

Práctica4

a) Estudiar la influencia del tamaño del melanoma en el tiempo de supervivencia

Antes de realizar este estudio lo primero que hay que hacer es inspeccionar los datos con los que estamos trabajando. Para ello como estamos acostumbrados usaremos head, tail y summary.

```
#CARGAMOS LOS DATOS MELANOMA DEL DIRECTORIO
setwd("~/Escritorio/R/Practica 1")
melanoma <- read.table("Datos_Melanoma.txt", sep = "\t", header = TRUE)
```

```
#INSPECCIONAMOS LA TABLA DE DATOS
```

```
head(melanoma)
```

```
##      tiempo status sexo edad grosor ulceracion
## 1         10      3    1   76   6.76          1
## 2          30      3    1   56   0.65          0
## 3          35      2    1   41   1.34          0
## 4          99      3    0   71   2.90          0
## 5         185      1    1   52  12.08          1
## 6         204      1    1   28   4.84          1
```

```
tail(melanoma)
```

```
##      tiempo status sexo edad grosor ulceracion
## 200     4479      2    0   19   1.13          1
## 201     4492      2    1   29   7.06          1
## 202     4668      2    0   40   6.12          0
## 203     4688      2    0   42   0.48          0
## 204     4926      2    0   50   2.26          0
## 205     5565      2    0   41   2.90          0
```

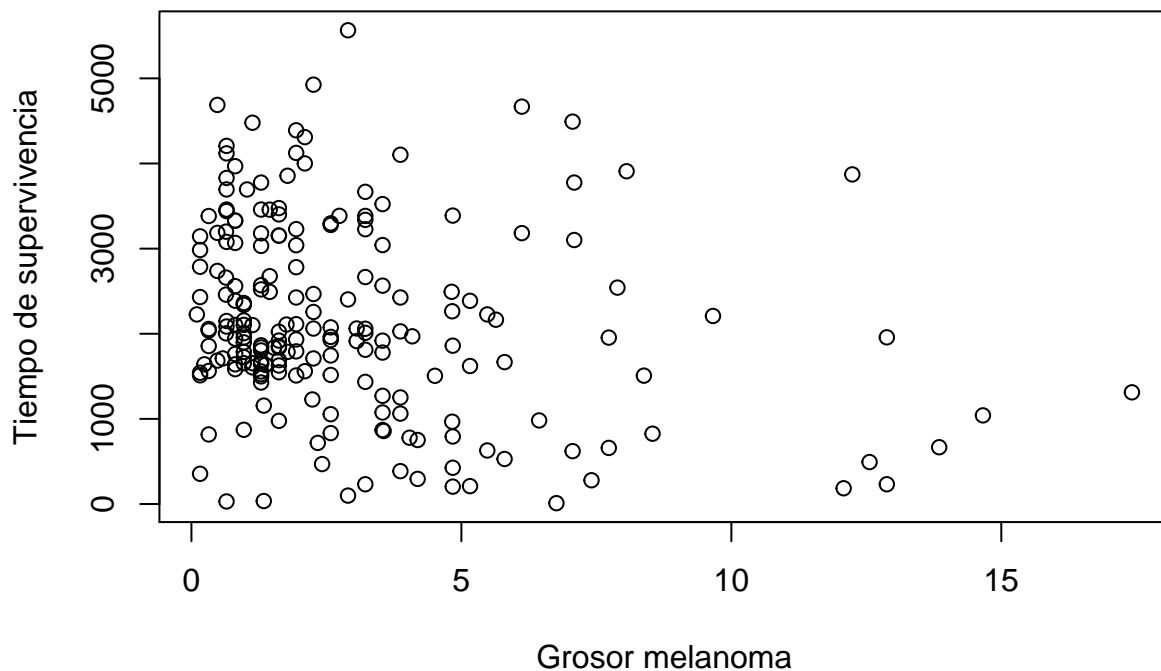
```
summary(melanoma)
```

```
##      tiempo      status      sexo      edad      grosor
## Min.   : 10   Min.   :1.00   Min.   :0.0000   Min.   : 4.00   Min.   : 0.10
## 1st Qu.:1525   1st Qu.:1.00   1st Qu.:0.0000   1st Qu.:42.00   1st Qu.: 0.97
## Median :2005   Median :2.00   Median :0.0000   Median :54.00   Median : 1.94
## Mean   :2153   Mean   :1.79   Mean   :0.3854   Mean   :52.46   Mean   : 2.92
## 3rd Qu.:3042   3rd Qu.:2.00   3rd Qu.:1.0000   3rd Qu.:65.00   3rd Qu.: 3.56
## Max.   :5565   Max.   :3.00   Max.   :1.0000   Max.   :95.00   Max.   :17.42
##      ulceracion
## Min.   :0.000
## 1st Qu.:0.000
## Median :0.000
## Mean   :0.439
## 3rd Qu.:1.000
## Max.   :1.000
```

Una vez realizado esto comenzaremos el estudio. Primero mostraremos gráficamente usando plot una gráfica muy sencilla para tener una vista previa

```
plot(melanoma$grosor, melanoma$tiempo,
     xlab="Grosor melanoma",
     ylab="Tiempo de supervivencia",
     main="Tiempo de supervivencia en funcion del grosor del melanoma")
```

Tiempo de supervivencia en funcion del grosor del melanoma



Desde aqui podemos observar ya que el tamaño del melanoma parece tener cierta relación con el tiempo de supervivencia, sigamos indagando. Para ello primero observemos los diferentes tamaños que tenemos. Y a su vez dividamos en grupos los mismos desde 0 mm a 20mm de 5 en 5 usando la funcion cut. De esta manera podremos estudiar mas detalladamente su relación

```
sort(unique(melanoma$grosor))
```

