818245
## X44  10.077017 10.009444 10.144591
## X45  10.050344 10.000974 10.099714
## X46  10.188217 10.112904 10.263530
## X47   9.802783  9.734670  9.870896
## X48   9.552618  9.485324  9.619912
## X49   9.917454  9.845548  9.989360
## X50  10.169843 10.103982 10.235703
## X51   9.856027  9.796372  9.915682
## X52   9.299442  9.230038  9.368846
## X53  10.045699  9.982847 10.108550
## X54  10.184347 10.116067 10.252626
## X55   9.885970  9.828380  9.943561
## X56   9.852738  9.779315  9.926161
## X57   9.838533  9.769245  9.907821
## X58   9.832220  9.758035  9.906405
## X59   9.674892  9.615962  9.733822
## X60   9.911734  9.844685  9.978783
## X61   9.669543  9.603999  9.735087
## X62   9.744509  9.682592  9.806427
## X63  10.245969 10.178037 10.313901
## X64   9.934396  9.877444  9.991347
## X65  10.177343 10.118152 10.236535
## X66  10.058952  9.982325 10.135578
## X67   9.857526  9.783671  9.931381
## X68   9.857805  9.787779  9.927832
## X69   9.600571  9.539648  9.661493
## X70   9.696385  9.630155  9.762614
## X71  10.233076 10.166985 10.299168
## X72  10.027390  9.953046 10.101734
## X73   9.812836  9.742619  9.883053
## X74  10.191553 10.119398 10.263708
## X75  10.054493  9.994409 10.114577
## X76  10.067387  9.998568 10.136206
## X77   9.817793  9.752289  9.883298
```

```
## X78   9.589308  9.517409  9.661206
## X79  10.188418 10.128915 10.247921
## X80  10.006802  9.937731 10.075872
## X81  10.301417 10.228406 10.374429
## X82   9.953892  9.873803 10.033982
## X83   9.708866  9.636945  9.780788
## X84  10.345541 10.275348 10.415735
## X85  10.195723 10.131167 10.260279
## X86   9.921911  9.854154  9.989669
## X87   9.897960  9.834373  9.961546
## X88   9.791028  9.721483  9.860574
## X89  10.344184 10.287827 10.400540
## X90   9.690279  9.622468  9.758090
## X91  10.294662 10.231820 10.357505
## X92  10.292739 10.237044 10.348433
## X93  10.163757 10.095359 10.232155
## X94  10.339230 10.264295 10.414165
## X95  10.244536 10.184325 10.304747
## X96   9.934854  9.863356 10.006352
## X97  10.167383 10.108450 10.226316
## X98  10.041437  9.970332 10.112543
## X99   9.438359  9.367825  9.508893
## X100 10.416802 10.350437 10.483168
```

4. Consider the built-in data set UCBAdmissions in R.

```
set.seed(123)
data(UCBAdmissions)

#(a) If we are interested in the proportion of people that apply to Berkeley
#University and get accepted, what is the population of interest and
#what is the parameter of interest?

#The population of interest is all peoole who apply to Berkeley University
#The parameter of interest is the proportion people who apply to Berkeley
#University and got accepted.

#(b) Using the command ? UCBAdmissions, determine the variables in the
dataset
#and describe what kind of variables they are.
#?UCBAdmissions

#There are 3 variables: Admit, Gender, and Dept.
#First of all, Admit is a categorical variable
#which describes if the applicants got admitted or not. This variable has two
levels: Admitted and Rejected
#Second of all, Gender is a categorical variable which describes genders of
applicants. There are two levels:Male and Female
#Lastly, Dept is also categorical variable. Dept shows different departments
which applicants applied to. This has 6 levels:A,B,C,D,E,and F.
```

```r
#(c) Create a variable in R called totalAdmissions which contains the total
number of students who were admitted to the university (across all genders
and departments).

totalAdmissions <- sum(UCBAdmissions["Admitted", , ])
totalAdmissions
```

```
## [1] 1755
```

```r
#(d) Create a variable in R called totalRejections which contains the total
#number of students who were rejected to the university (across all genders a
#                                                        nd departments).
totalRejections <- sum(UCBAdmissions["Rejected",,])
totalRejections
```

```
## [1] 2771
```

```r
#(e) Create a variable in R called totalApplicants which contains the total n
#umber of students who applied to the university in our sample.
totalApplicants <- sum(UCBAdmissions)
totalApplicants
```

```
## [1] 4526
```

```r
#(f) What is the observed value of the statistic we should use to estimate
the population parameter of interest?
pop_prop<- totalAdmissions/totalApplicants
pop_prop
```

```
## [1] 0.3877596
```

```r
#(g) What is the estimated standard error for ^p?
est_stdError<- sqrt(pop_prop*(1-pop_prop))/totalApplicants
est_stdError
```

```
## [1] 0.0001076534
```

```r
#(h) What is the critical value for a 92% con_dence interval for p?
#1-a = 0.92
#a = 0.08
#a/2 = 0.04 = 96%
cri_p<-qnorm(0.96)
cri_p
```

```
## [1] 1.750686
```

```r
#(i) What is the margin of error for our estimate?
est_stdError*cri_p
```

```
## [1] 0.0001884673
```

```r
#(j) Compare that to result of the approximate margin of error formula we
#learned earlier in the course.
```

```r
app_me<- cri_p*sqrt((pop_prop*(pop_prop))/totalApplicants)
app_me
```

```
## [1] 0.01009052
```

```r
#(k) Determine a 92% con_dence interval for the true value of the population
proportion
upper<- 0.3877596 + 0.0001884673
lower<-0.3877596 - 0.0001884673
print(paste("The 92% CI is betweeen",lower,"and", upper))
```

```
## [1] "The 92% CI is betweeen 0.3875711327 and 0.3879480673"
```

#(a) If we are interested in the proportion of people that apply to Berkeley #University and get accepted, what is the population of interest and #what is the parameter of interest?

#The population of interest is all peoole who apply to Berkeley University #The parameter of interest is the proportion people who apply to Berkeley #University and got accepted.

#(b) Using the command ? UCBAdmissions, determine the variables in the dataset #and describe what kind of variables they are. #?UCBAdmissions

#There are 3 variables: Admit, Gender, and Dept. #First of all, Admit is a categorical variable #which describes if the applicants got admitted or not. This variable has two levels: Admitted and Rejected #Second of all, Gender is a categorical variable which describes genders of applicants. There are two levels:Male and Female #Lastly, Dept is also categorical variable. Dept shows different departments which applicants applied to. This has 6 levels:A,B,C,D,E,and F.