

Guia 22 Clase

Agustin Muñoz González

6/7/2020

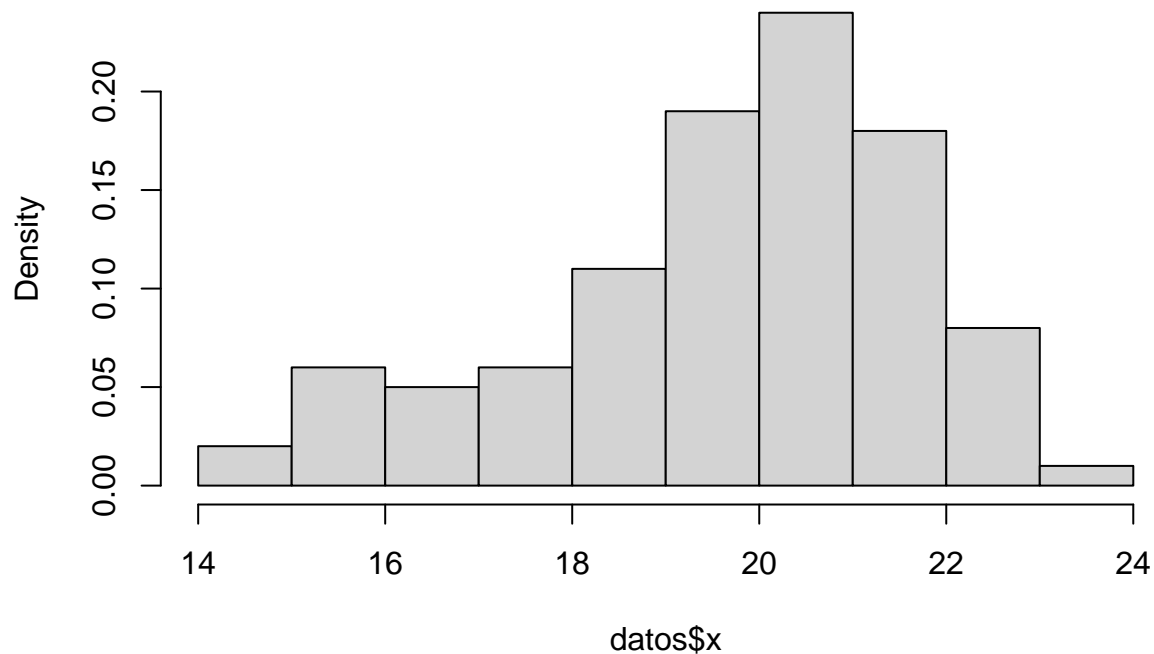
Preparamos el entorno.

```
rm(list=ls())  
library(ggplot2)  
library(tidyr)  
library(gganimate)
```

1.

```
datos=read.csv('datos.csv',header=TRUE)  
hist(datos$x,freq=F)
```

Histogram of datos\$x

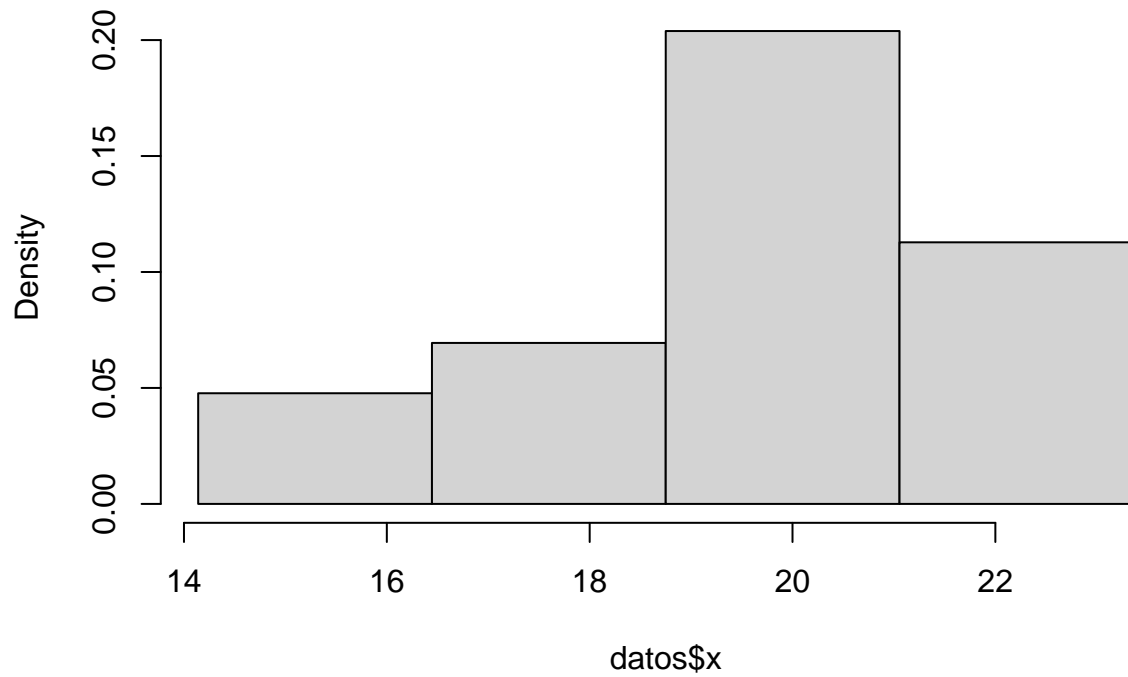


2.