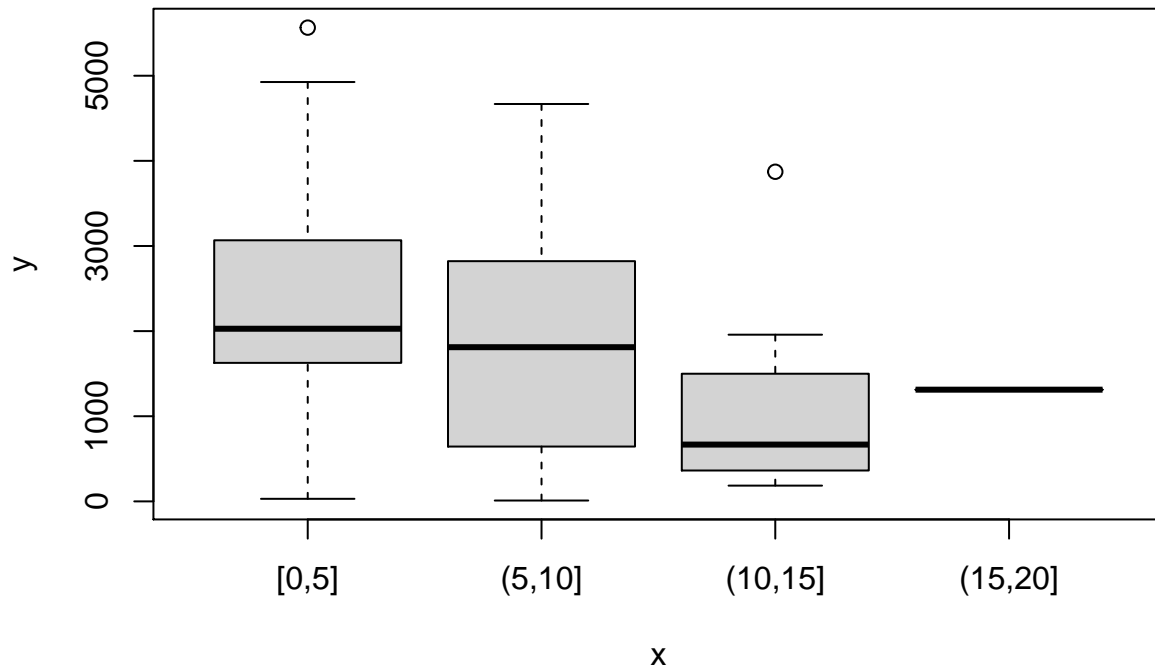
```
## [1] 0.10 0.16 0.24 0.32 0.48 0.58 0.64 0.65 0.81 0.97 1.03 1.13
## [13] 1.29 1.34 1.37 1.45 1.53 1.62 1.76 1.78 1.94 2.10 2.24 2.26
## [25] 2.34 2.42 2.58 2.74 2.90 3.06 3.22 3.54 3.56 3.87 4.04 4.09
## [37] 4.19 4.51 4.82 4.83 4.84 5.16 5.48 5.64 5.80 6.12 6.44 6.76
## [49] 7.06 7.09 7.41 7.73 7.89 8.06 8.38 8.54 9.66 12.08 12.24 12.56
## [61] 12.88 13.85 14.66 17.42
```

```
melanoma_agrupado <- cut(melanoma$grosor, breaks=seq(0, 20, 5), include.lowest=TRUE, right=TRUE)
```

Por ultimo creamos una tabla usando los datos del tamaño acotado. Por ultimo mostraremos en una gráfica la relacion entre el tamaño y el tiempo

```
melanoma_agrupado_edad <- cut(melanoma$edad, breaks=seq(0, 100, 15), include.lowest=TRUE, right=TRUE)
```

```
tabla_agrupado <- data.frame(size = melanoma_agrupado, tiempo = melanoma$tiempo, ulceracion = melanoma$ulceracion)
plot(tabla_agrupado$size, tabla_agrupado$tiempo)
```



Aquí ya vemos claramente que de 0 a 5mm la tasa de mortalidad por lo general es bastante mas alta que en el resto, llegando a de 10mm a 15mm ser la menor. También cabe destacar que por encima de 15 hay muy pocos datos que se hayan tomado por lo tanto sera difícil de estudiar este caso.

Continuemos calculando las medias, medianas y desviaciones típicas que nos daran mas información.

```
periodos <- unique(tabla_agrupado$size)

#CALCULAMOS MEDIAS DE DURACION EN CADA PERIODO

#las medianas indican que la gran diferencia
#se nota a partir de los 10mm de grosor
#Aunque los datos la mayoría se tomaron de melanomas menores a 5mm
mean(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[1]])

## [1] 1923.083
mean(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[2]])

## [1] 2227.803
mean(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[3]])

## [1] 1206.857
```

```

mean(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[4]])

## [1] 1312
#CALCULAMOS LAS MEDIANAS EN CADA PERIODO
median(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[1]])

## [1] 1811
median(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[2]])

## [1] 2028
median(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[3]])

## [1] 667
median(tabla_agrupado$tiempo[tabla_agrupado$size == periodos[4]])

## [1] 1312

```

Como podemos observar existe diferencia sustancial Entre los tamaños mayores y menores aunque también observamos un fenomeno curioso la media de 0 a 5 mm es menor que de 5 a 10 mm esto puede ser cierto sin embargo cabe destacar la gran dispersión que existe de 5 a 10mm en cuanto a los datos se refiere y la gran concentración en cuanto de 0 a 5mm. También podemos observar como la media de 15 a 20mm es mayor (aunque no mucho) a la de 10 a 15mm esta muestra no la consideraría representativa puesto que hay muy pocos datos que estudiar en el último caso.

A su vez es importante analizar la cantidad de muestras de cada grupo que tenemos

```

#CONTAMOS LA CANTIDAD DE DATOS RECOGIDOS DE CADA
length(melanoma$grosor[melanoma$grosor < 5]) #demasiados datos de este

## [1] 173
length(melanoma$grosor[melanoma$grosor > 5]) #muy pocos de este

## [1] 32

```

Como podemos observar 173 de los casos son de personas con melanomas por debajo de 5mm y 32 e por encima de 5mm. Debido a la desigualdad en la cantidad de datos entre grupos podemos considerar que el estudio no ha sido realizado homogéneamente ya que tenemos mas individuos en unos estratos que en otros. Esto puede ser un problema a la hora de sacar conclusiones.

A pesar de lo dicho anteriormente podemos observar claramente que mientras menor sea el tamaño del melanoma mas esperanza de vida tras la operación se tendrá esto lo muestran tanto las gráficas como la amplia diferencia en las medias de los estratos por debajo de 10mm y por encima de 10mm.

b) Explicar las diferencias sustantivas entre los casos que conlleven un tamaño inferior a 10mm y los que tienen un tamaño mayor.

Como hemos dicho anteriormente existe una gran diferencia entre los datos tomados en los anteriores grupos sin embargo merece la pena realizar este estudio. Para ello primero dividiremos los datos en un grupo de 0 a 10mm y otro de 10 a 20mm. Mostraremos usando plot ambos.

