Practical 4: Visualisation using qplot()

Laura Rodriguez Navas

February 2020

Keratoconus is a disorder that affects the cornea through an abnormal growth of collagen fibres. This makes the cornea become conical with an important vision loss. There are many possible treatments, but one common solution is the insertion of intrastromal corneal ring segments, such that the cornea is flattened.

TODO insert img in LaTeX

Dataset

The file "queratocono.csv" includes information about 394 patients with Keratoconus who were treated with ring placement. The variables that were recorded are:

1. K1: keratometry or main corneal curvature.
2. K2: perpendicular curvature to K1.
3. Ch: corneal hysteresis.
4. Na: number of rings (1 or 2).
5. Incision: angle in which the cornea is cut.
6. Prof: depth of the incision.
7. Diam: diameter of the incision.
8. Grosor: Incision thickness.
9. Longitud1: Angle of placement of the first ring (surgical parameter).
10. Longitud2: Angle of placement of the second ring (surgical parameter).
11. grosor1: Thickness of the first ring.
12. grosor2: Thickness of the second ring.
13. long1: arc length of the first ring.
14. long2: arc length of the second ring.
15. K1.salida: keratometry or main corneal curvature after the placement of the ring(s).
16. Astig: astigmatism curvature after the placement of the ring(s) (K1.salida – K2.salida).

A continuation, is provided a sample of the content of each variable of dataset. It is also checked that there are no NA values and the dataset is ordered by the column na.

```r
str(queratocono)
```

```
## 'data.frame':    394 obs. of  16 variables:
##  $ K1       : num  38.1 39.4 43.3 45.7 44.6 44.2 44.2 53.1 45.1 40.7 ...
##  $ K2       : num  48.6 53.6 50.4 50.1 50.4 44.8 45.8 53.8 52.5 44.5 ...
##  $ ch       : num  6.9 5.7 8.8 11.1 11.1 8.9 11.1 7.5 7.7 9.4 ...
##  $ na       : int  2 2 2 1 2 1 1 1 2 1 ...
##  $ Incision : int  60 140 120 30 60 180 130 60 60 30 ...
##  $ Prof     : int  390 400 381 370 440 327 380 400 380 477 ...
##  $ diam     : int  5 5 5 5 5 6 6 5 6 5 ...
##  $ grosor   : int  300 300 200 250 200 200 200 200 225 200 ...
##  $ Longitud1: int  90 105 120 160 120 150 210 160 120 160 ...
##  $ Longitud2: int  180 210 240 160 240 150 210 160 240 160 ...
```

```
##  $ grosor1  : int  300 300 200 250 200 200 200 200 200 200 ...
##  $ grosor2  : int  300 300 200 0 200 0 0 0 250 0 ...
##  $ long1    : int  90 90 120 160 120 150 210 160 90 160 ...
##  $ long2    : int  90 120 120 0 120 0 0 0 150 0 ...
##  $ K1.salida: num  35.5 42 42.7 45.8 44.4 42.8 43.1 48.7 45.7 39.8 ...
##  $ Astig    : num  10.1 8.9 3 1.9 2.9 2.3 3.9 2.8 3.2 3.8 ...
```

```r
any(is.na(queratocono))
```

```
## [1] FALSE
```

```r
queratocono <- queratocono[order(queratocono$na), ]
str(queratocono)
```