i)

```
hist(datos$x,freq=F,  
      breaks=seq(min(datos$x),max(datos$x),length.out=5))
```

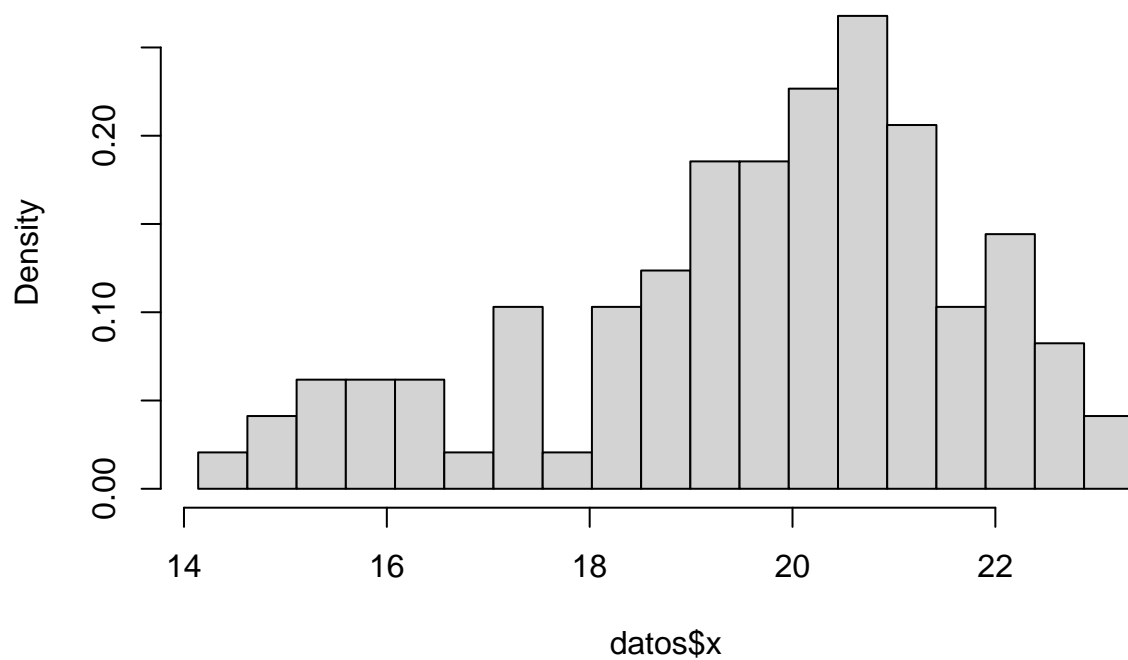
Histogram of datos\$x



ii)

```
hist(datos$x,freq=F,  
      breaks=seq(min(datos$x),max(datos$x),length.out=20))
```