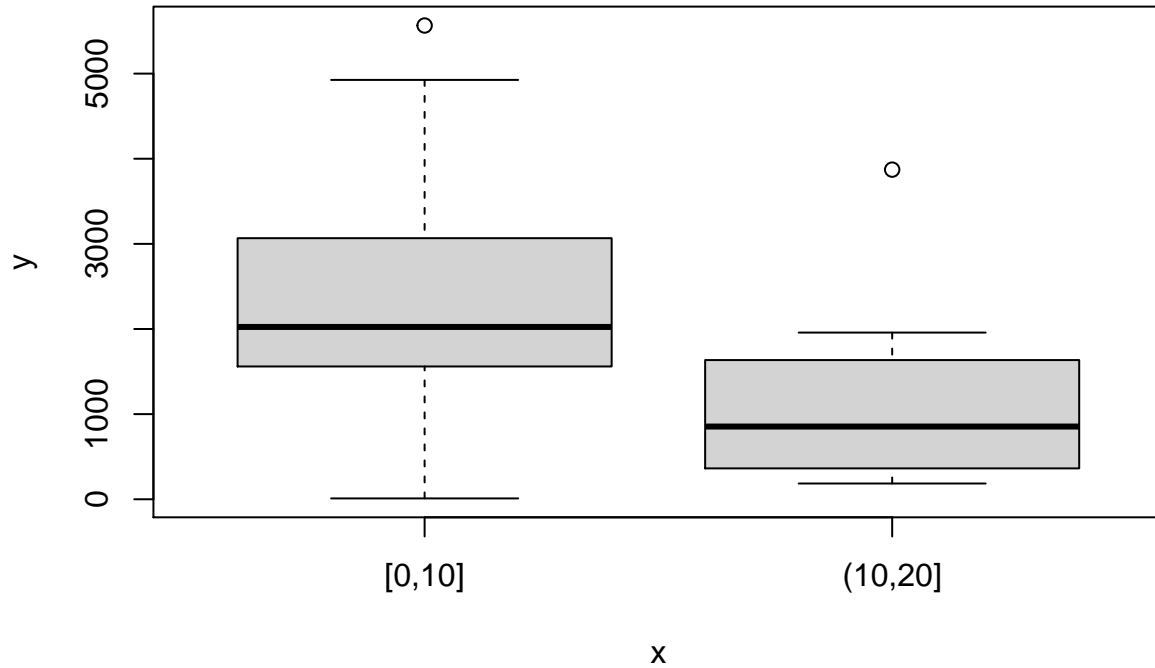
```

melanoma_agrupado_2 <- cut(melanoma$grosor, breaks=seq(0, 20, 10), include.lowest=TRUE, right=TRUE)

tabla_agrupado <- data.frame(size = melanoma_agrupado_2, tiempo = melanoma$tiempo, ulceracion = melanoma$ulceracion)

plot(tabla_agrupado$size, tabla_agrupado$tiempo)

```



Como podemos observar en la gráfica el primer grupo tiene una gran diferencia en cuanto a esperanza de vida respecto al segundo.

```
mean(melanoma$tiempo[melanoma$grosor < 10])
```

```
## [1] 2190.68
```

```
mean(melanoma$tiempo[melanoma$grosor > 10])
```

```
## [1] 1220
```

A su vez como podemos observar en las medias esta diferencia sigue siendo considerable. Aunque otro factor que merece la pena estudiar es la media del grosor de ambos grupos.

```
mean(melanoma$grosor[melanoma$grosor < 10])
```

```
## [1] 2.48731
```

```
mean(melanoma$grosor[melanoma$grosor > 10])
```

```
## [1] 13.57125
```

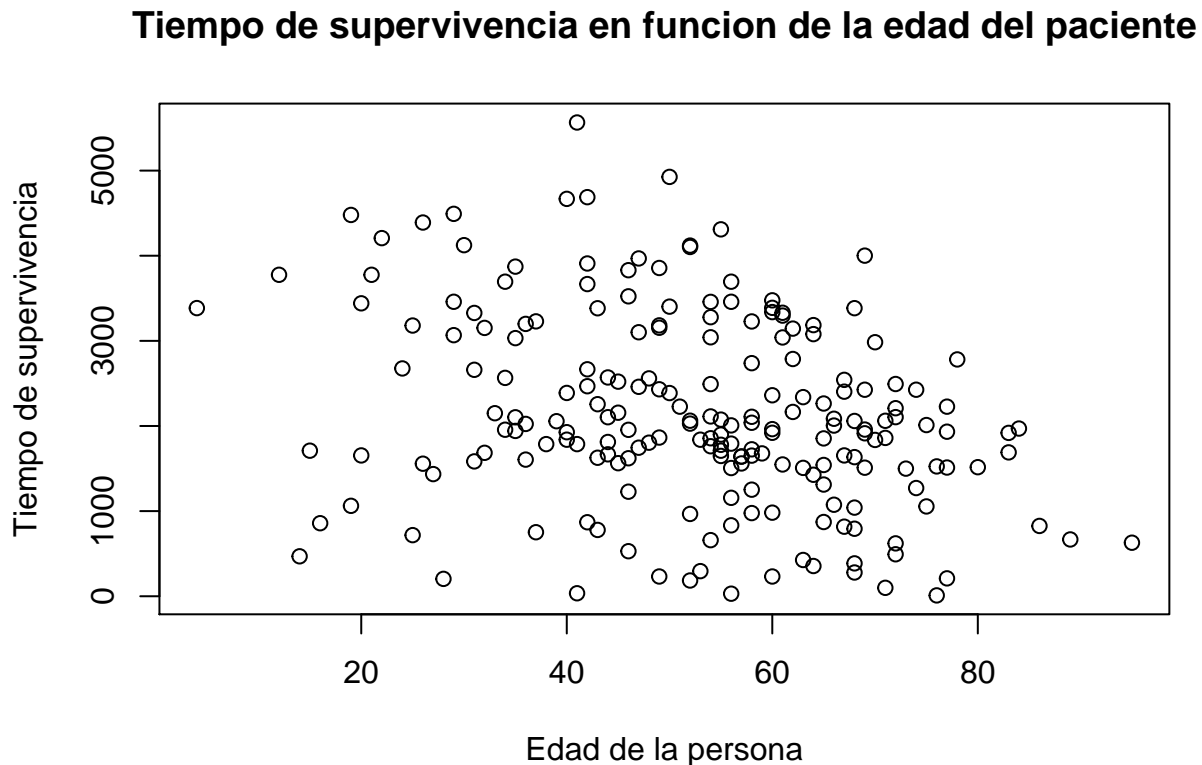
Como podemos observar en nuestro primer grupo la mayoría de datos son bastante alejados a los 10mm, esto vuelve a demostrar lo que hablabamos antes, las muestras no son homogéneas.