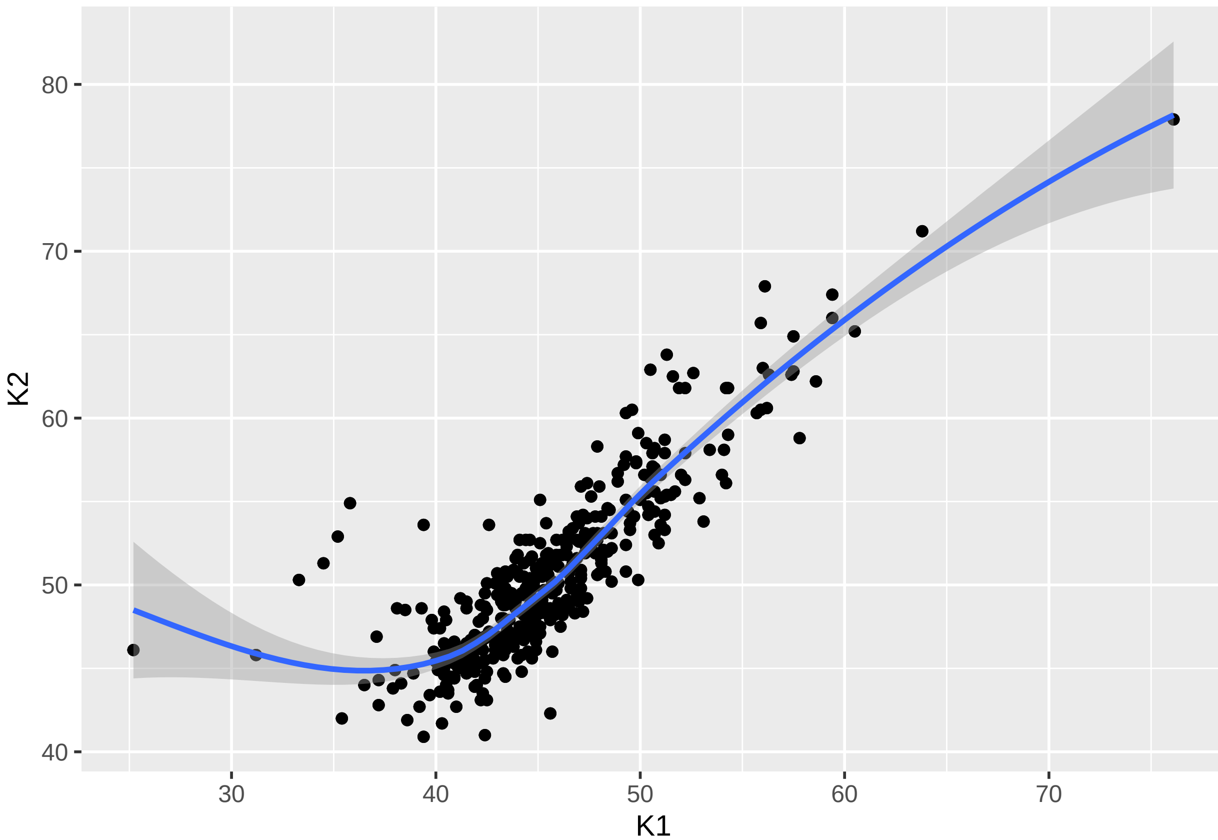
```
## 'data.frame':    394 obs. of  16 variables:
##  $ K1       : num  45.7 44.2 44.2 53.1 40.7 35.4 51.5 40.6 49.5 44.9 ...
##  $ K2       : num  50.1 44.8 45.8 53.8 44.5 42 55.4 45.4 53.7 47.5 ...
##  $ ch       : num  11.1 8.9 11.1 7.5 9.4 8.2 10.1 8.6 7.2 7.6 ...
##  $ na       : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Incision : int  30 180 130 60 30 170 165 30 180 60 ...
##  $ Prof     : int  370 327 380 400 477 420 400 400 434 367 ...
##  $ diam     : int  5 6 6 5 5 5 6 6 5 5 ...
##  $ grosor   : int  250 200 200 200 200 250 200 250 200 200 ...
##  $ Longitud1: int  160 150 210 160 160 160 150 150 210 160 ...
##  $ Longitud2: int  160 150 210 160 160 160 150 150 210 160 ...
##  $ grosor1  : int  250 200 200 200 200 250 200 250 200 200 ...
##  $ grosor2  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ long1    : int  160 150 210 160 160 160 150 150 210 160 ...
##  $ long2    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ K1.salida: num  45.8 42.8 43.1 48.7 39.8 35.6 48.7 42.7 47.4 43.7 ...
##  $ Astig    : num  1.9 2.3 3.9 2.8 3.8 4 5.3 1.2 0.3 1.8 ...
```

