SCS 2111 | Statistical Methods using R
Take home Assignment

Y.D.K.Medis

16000897

2016/CS/089

a)

```
> data(package = "boot","channing")
> channing
       sex entry exit time cens
1     Male   782  909  127    1
2     Male  1020 1128  108    1
3     Male   856  969  113    1
4     Male   915  957   42    1
5     Male   863  983  120    1
6     Male   906 1012  106    1
7     Male   955 1055  100    1
8     Male   943 1025   82    1
9     Male   943 1043  100    1
10    Male   837  945  108    1
11    Male   966 1009   43    1
12    Male   936  971   35    1
13    Male   919 1033  114    1
14    Male   852  869   17    1
15    Male  1073 1139   66    1
16    Male   925 1036  111    1
```

.

.

.

```
445 Female   930  936    6    1
446 Female   943  994   51    1
447 Female  1024 1063   39    1
448 Female   802  821   19    0
449 Female   811  819    8    0
450 Female   927 1001   74    1
451 Female   967  975    8    1
452 Female   943  982   39    1
453 Female   840  905   65    0
454 Female   979 1040   61    1
455 Female   921  926    5    0
456 Female   986 1030   44    1
457 Female  1039 1132   93    1
458 Female   968  990   22    1
459 Female   955  990   35    1
460 Female   837  911   74    1
461 Female   861  915   54    1
462 Female   967  983   16    1
```

b)

```
> str(channing)
'data.frame':   462 obs. of  5 variables:
 $ sex  : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 2 ...
 $ entry: num  782 1020 856 915 863 906 955 943 943 837 ...
 $ exit : num  909 1128 969 957 983 ...
 $ time : num  127 108 113 42 120 106 100 82 100 108 ...
 $ cens : num  1 1 1 1 1 1 1 1 1 1 ...
> 
```

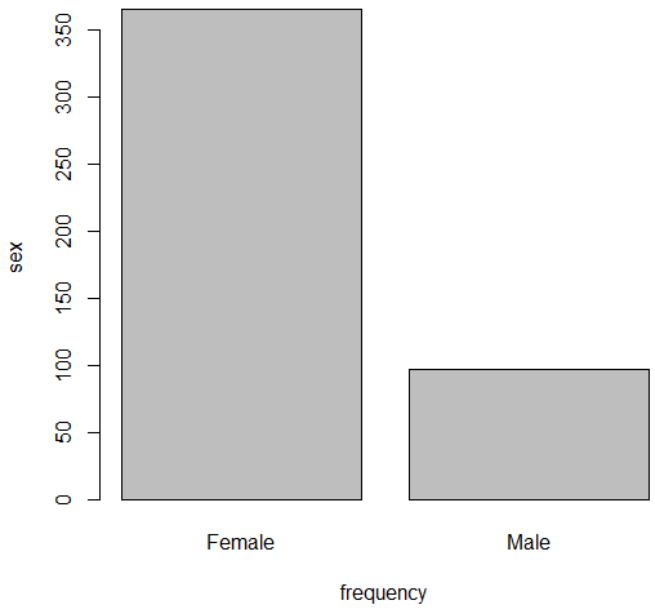
c)

quality variables = sex, entry, exit, time, cens

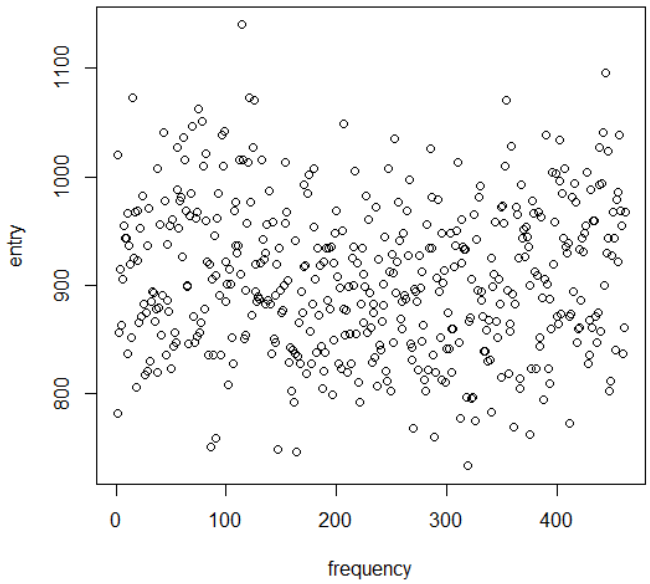
d)

```
> summary(channing$entry)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  733.0   854.0   900.5   905.9   956.0  1140.0
> summary(channing$exit)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    777     939     990     986    1031    1207
> summary(channing$time)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00   35.00   82.00   80.34  137.00  137.00
> summary(channing$cesn)
Length  Class   Mode
     0   NULL   NULL
> summary(channing$sex)
Female   Male
   365     97
>
```