Aun así podemos concluir que la diferencia de esperanza de vida tras la operación en los valores por debajo de 10mm y por encima es abismal.

c) Estudiar el efecto de la edad en la supervivencia a la enfermedad tras la operación

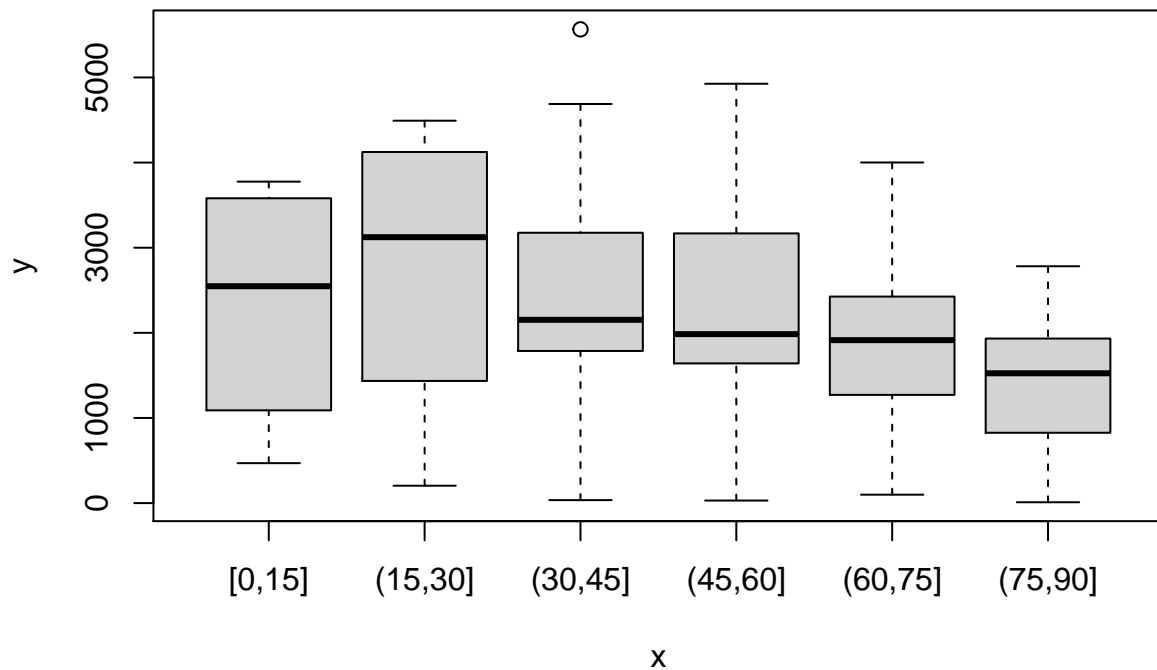
Para llevar a cabo esta tarea realizaremos un estudio muy parecido al que hemos realizado anteriormente. Primero mostraremos por pantalla un gráfico que nos muestre por encima los datos.

```
plot(melanoma$edad, melanoma$tiempo,  
     xlab="Edad de la persona",  
     ylab="Tiempo de supervivencia",  
     main="Tiempo de supervivencia en funcion de la edad del paciente")
```



El gráfico mostrado es complicado de estudiar debido a que los datos estan muy dispersos, por ello crearemos grupos de edad como realizamos anteriormente con los tamaños.

```
melanoma_agrupado_edad <- cut(melanoma$edad, breaks=seq(0, 100, 15), include.lowest=TRUE, right=TRUE)  
  
sort(unique(melanoma$edad))  
  
## [1]  4 12 14 15 16 19 20 21 22 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39  
## [26] 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64  
## [51] 65 66 67 68 69 70 71 72 73 74 75 76 77 78 80 83 84 86 89 95  
  
melanoma_agrupado_frame <- na.omit(data.frame(edad = melanoma_agrupado_edad, tiempo = melanoma$tiempo))  
  
plot(melanoma_agrupado_frame$edad, melanoma_agrupado_frame$tiempo)
```



Ahora podemos observar mas claramente los diferentes datos y como vemos tenemos una gráfica que muestra como mientras mas edad se tenga menos supervivencia se tendrá aunque de 0 a 15 años, aunque la esperanza de vida sea mayor a los estratos mas adultos, se tiene menores posibilidades de sobrevivir que de 15 a 30 años. Y como bien se menciona antes el grupo con mayor probabilidad de sobrevivir mas tiempo después de la operación son aquellos de 15 a 30 años.

Continuemos calculando las medaias de todos los grupos

```
m_tp_vida_edad = c()

for(x in 1:length(unique(melanoma_agrupado_edad))){
  m_tp_vida_edad[x] = mean(melanoma_agrupado_frame$tiempo[melanoma_agrupado_frame$edad == unique(melanoma_agrupado_edad)[x]])
}

m_tp_vida_edad <- m_tp_vida_edad

mean_edad_mortalidad <- na.omit(data.frame(edad = unique(melanoma_agrupado_edad), media_vida = m_tp_vida_edad))

library(knitr)

kable(mean_edad_mortalidad)
```

edad	media_vida
(75,90]	1445.154
(45,60]	2195.000
(30,45]	2438.341
(60,75]	1857.774