Analysis

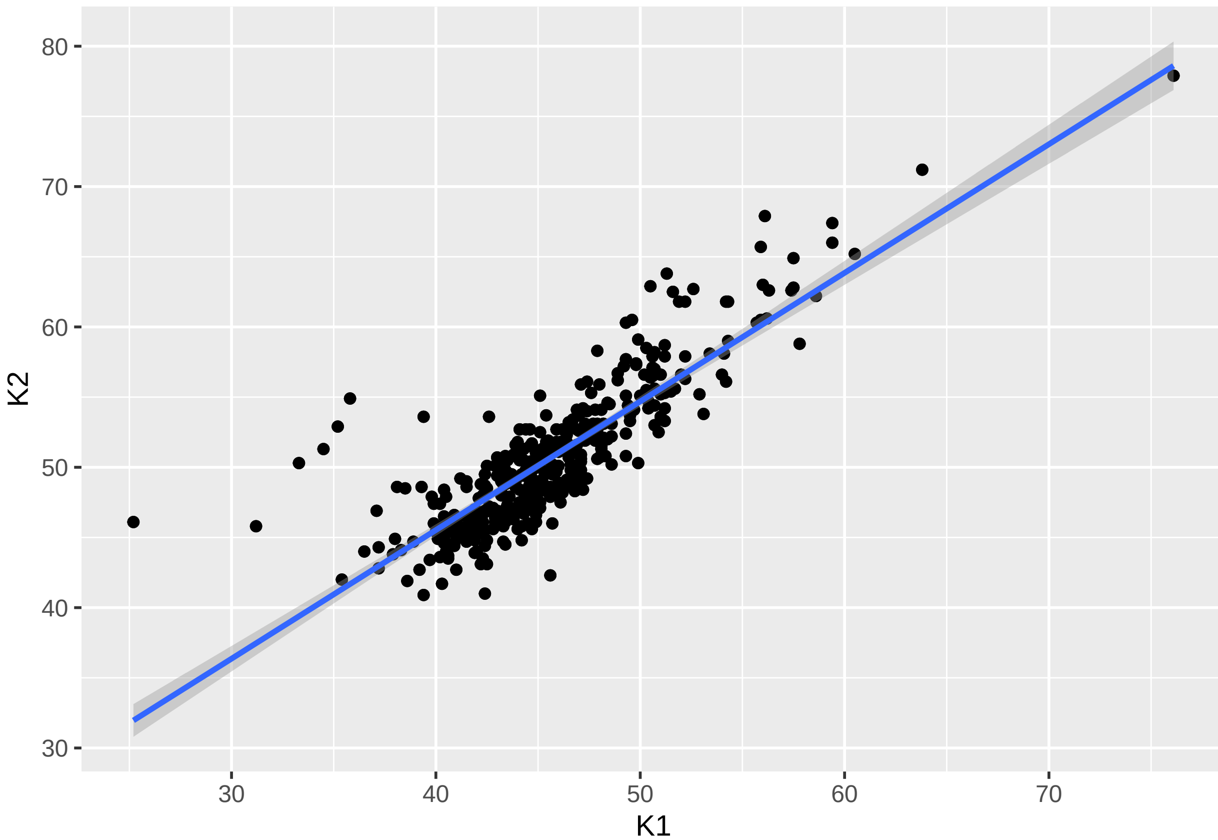
In order to analyse the information in a visual way:

1. Study the relation between K1 and K2 with smoother (by default and using linear regression).

```r
qplot(K1, K2, data = queratocono) +
  geom_point() +
  geom_smooth(method = "loess") +
  xlab("K1") + ylab("K2")
```