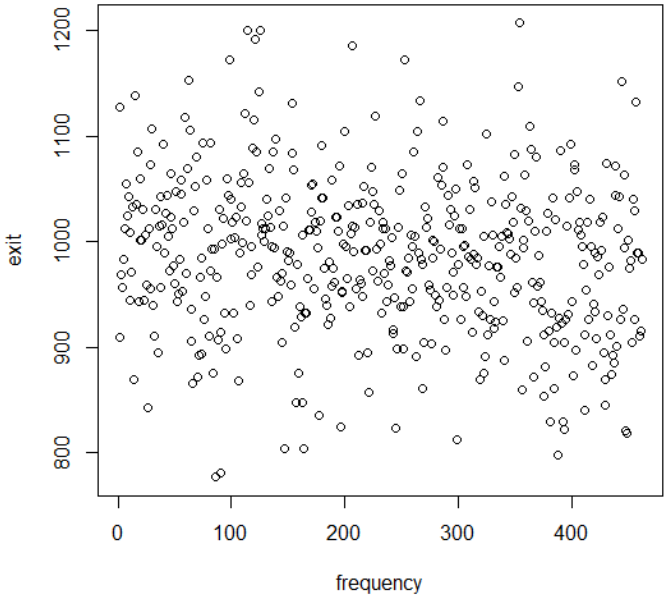
```
> plot(sex,xlab="frequency",ylab="sex")
```



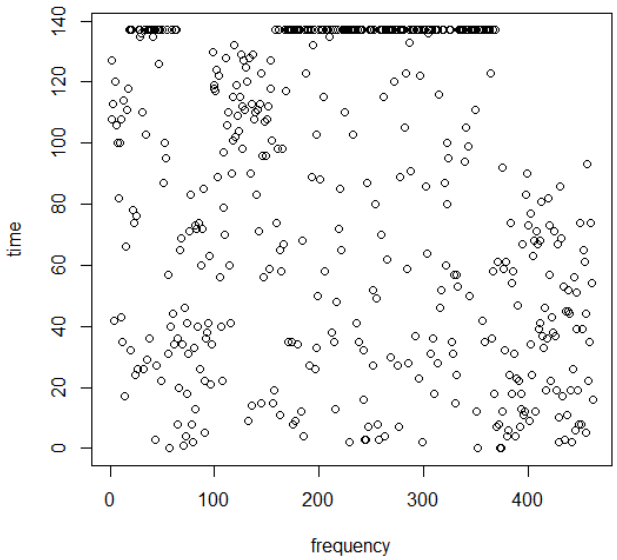
```
> plot(entry,xlab="frequency",ylab="entry")
```



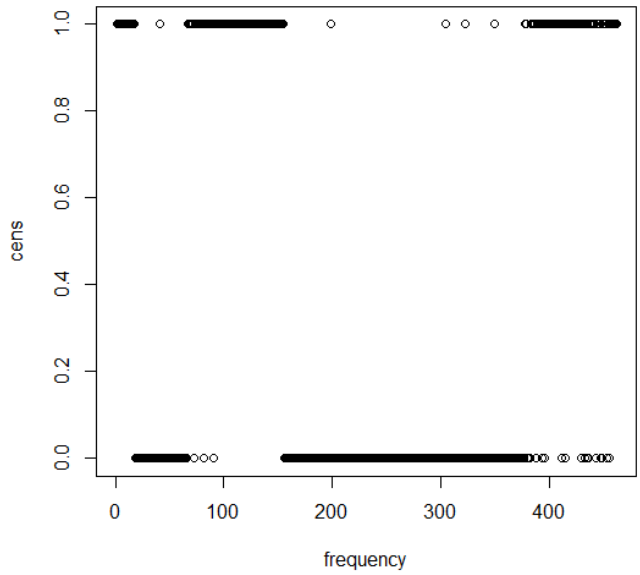

```
> plot(exit,xlab="frequency",ylab="exit")
```



```
> plot(time,xlab="frequency",ylab="time")
```