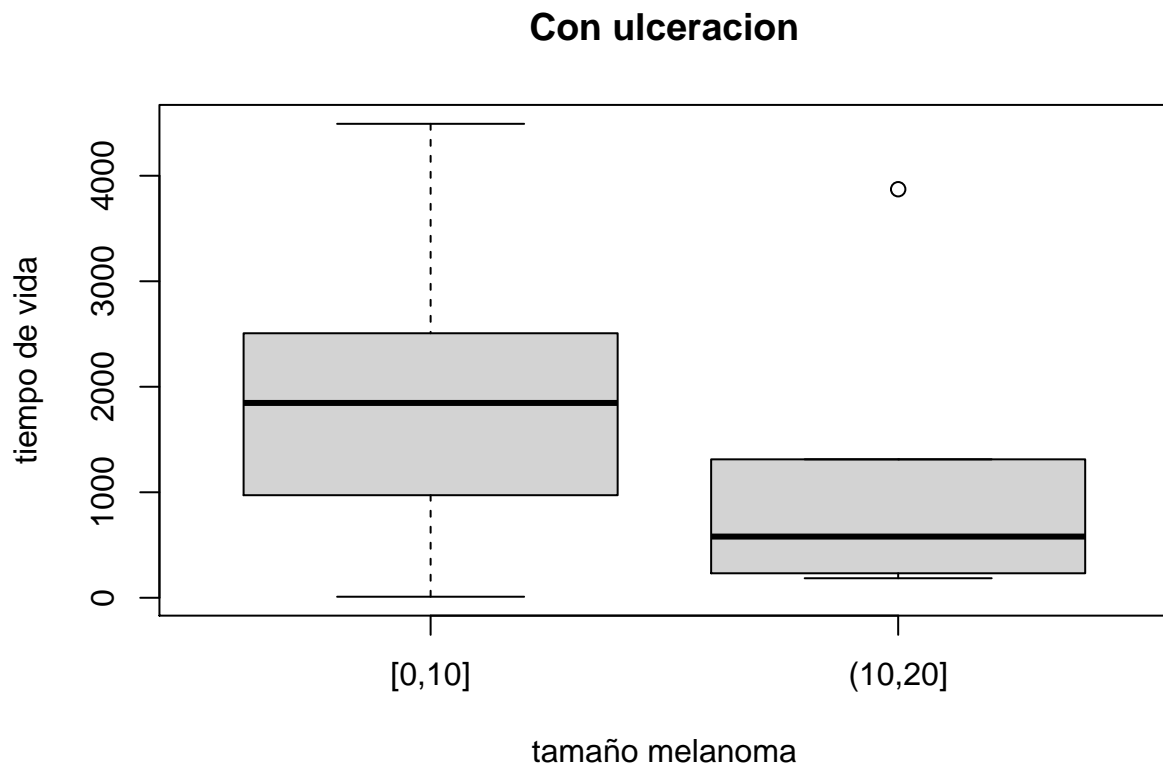
edad	media_vida
(15,30]	2709.944
[0,15]	2335.000

Como observamos en la tabla mostrada anteriormente se ve claramente como mientras mayor sea la edad menor sera el tiempo de supervivencia tras la operación. El grupo de 75 a 90 años se muestra claramente como el que mas complicaciones tiene, siendo el grupo de 15 a 30 años el que menos como dijimos anteriormente.

d) Analizar si la presencia y diagnóstico de ulceración afecta a los resultados a) y b)

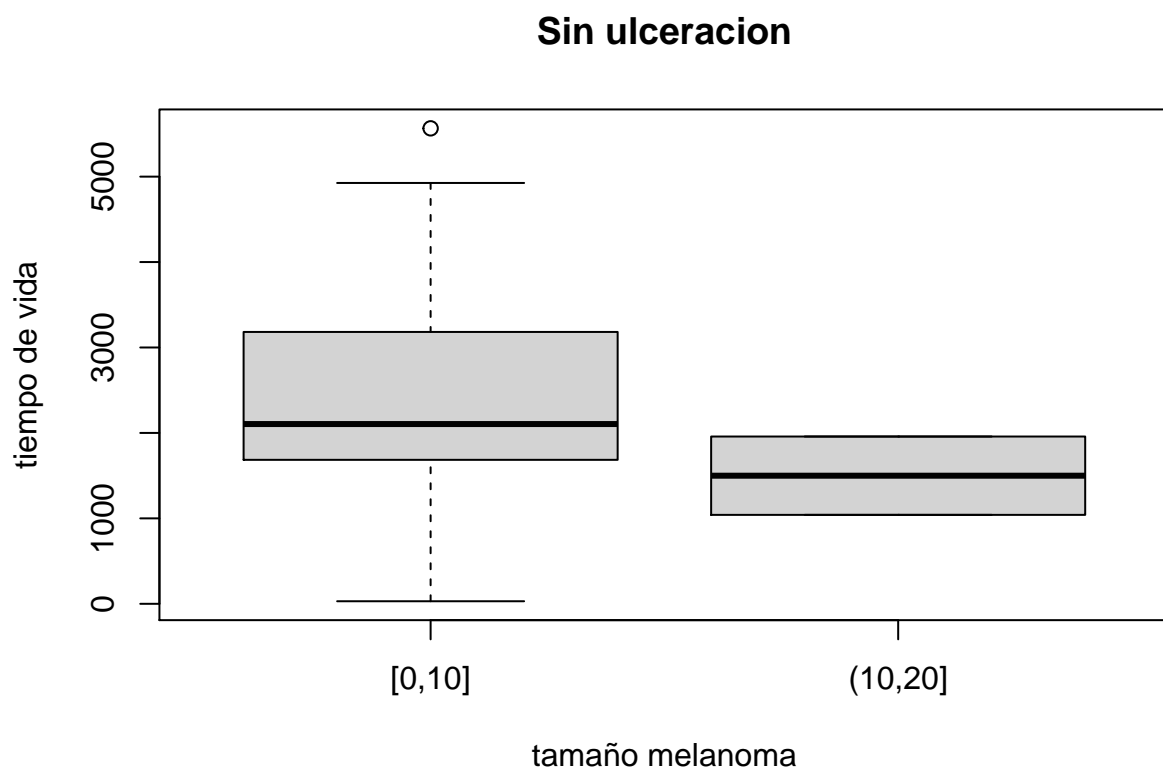
#CON ULCERACION

```
plot(tabla_agrupado$size[tabla_agrupado$ulceracion == 1], tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 1])
```