Histogram of datos\$x

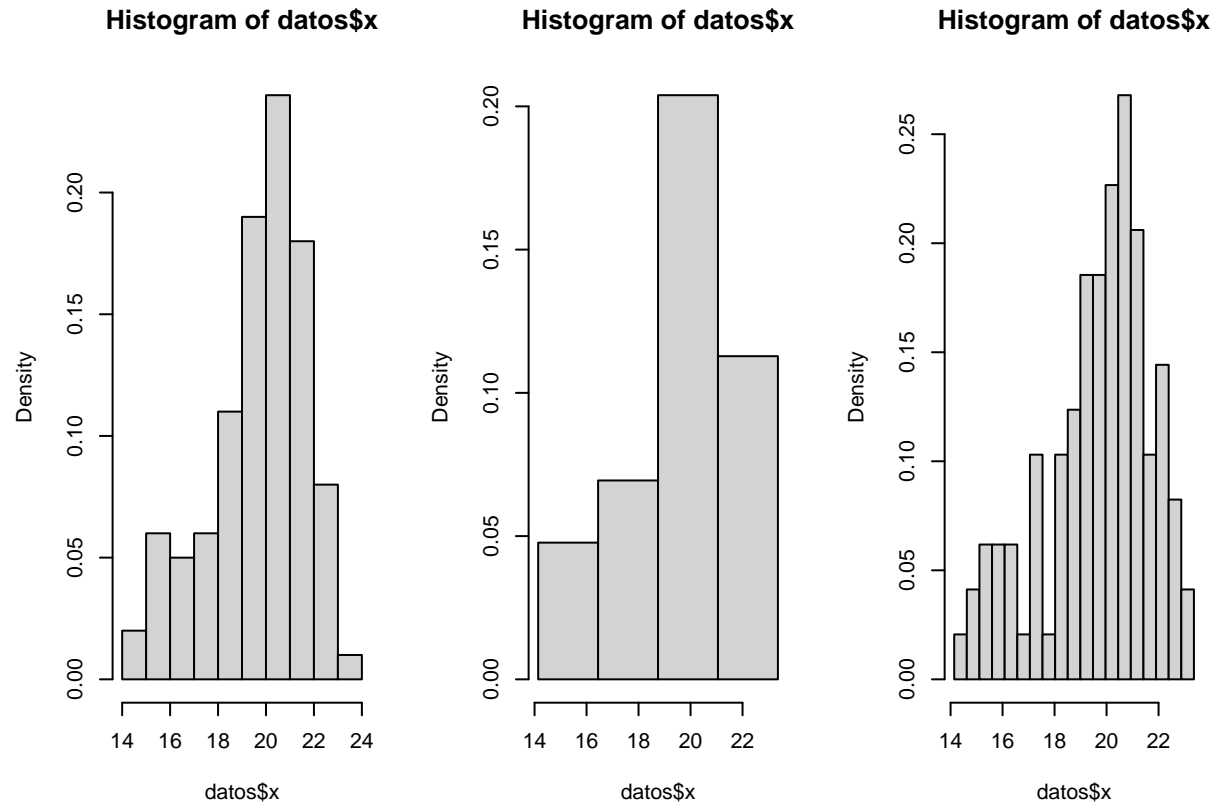


Ploteamos los 3 juntos

```

par(mfrow=c(1,3))
hist(datos$x,freq=F)
hist(datos$x,freq=F,
      breaks=seq(min(datos$x),max(datos$x),length.out=5))
hist(datos$x,freq=F,
      breaks=seq(min(datos$x),max(datos$x),length.out=20))

```



```

par(mfrow=c(1,1))

```

Claramente observamos que sí hay un efecto y es que al tener barras mas finas tenemos un histograma mas preciso.

3.

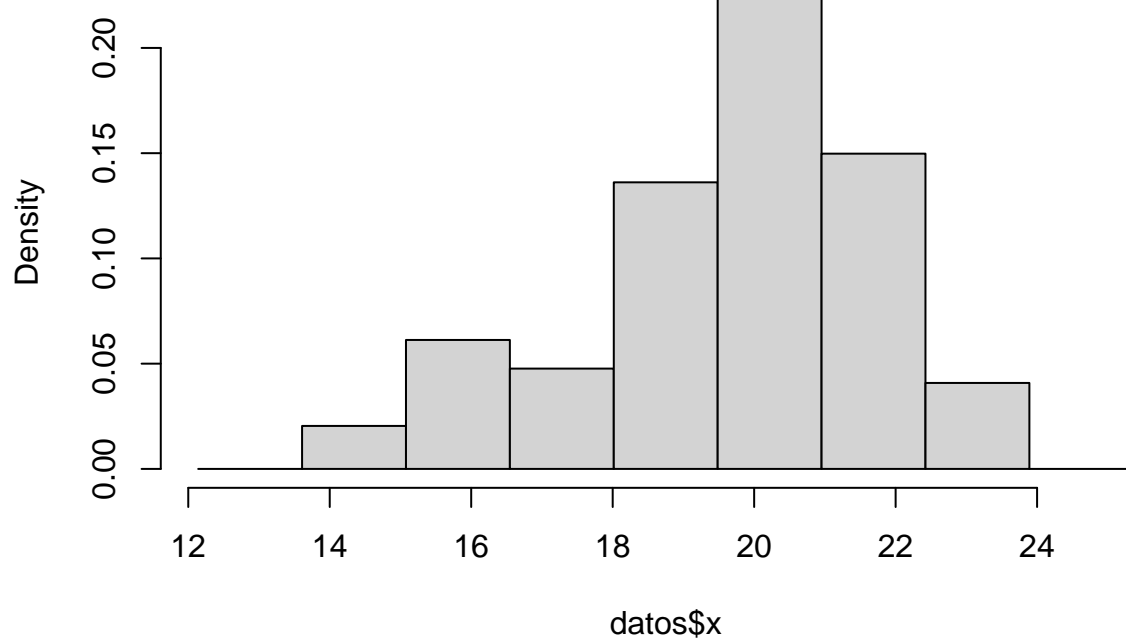
i)

```

hist(datos$x,freq=F,
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10))