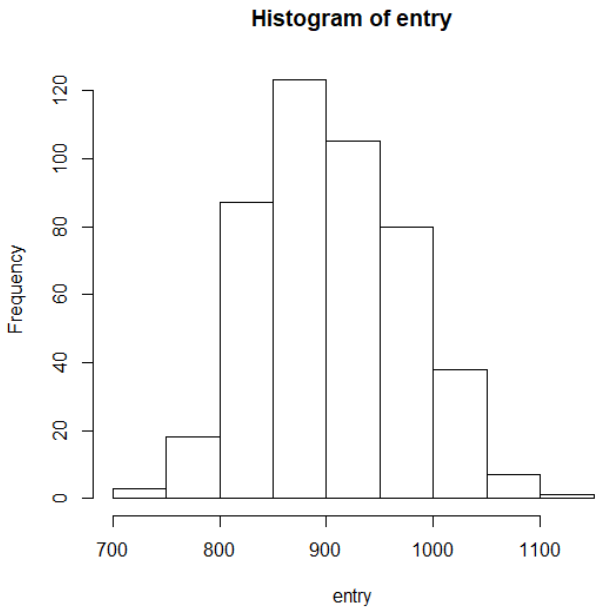

```
> plot(cesn,xlab="frequency",ylab="cens")
```



e)

```
> hist(entry)
```

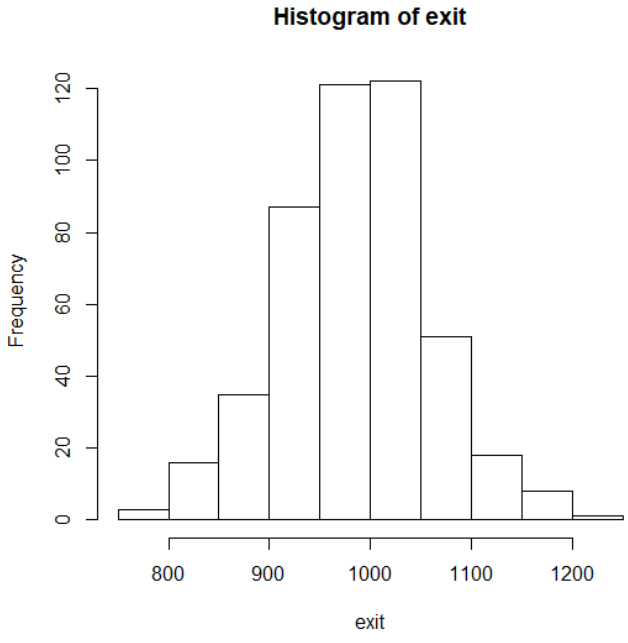
negative Skewness




```
> hist(exit)
```