#SIN ULCERACION

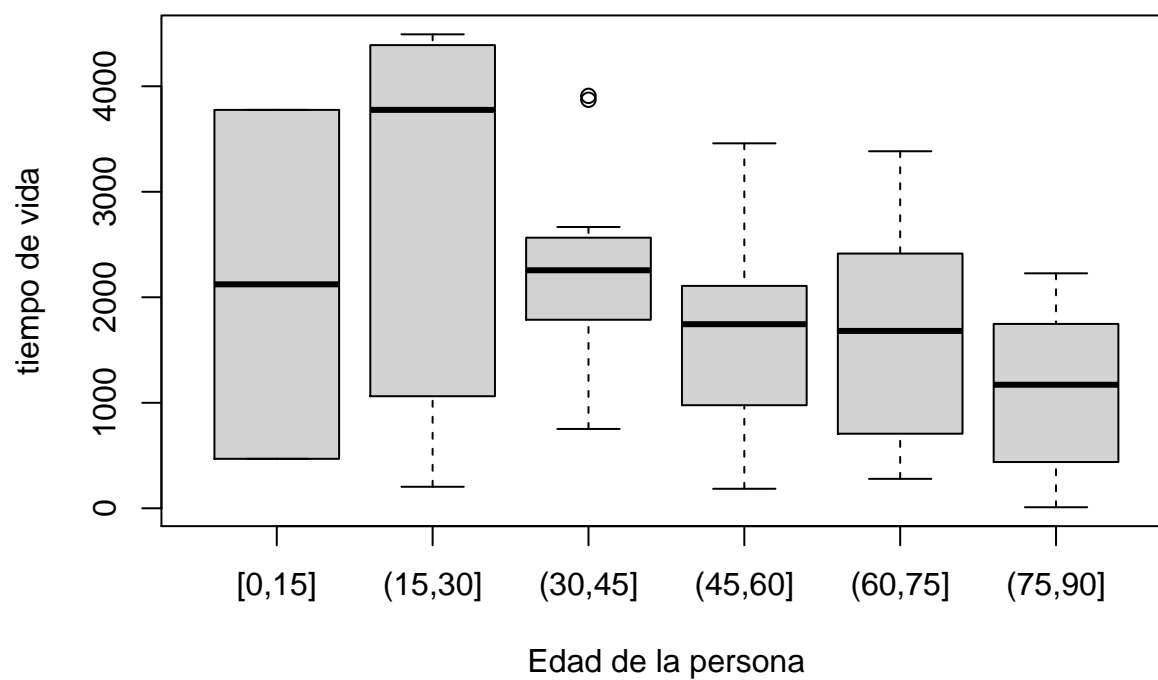
```
plot(tabla_agrupado$size[tabla_agrupado$ulceracion == 0], tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 0])
```

#CON ULCERACION

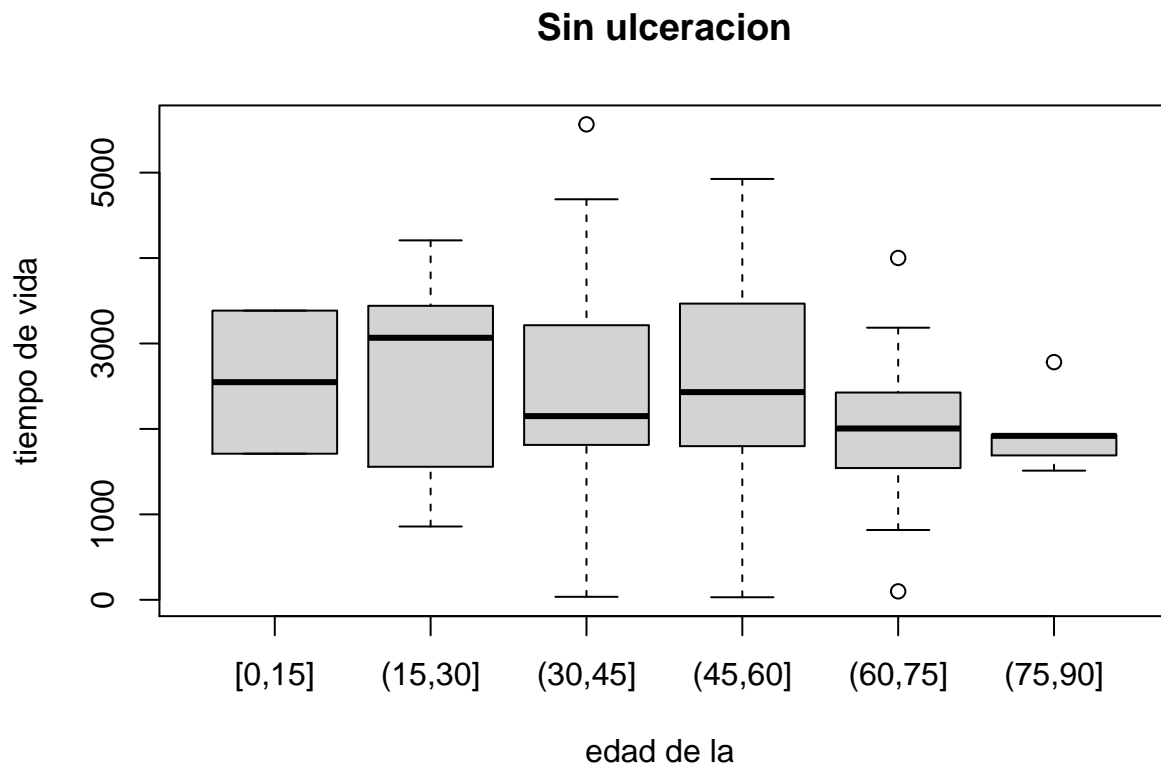
```
plot(tabla_agrupado$edad[tabla_agrupado$ulceracion == 1], tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 1])
```

Con ulceracion



#sin ulceracion

```
plot(tabla_agrupado$edad[tabla_agrupado$ulceracion == 0], tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 0])
```



#CON ULCERACION

```
mean(tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 1])
```

```
## [1] 1817.811
```

```
median(tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 1])
```

```
## [1] 1799.5
```

```
length(melanoma$tiempo[melanoma$ulceracion == 1])
```

```
## [1] 90
```

#SIN ULCERACION

```
mean(tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 0])
```

```
## [1] 2414.965
```

```
median(tabla_agrupado$tiempo[tabla_agrupado$ulceracion == 0])
```

```
## [1] 2103
```

```
length(melanoma$tiempo[melanoma$ulceracion == 0])
```

```
## [1] 115
```

Como podemos observar en las gráficas expuestas la ulceración afecta negativamente a la supervivencia tras la enfermedad, mostrando no solo datos de mayor tiempo de supervivencia sino mas concentrados cuando no existe la ulceración, tanto para personas de mayor y menor edad y melanomas de menor y mayor tamaño.