```

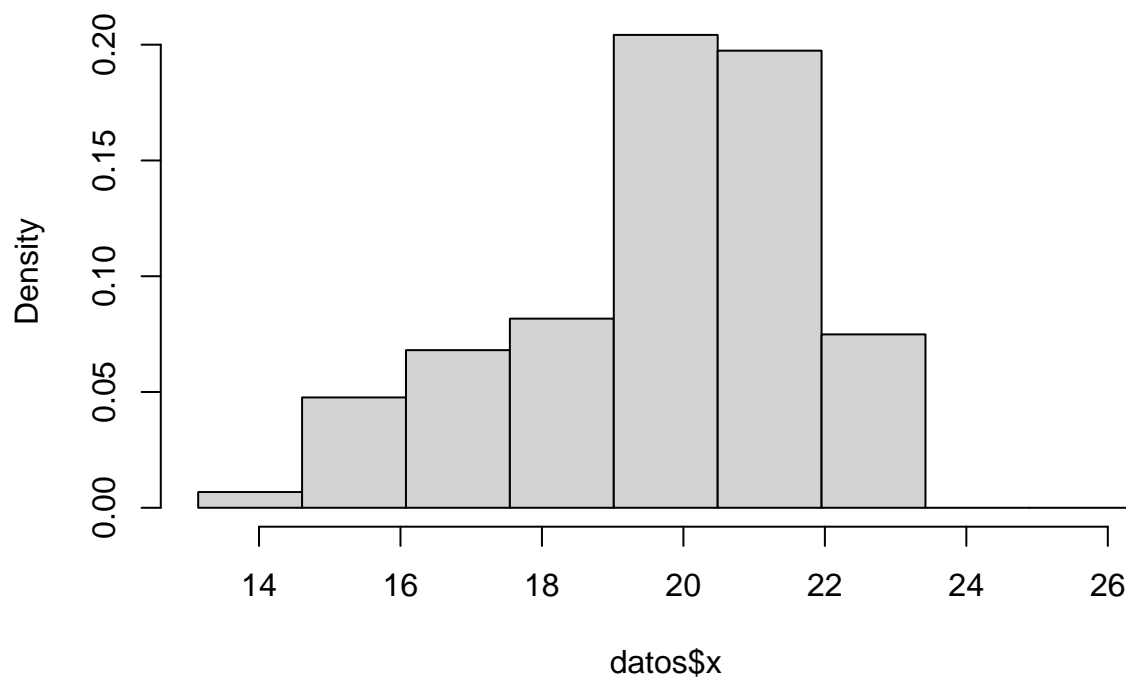
Histogram of datos\$x



ii)

```
hist(datos$x,freq=F,  
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10)+1)
```

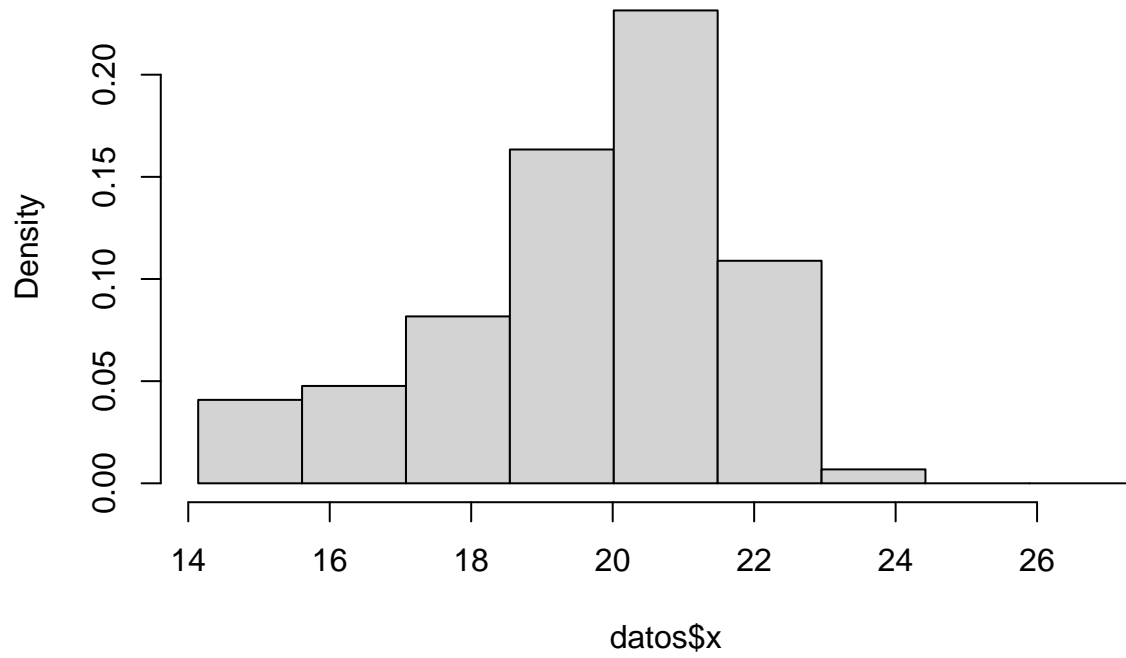
Histogram of datos\$x



iii)

