9. Group fr distn and descriptive statistics

Grouped frequency ditribution

```
r1=c(170, 151, 154, 160, 158, 154, 171, 156, 160, 157, 148, 165, 158, 159, 155, 151)
r2=c(152, 161, 156, 164, 156, 163, 174, 153, 170, 149, 166, 154, 166, 160, 160, 161)
r3=c(154, 163, 164, 160, 148, 162, 167, 165, 158, 158, 176, 160, 157)
height=c(r1,r2,r3)
#height=scan() also use for data entry
#class= 147.5-152.5,152.5-157.5,157.5-162.5,162.5-167.5,167.5-172.5,172.5-177.5
n=length(height)
rg=max(height)-min(height) #Range
k=6 #number of classes
cls=rg/k #4.67
w=5 #class width
i=1:6
x=c(150+w*(i-1)) #mid point
x
```

```
[1] 150 155 160 165 170 175
```

```
f=1:6
for (i in 1:6) {
  f[i]=length(height[height>=(x[i]-w/2) & height<=(x[i]+w/2)])
}
f
```

```
[1]  6 11 14  9  3  2
```

```
n=sum(f)
n
```

```
[1] 45
```

```
ht.cls=c("147.5-152.5","152.5-157.5","157.5-162.5","162.5-167.5","167.5-172.5","172.5-177.5")
fr.dist=data.frame(ht.cls,f)
fr.dist
```

```
       ht.cls  f
1 147.5-152.5  6
2 152.5-157.5 11
3 157.5-162.5 14
4 162.5-167.5  9
5 167.5-172.5  3
6 172.5-177.5  2
```

Second method for finding frequency

```
brk=seq(147.5,177.5,5)
temp=cut(height,brk,Right=F)
temp
```

```
 [1] (168,172] (148,152] (152,158] (158,162] (158,162] (152,158] (168,172]
 [8] (152,158] (158,162] (152,158] (148,152] (162,168] (158,162] (158,162]
[15] (152,158] (148,152] (148,152] (158,162] (152,158] (162,168] (152,158]
[22] (162,168] (172,178] (152,158] (168,172] (148,152] (162,168] (152,158]
[29] (162,168] (158,162] (158,162] (158,162] (152,158] (162,168] (162,168]
[36] (158,162] (148,152] (158,162] (162,168] (162,168] (158,162] (158,162]
[43] (172,178] (158,162] (152,158]
Levels: (148,152] (152,158] (158,162] (162,168] (168,172] (172,178]
```

```
temp1=table(temp)
temp1
```

```
temp
(148,152] (152,158] (158,162] (162,168] (168,172] (172,178]
        6        11        14         9         3         2
```

```
height=data.frame(fr.dist,x)
height
```

```
       ht.cls  f   x
1 147.5-152.5  6 150
2 152.5-157.5 11 155
3 157.5-162.5 14 160
4 162.5-167.5  9 165
5 167.5-172.5  3 170
6 172.5-177.5  2 175
```

Finding mean of the given data

```
attach(height)
```

```
The following objects are masked _by_ .GlobalEnv:

    f, ht.cls, x
```

```
mn=sum(f*x)/n
mn
```

```
[1] 159.7778
```

Finding median of the given data

```
cf=cumsum(f)
cf
```

```
[1]  6 17 31 40 43 45
```

```
m1=min(which(cf>=n/2))
m1
```

```
[1] 3
```

```
x
```

```
[1] 150 155 160 165 170 175
```

```
lower_class=x[m1]-w/2
lower_class
```

```
[1] 157.5
```

```
f1<-f[m1]
f1
```

```
[1] 14
```

```
cum=cf[m1-1]
cum
```

```
[1] 17
```

```
median=lower_class+(((n/2)-cum)/f1)*w
median
```

```
[1] 159.4643
```