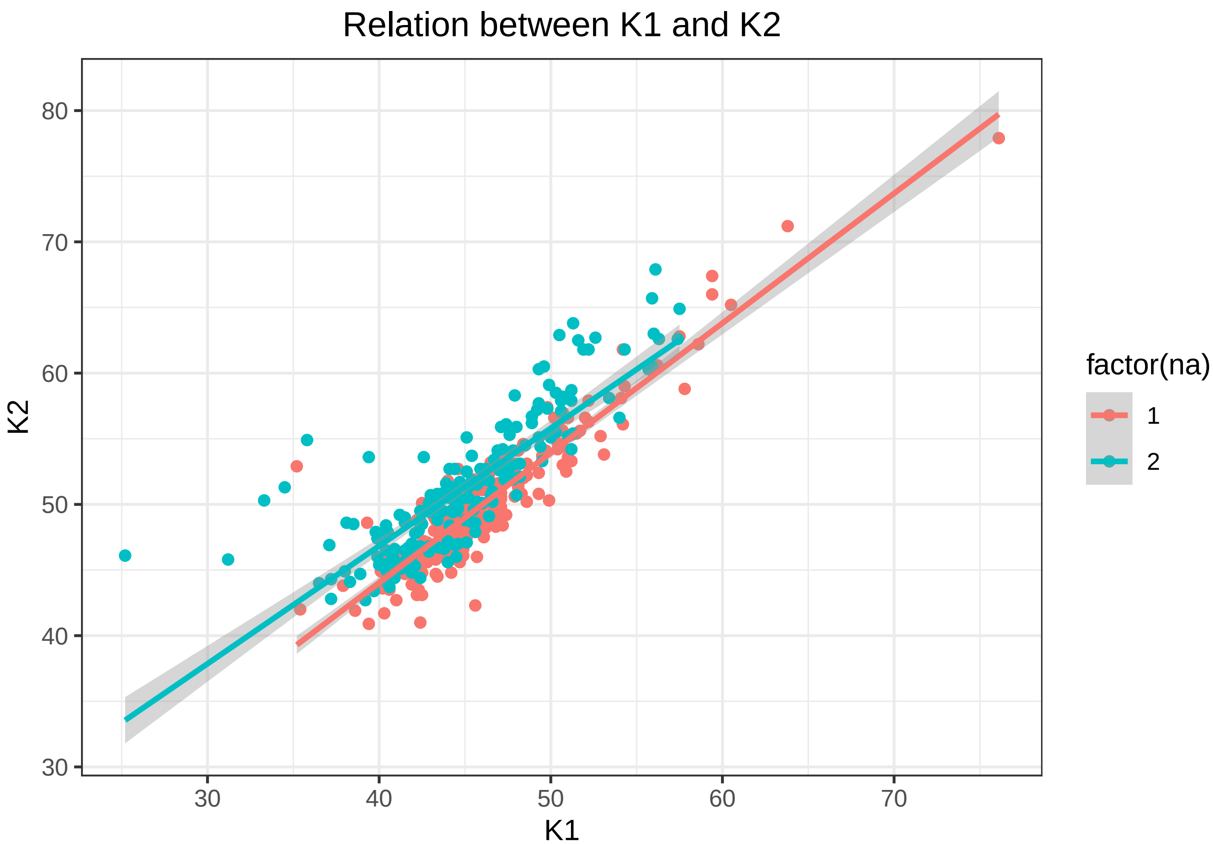



```
qplot(K1, K2, data = queratocono) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("K1") + ylab("K2")
```