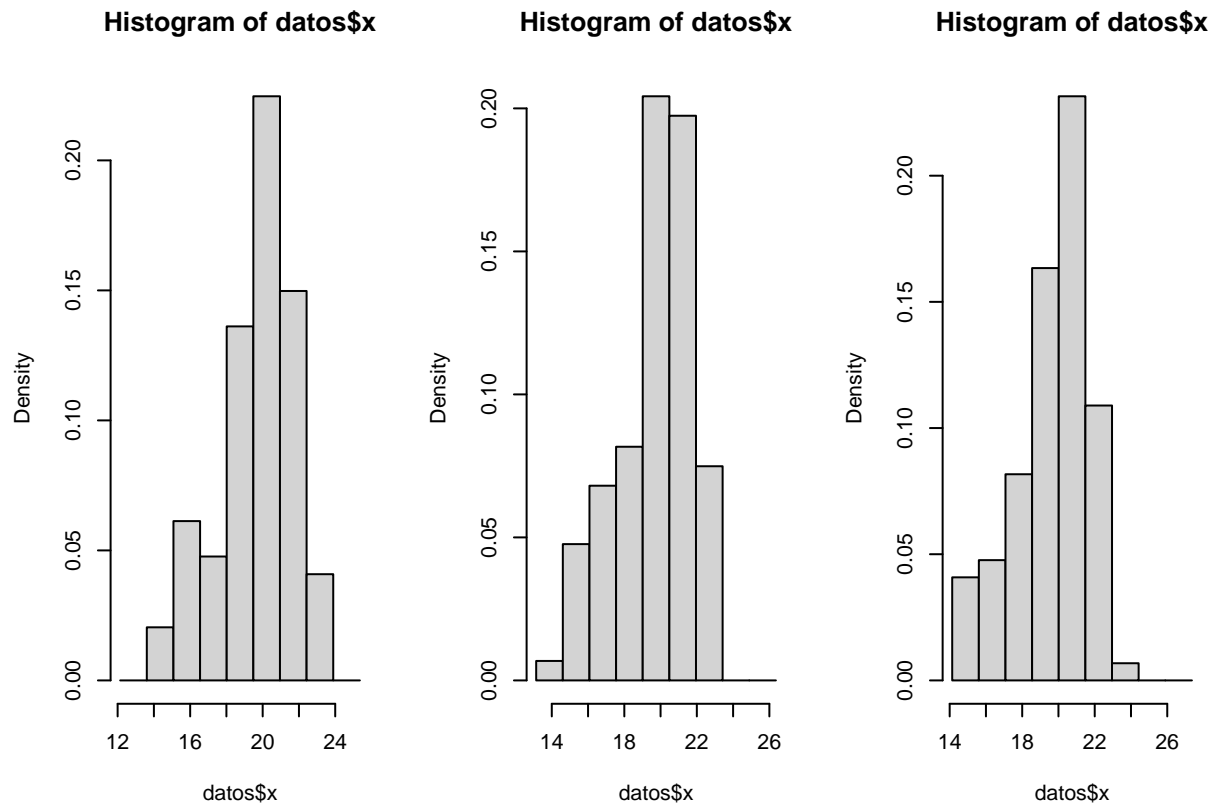
```
hist(datos$x,freq=F,
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10)+1+1)
```

Histogram of datos\$x



Ploteamos los 3

```
par(mfrow=c(1,3))
hist(datos$x,freq=F,
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10))
hist(datos$x,freq=F,
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10)+1)
hist(datos$x,freq=F,
      breaks=seq(min(datos$x)-2,max(datos$x)+2,length.out=10)+1+1)
```



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

Vemos que el efecto es el de mover el histograma hacia la izquierda o hacia la derecha.