Finding mode of the given data

```
m=which(f==max(f))
m
```

```
[1] 3
```

```
fm=f[m]
fm
```

```
[1] 14
```

```
f1=f[m-1]
f1
```

```
[1] 11
```

```
f2=f[m+1]
f2
```

```
[1] 9
```

```
lower_class=x[m]-w/2
lower_class
```

```
[1] 157.5
```

```
mode=lower_class+((fm-f1)/(2*fm-f1-f2))*w
mode
```

```
[1] 159.375
```

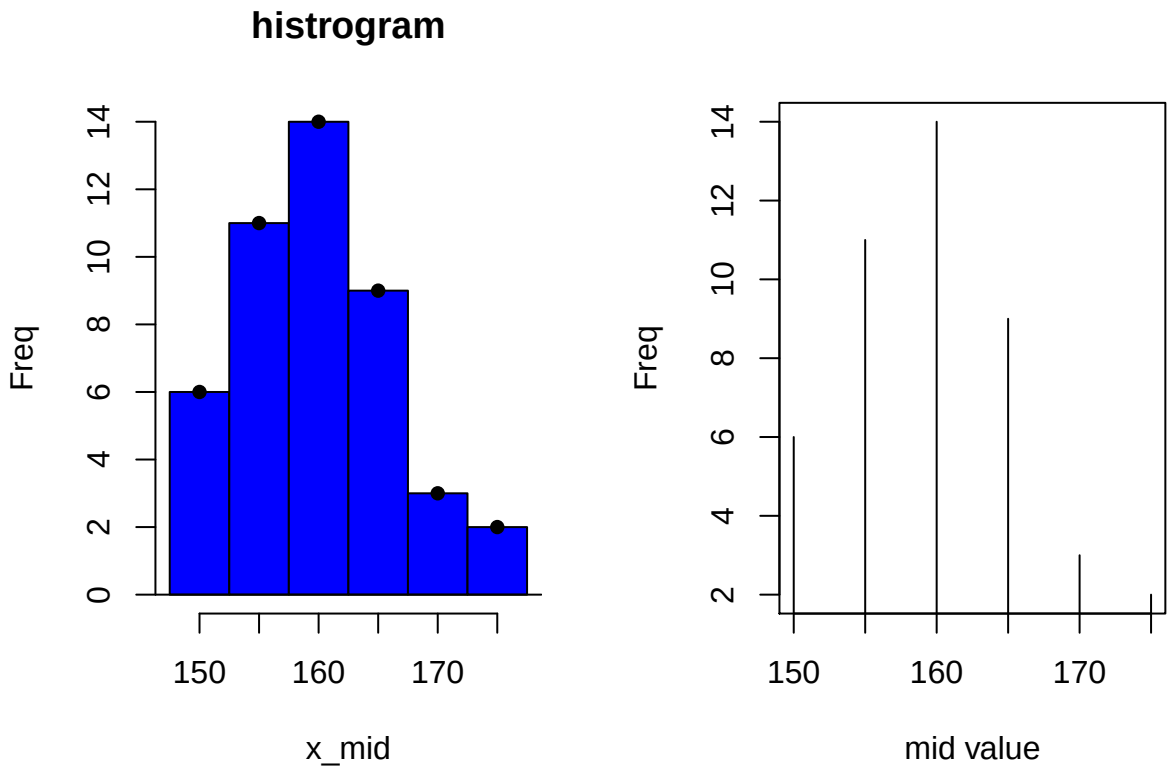
Ploting graph of the given data

```
par(mfrow=c(1,2))
pol=rep(x,f)
pol
```

```
 [1] 150 150 150 150 150 150 155 155 155 155 155 155 155 155 155 155 155 160 160
[20] 160 160 160 160 160 160 160 160 160 160 160 160 165 165 165 165 165 165 165
[39] 165 165 170 170 170 175 175
```