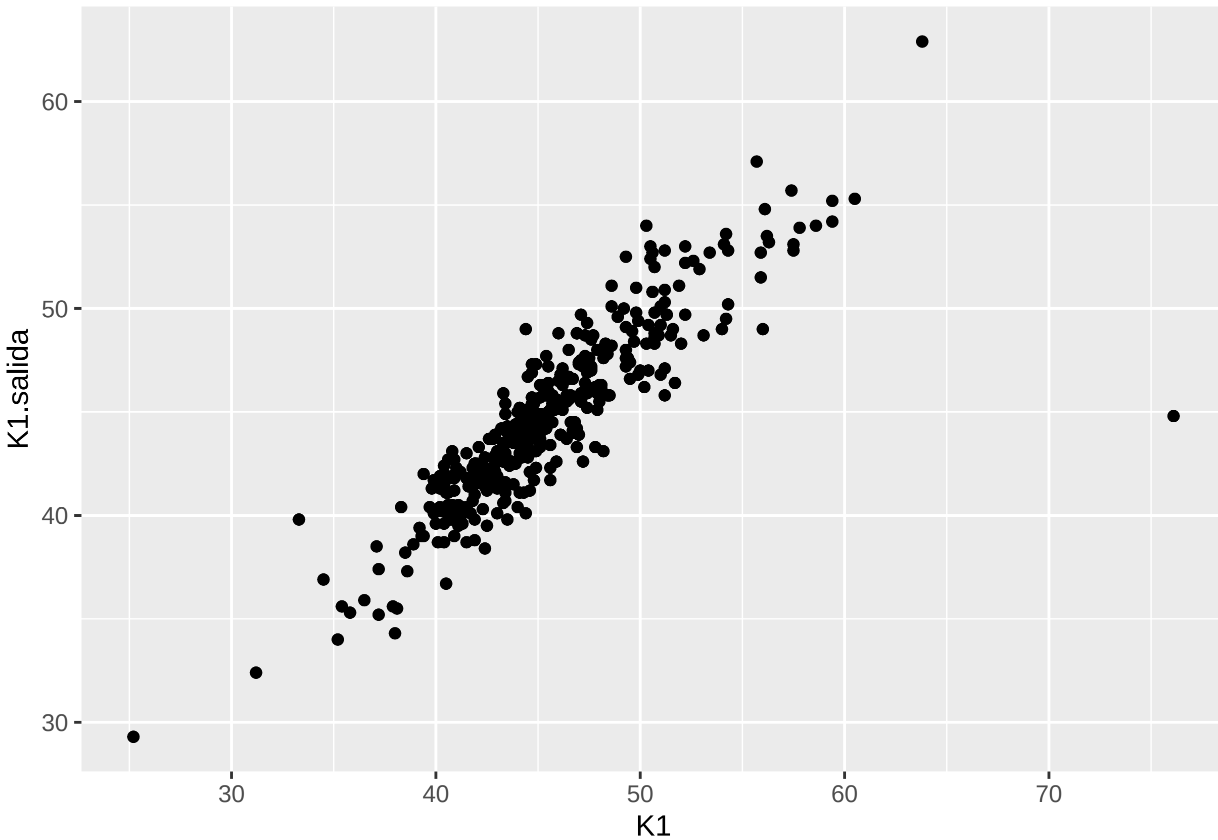
2. Study the relation between K1 and K2 distinguishing by factor na.

```
qplot(K1, K2, data = queratocono, colour = factor(na)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("K1") + ylab("K2") +
  ggtitle("Relation between K1 and K2") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```



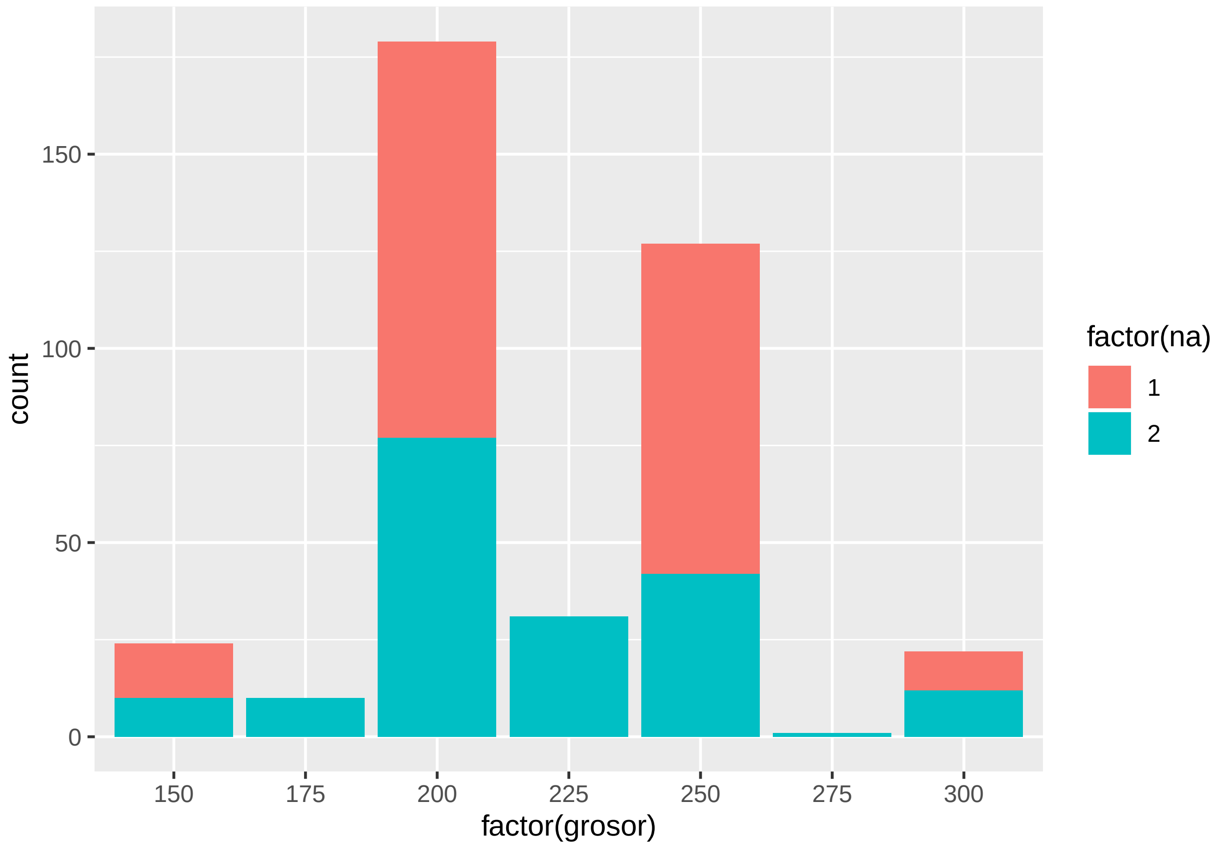
3. Study the relation between K1 and K1.salida.