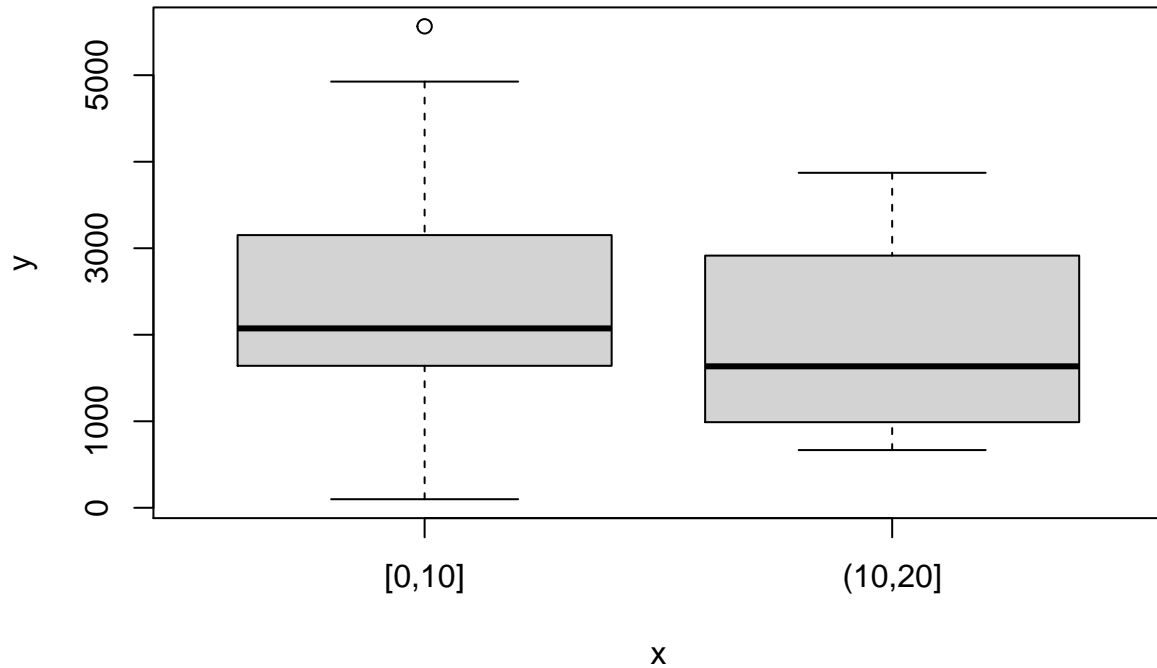
e) ¿Existe alguna diferencia de los resultados si se trata de un hombre o una mujer?

Para averiguar esto realizaremos lo mismo que en el anterior apartado.

```
#HOMBRE
```

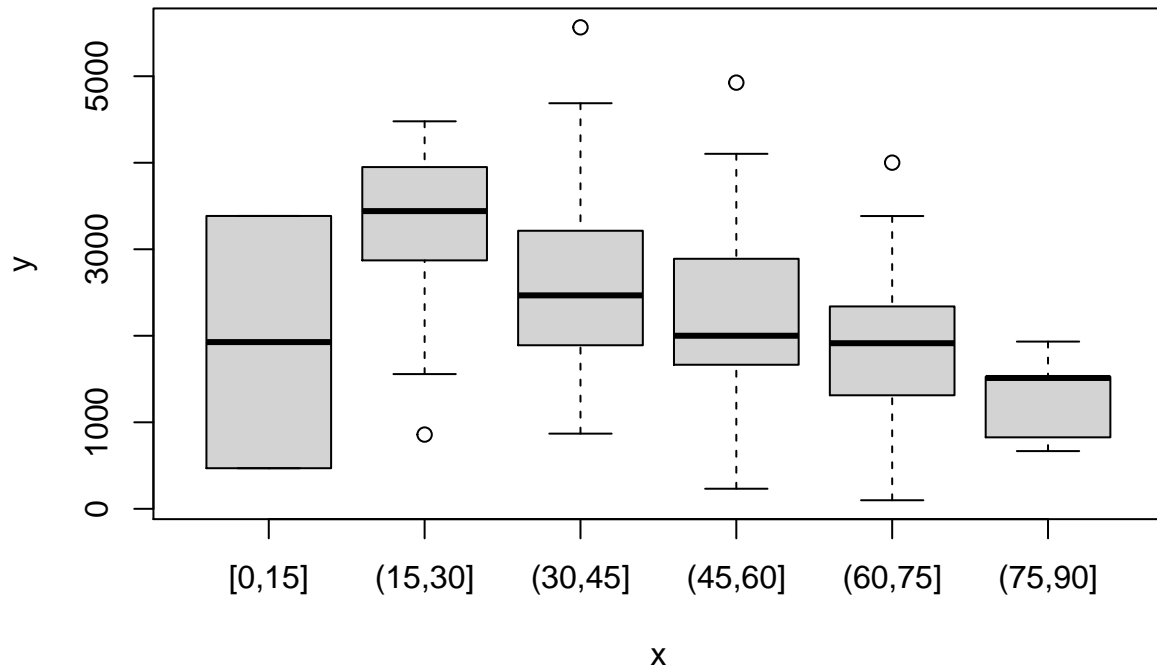
```
plot(tabla_agrupado$size[tabla_agrupado$sexo == 0], tabla_agrupado$tiempo[tabla_agrupado$sexo == 0], ma
```

Tamaño respecto tiempo de supervivencia en mujeres



```
plot(tabla_agrupado$edad[tabla_agrupado$sexo == 0], tabla_agrupado$tiempo[tabla_agrupado$sexo == 0], m
```

edad respecto tiempo de supervivencia en mujeres



```
#MEDIAS PARA LAS MUJERES
```

```
mean(tabla_agrupado$tiempo[tabla_agrupado$sexo == 0])
```

```
## [1] 2282.643
```

```
median(tabla_agrupado$tiempo[tabla_agrupado$sexo == 0])
```