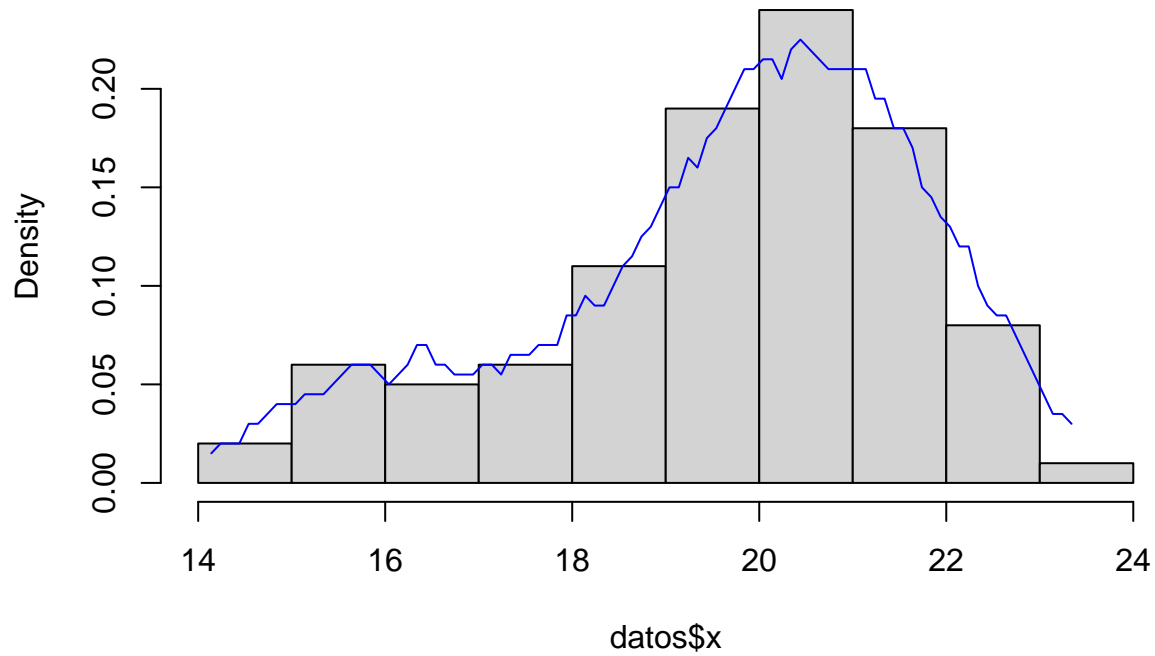
4.

```
K_rect=function(t){
  1/2*ifelse(-1<= t & t<= 1,1,0)
}
densidad.est.parzen=function(datos,h,z){
  mean(K_rect((z-datos)/h))/h
}
```

5. Aproximamos la densidad con densidad.est.parzen y la superponemos al histograma.

```
grilla=seq(min(datos$x),max(datos$x),0.1)
hist(datos$x,freq=F)
lines(grilla, sapply(grilla,densidad.est.parzen,datos=datos$x,h=1),
      col='blue')
```

Histogram of datos\$x



Podemos aproximar $f(20)$ con $h=1$ haciendo

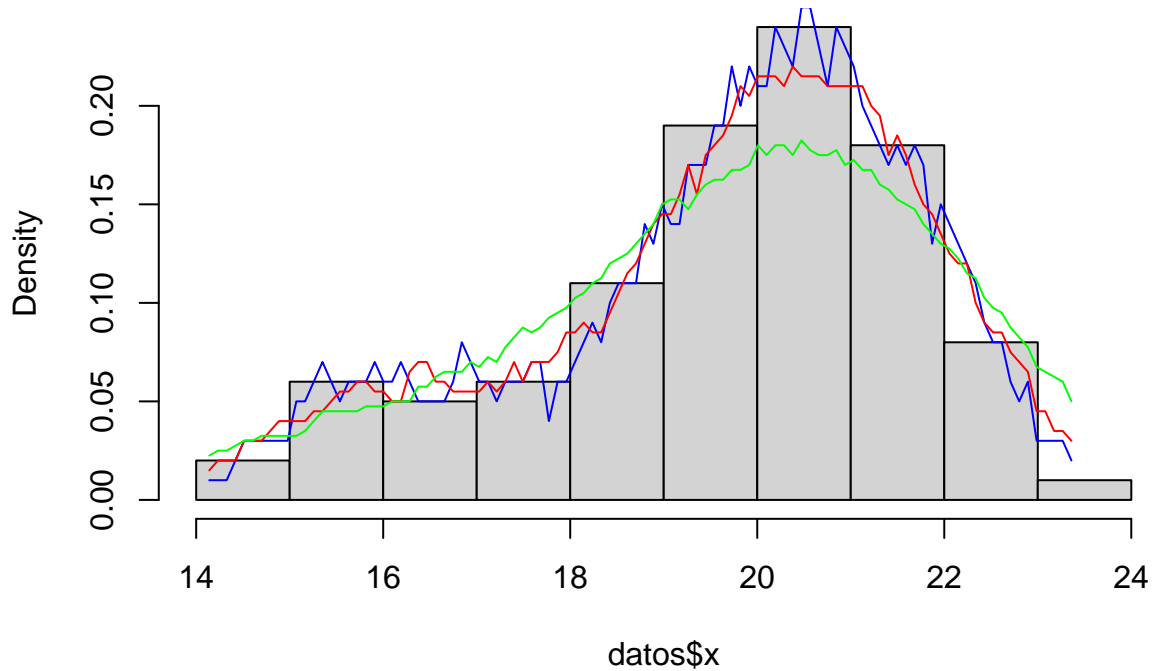
```
densidad.est.parzen(datos$x,1,20)
```

```
## [1] 0.215
```