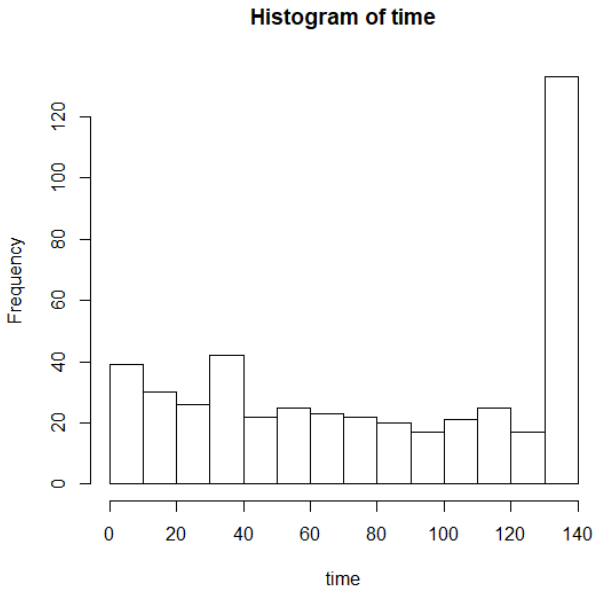
symmetrically bell shaped

Normal distribution



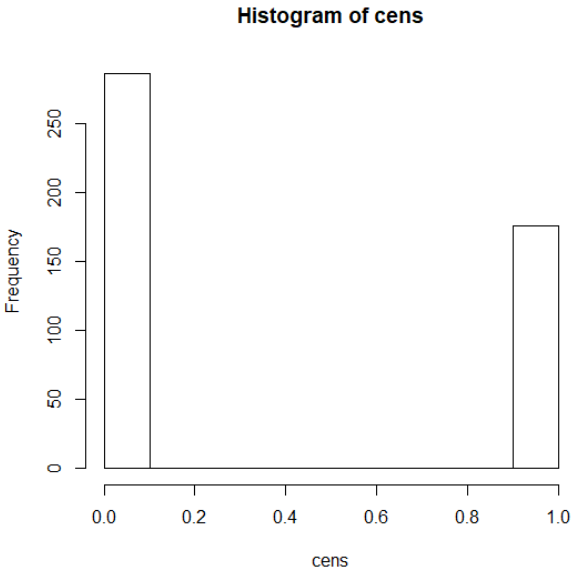
```
> hist(time)
```

Other distribution




```
> hist(cens)
```

Other distribution



f)

```
> R<-set.seed(1600897)
```

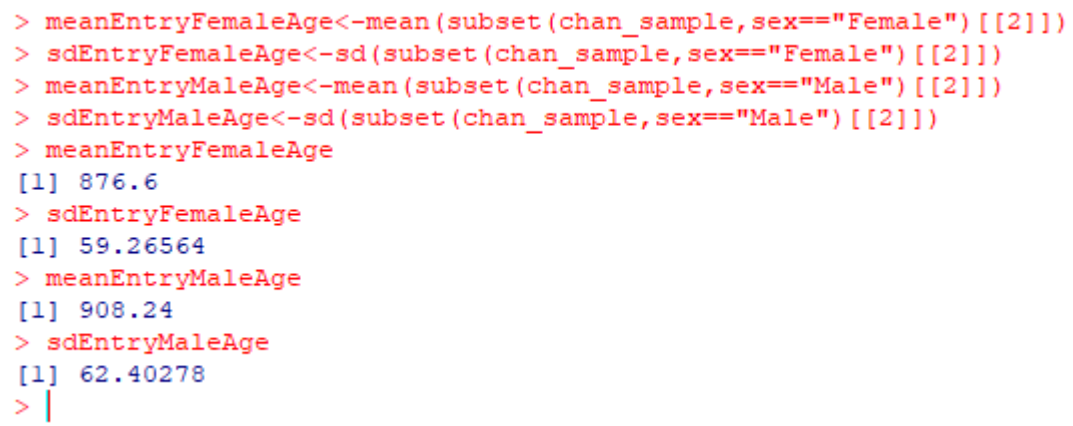
g)

```
> df<-split(channing,sex)
> chan_female<-df[[1]][sample(nrow(df[[1]]),25),]
> chan_male<-df[[2]][sample(nrow(df[[2]]),25),]
> chan_sample<-rbind(chan_male,chan_female)
```

```
> df<-split(channing,sex)
> chan_female<-df[[1]][sample(nrow(df[[1]]),25),]
> chan_male<-df[[2]][sample(nrow(df[[2]]),25),]
> chan_sample<-rbind(chan_male,chan_female)
> chan_sample
       sex entry exit time cens
11    Male   966 1009   43    1
97    Male  1010 1044   34    1
42    Male   890 1027  137    0
84    Male   836  876   40    1
48    Male   876 1013  137    0
70    Male   871  872    1    1
66    Male   846  866   20    0
28    Male   821  956  135    0
87    Male   906  966   60    1
6     Male   906 1012  106    1
45    Male   836  973  137    0
77    Male   865  948   83    1
89    Male   946 1031   85    1
4     Male   915  957   42    1
31    Male   830  940  110    0
17    Male   967 1085  118    1
41    Male   854  989  135    1
51    Male   960 1047   87    0
5     Male   863  983  120    1
15    Male  1073 1139   66    1
33    Male   894 1031  137    0
24    Male   982 1006   24    0
12    Male   936  971   35    1
83    Male   921  993   72    1
29    Male   936 1073  137    0
349 Female   815  952  137    0
```

h)

```
> meanEntryFemaleAge<-mean(subset(chan_sample,sex=="Female")[[2]])
> sdEntryFemaleAge<-sd(subset(chan_sample,sex=="Female")[[2]])
> meanEntryMaleAge<-mean(subset(chan_sample,sex=="Male")[[2]])
> sdEntryMaleAge<-sd(subset(chan_sample,sex=="Male")[[2]])
> meanEntryFemaleAge
[1] 876.6
> sdEntryFemaleAge
[1] 59.26564
> meanEntryMaleAge
[1] 908.24
> sdEntryMaleAge
[1] 62.40278
>
```



i)

```
> mf<-meanEntryFemaleAge
> mm<-meanEntryMaleAge
> sf<-sdEntryFemaleAge
> sm<-sdEntryMaleAge
>
>
> n<-25
>
> sp<- sqrt(((n-1)*sf^2+(n-1*sm^2))/(2*n-2))
> tstVal<-(mf-mm)/(sp*sqrt(2/n))
> tstVal
[1] -2.732789
> popVal<-2*pt(tstVal,df=2*n-2)
> popVal
[1] 0.008764895
>
```