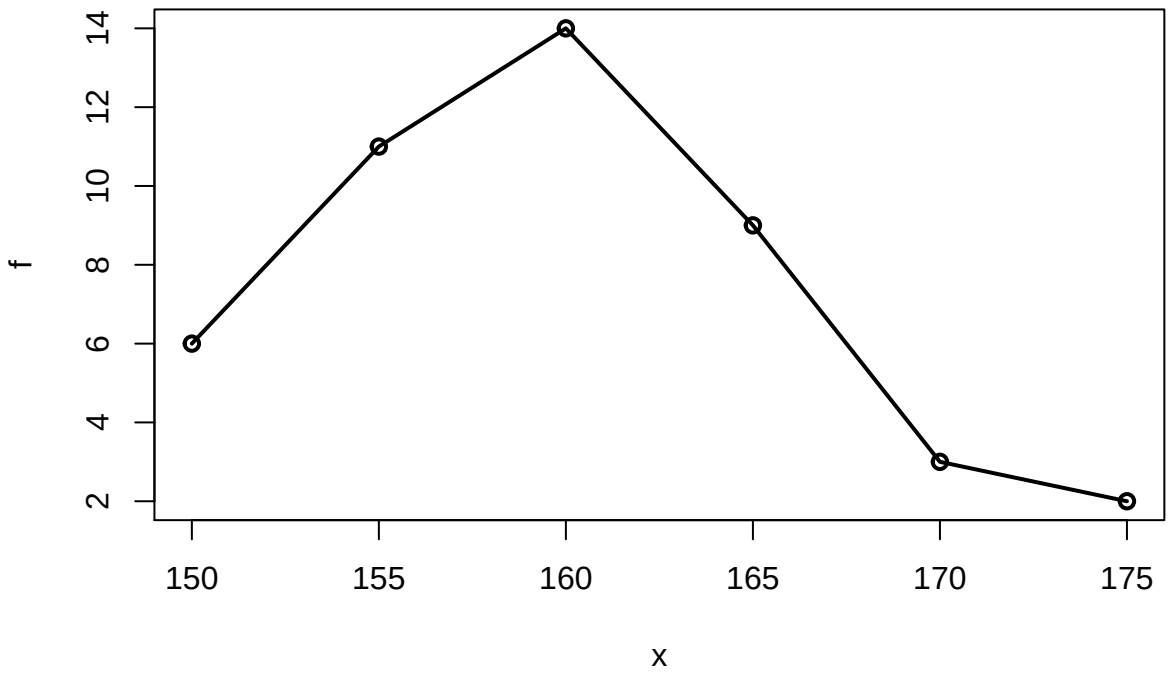
```
brk=seq(147.5,177.5,5)
hist(pol,breaks=brk,col="blue",main="histrogram",xlab="x_mid",ylab="Freq",cex=7)
```

```
abline(h=0)
plot(x,f,"h",xlab="mid value",ylab="Freq",points(x,f, pch=16))
```



Frequency Polygon of the given data

```
plot(x,f,type="o",lwd=2)
```




```
par(mfrow=c(1,1))
```

More than and less than ogive curves

```
f
```

```
## [1]  6 11 14  9  3  2
```

```
lc=c(0,cumsum(f))
lc
```

```
## [1]  0  6 17 31 40 43 45
```

```
uc=1:6
for(i in 1:6)
{uc[i]<-sum(f[6:i])}
uc=c(uc,0)
uc
```

```
## [1] 45 39 28 14  5  2  0
```

```
lbx=seq(147.5,177.5,5)
lbx
```

```
## [1] 147.5 152.5 157.5 162.5 167.5 172.5 177.5
```

```
ubx=seq(147.5,177.5,5)
ubx
```

```
## [1] 147.5 152.5 157.5 162.5 167.5 172.5 177.5
```

```
plot(ubx,lc,type="l",xlim=c(147.5,172.5),xlab="class interval",ylab="cumulative frequency",main="",lwd=4
lines(lbx,uc,lty=2,lwd=2)
```