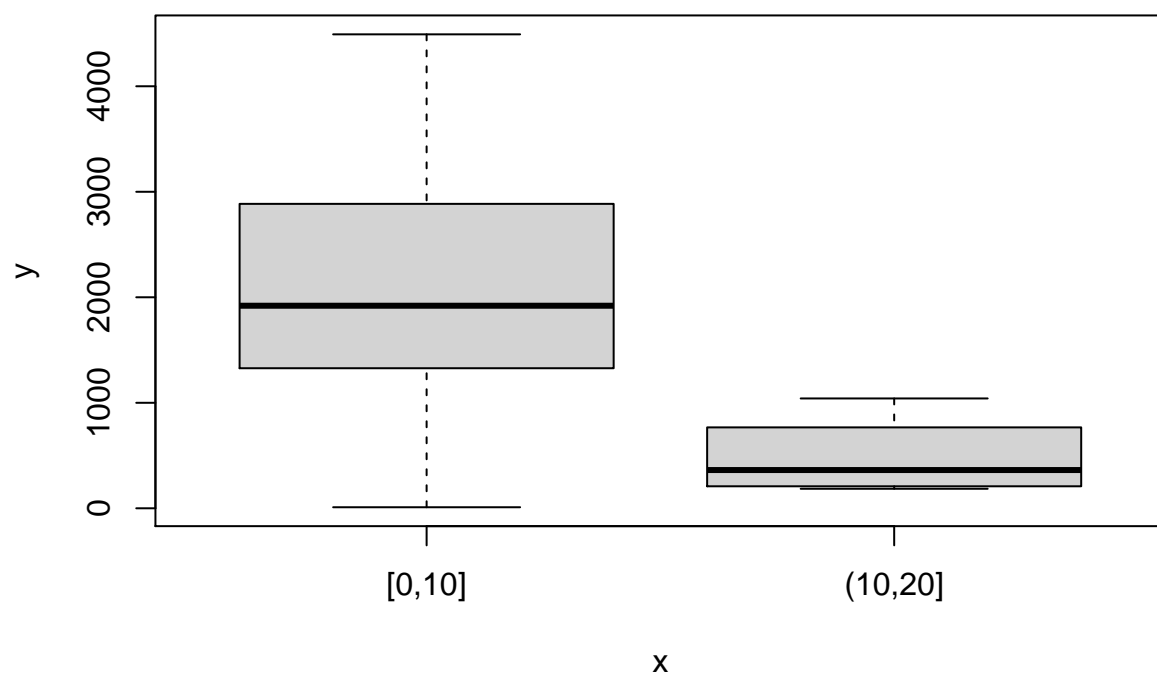
```
## [1] 2059
```

```
length(melanoma$tiempo[melanoma$sexo == 0])
```

```
## [1] 126
```

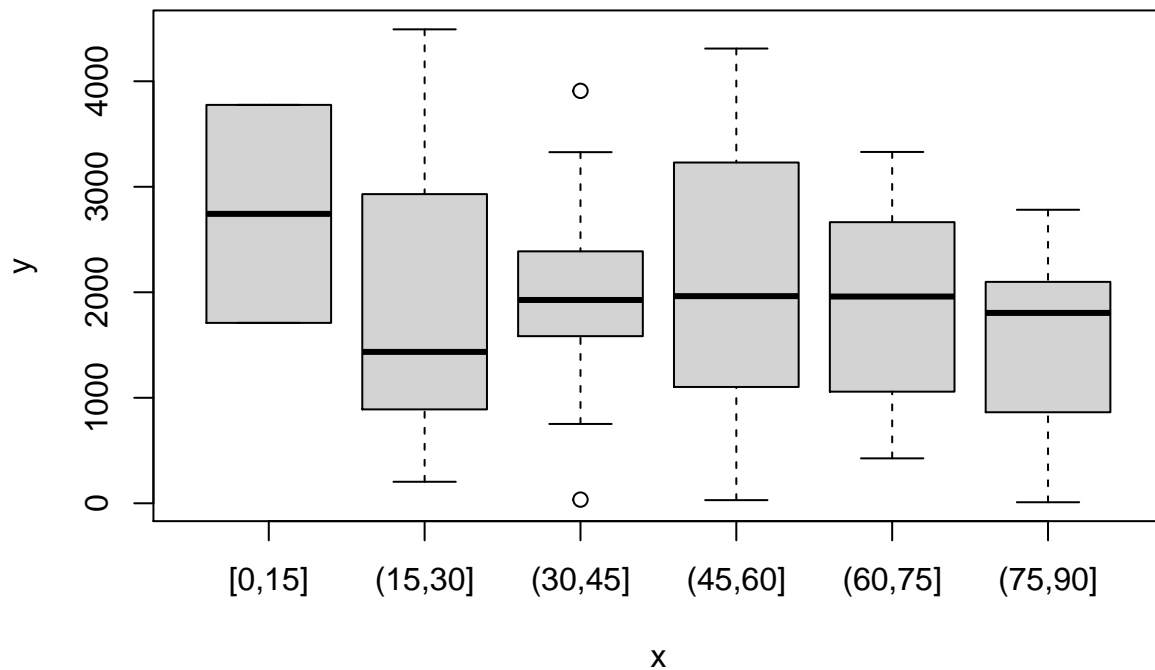
```
plot(tabla_agrupado$size[tabla_agrupado$sexo == 1], tabla_agrupado$tiempo[tabla_agrupado$sexo == 1], m
```

Tamaño respecto tiempo de supervivencia en hombres



```
plot(tabla_agrupado$edad[tabla_agrupado$sexo == 1], tabla_agrupado$tiempo[tabla_agrupado$sexo == 1], ma
```

edad respecto tiempo de supervivencia en hombres



#MEDIAS PARA LOS HOMBRES

```
mean(tabla_agrupado$tiempo[tabla_agrupado$sexo == 1])
```

```
## [1] 1945.709
```

```
median(tabla_agrupado$tiempo[tabla_agrupado$sexo == 1])
```

```
## [1] 1860
```

```
length(melanoma$tiempo[melanoma$sexo == 1])
```

```
## [1] 79
```

Como podemos observar en la gráfica del tamaño respecto al tiempo de supervivencia, las mujeres mantienen una esperanza de vida bastante parecida en todos los tamaños de melanomas, cosa totalmente diferente en hombres mostrando que si el tamaño es inferior a 10mm la sobreviven mas tiempo que las mujeres sin embargo cuando es por encima de 10mm decrece drásticamente.

En cuanto a la edad se ve que los hombres mas jovenes sobreviven mas que las mujeres y los hombres de mayor edad (por encima de los 60 años) igual. Aunque las mujeres de edad entre 15 y 45 años sobreviven mas que los hombres.

Aun asi de media se ve que las mujeres sobreviven mas tiempo que los hombres, aunque la diferencia no es sustancial.