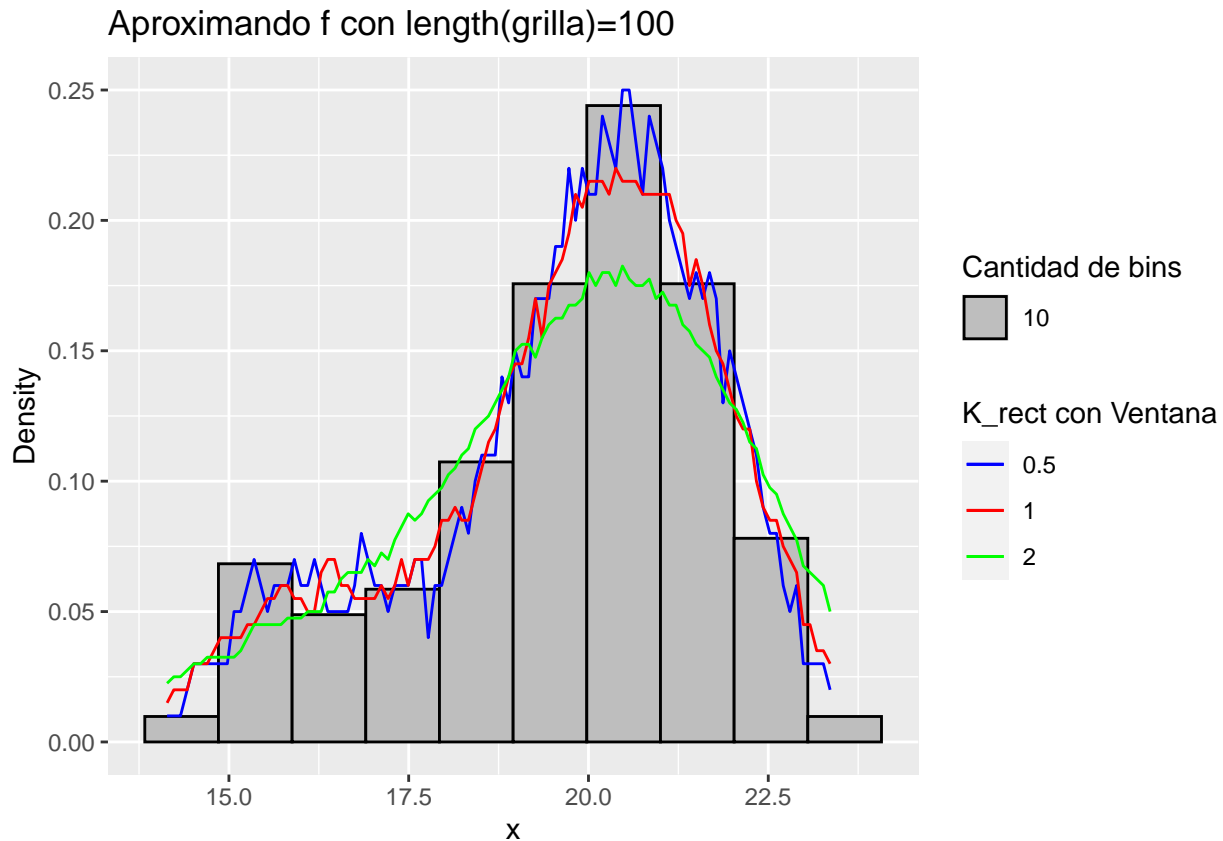
6. Ploteamos las 3 aproximaciones, $h=0.5, 1, 2$

```
grilla=seq(min(datos$x),max(datos$x),length.out = 100)
h=c(0.5,1,2)
hist(datos$x,freq=F)
lines(grilla, sapply(grilla,densidad.est.parzen,datos=datos$x,h=h[1]),
      col='blue')
lines(grilla, sapply(grilla,densidad.est.parzen,datos=datos$x,h=h[2]),
      col='red')
lines(grilla, sapply(grilla,densidad.est.parzen,datos=datos$x,h=h[3]),
      col='green')
```

Histogram of datos\$x



```
#####
aproximaciones_h=c()
for(i in h){
  aproximaciones_h=cbind(aproximaciones_h,
                          supply(grilla,densidad.est.parzen,datos=datos$x,h=i))
}
datos_plot=data.frame(cbind('Grilla'=grilla,
                             aproximaciones_h))
names(datos_plot)=c('Grilla',paste('Ventana_',h,sep = ""))
datos_plot %>%
  ggplot()+
  geom_histogram(data=datos,aes(x=x,y=..density..,fill='Datos'),
                 bins=10,col='black')+
  geom_line(aes(x=Grilla,y=Ventana_0.5,col='0.5'))+
  geom_line(aes(x=Grilla,y=Ventana_1,col='1'))+
  geom_line(aes(x=Grilla,y=Ventana_2,col='2'))+
  scale_color_manual("K_rect con Ventana",
                     breaks=c('0.5','1','2'),
                     values=c('blue','red','green'))+
  scale_fill_manual("Cantidad de bins",
                    breaks=c('Datos'),
                    values=c('grey'),
                    labels='10')+
  ylab("Density") +
  labs(title="Aproximando f con length(grilla)=100")
```

```
# ggsave('Item 6.png')
```

Podemos ver que a menor ventana la aproximación es mas fiel a los datos pero tiene mas picos y a mayor ventana la densidad aproximada es mas suave pero ya no es tan precisa.

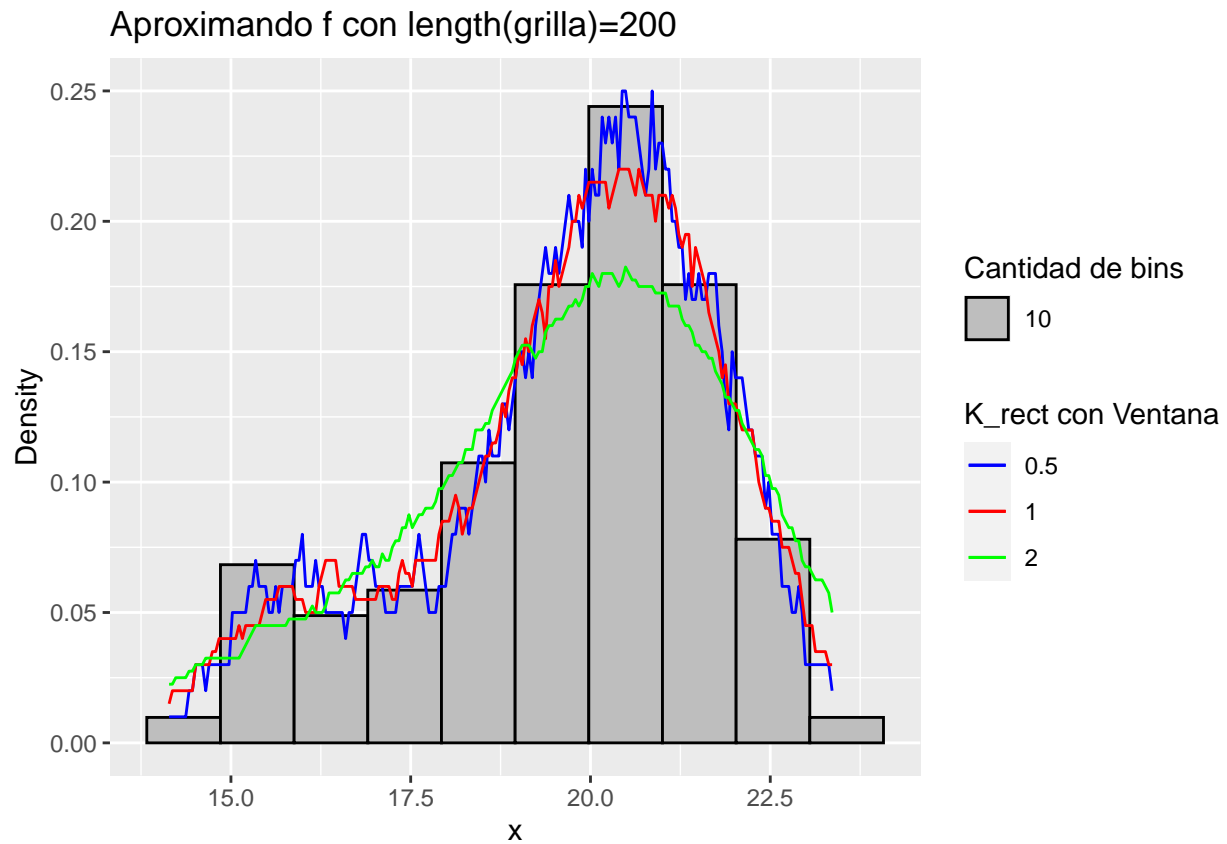
7. Hicimos el item anterior para una grilla equiespaciada con paso 0.1 y eso nos daba un total de 93 puntos equiespaciados. Repitamos lo que hicimos antes pero para la grilla que nos piden.

```
grilla=seq(min(datos$x),max(datos$x),length.out = 200)
#####3
aproximaciones_h=c()
for(i in h){
  aproximaciones_h=cbind(aproximaciones_h,
                          sapply(grilla,densidad.est.parzen,datos=datos$x,h=i))
}
datos_plot=data.frame(cbind('Grilla'=grilla,
                             aproximaciones_h))
names(datos_plot)=c('Grilla',paste('Ventana_',h,sep = ""))
datos_plot %>%
  ggplot()+
  geom_histogram(data=datos,aes(x=x,y=..density..,fill='Datos'),
                 bins=10,col='black')+
  geom_line(aes(x=Grilla,y=Ventana_0.5,col='0.5'))+
  geom_line(aes(x=Grilla,y=Ventana_1,col='1'))+
  geom_line(aes(x=Grilla,y=Ventana_2,col='2'))+
  scale_color_manual("K_rect con Ventana",
                     breaks=c('0.5','1','2'),
                     values=c('blue','red','green'))+
```

```

scale_fill_manual("Cantidad de bins",
  breaks=c('Datos'),
  values=c('grey'),
  labels='10')+
ylab("Density") +
labs(title="Aproximando f con length(grilla)=200")

```



```

# ggsave('Item 7.png')

```