```
qplot(K1, K1.salida, data = queratocono) +
  geom_point() +
  xlab("K1") + ylab("K1.salida")
```



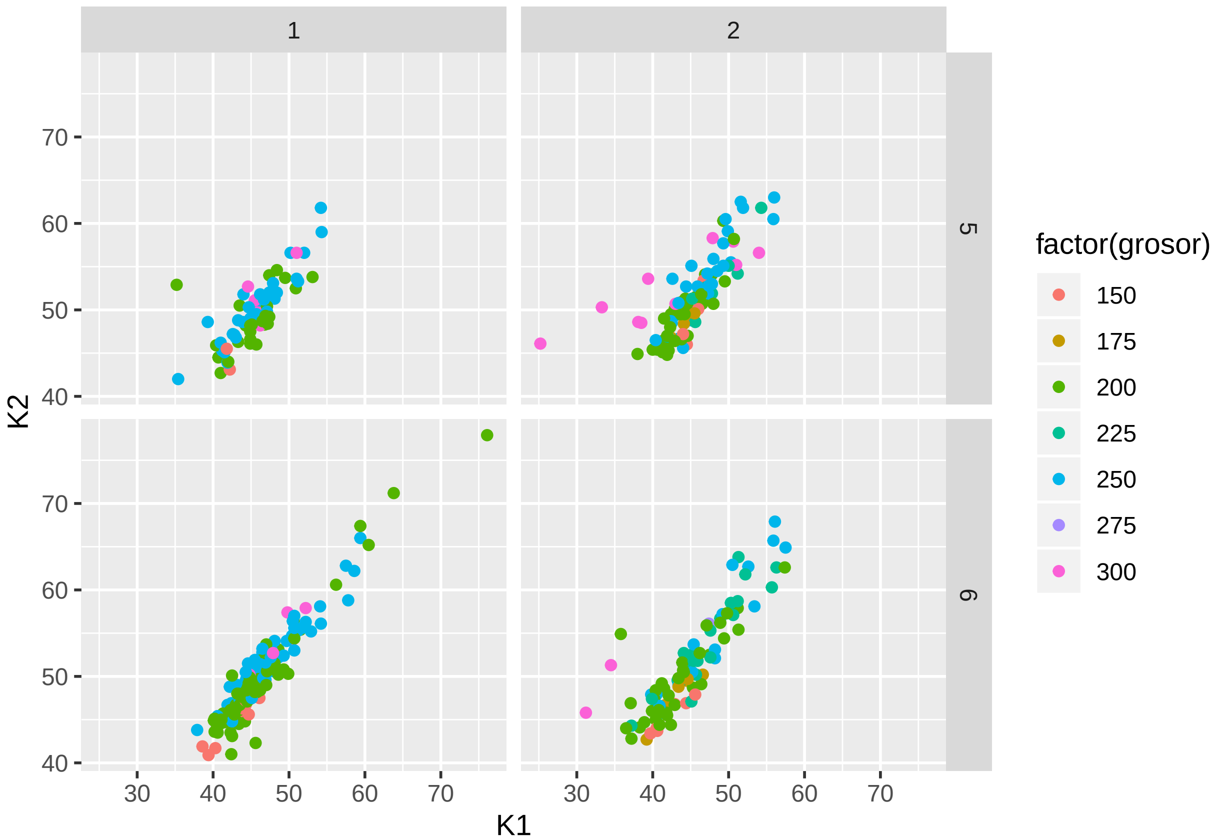
4. Build a histogram in terms of grosor (note that grosor should be taken as a factor) of the inserted ring.

```
ggplot(queratocono, aes(factor(grosor), geom = "bar" , fill = factor(na))) +
  geom_bar()
```



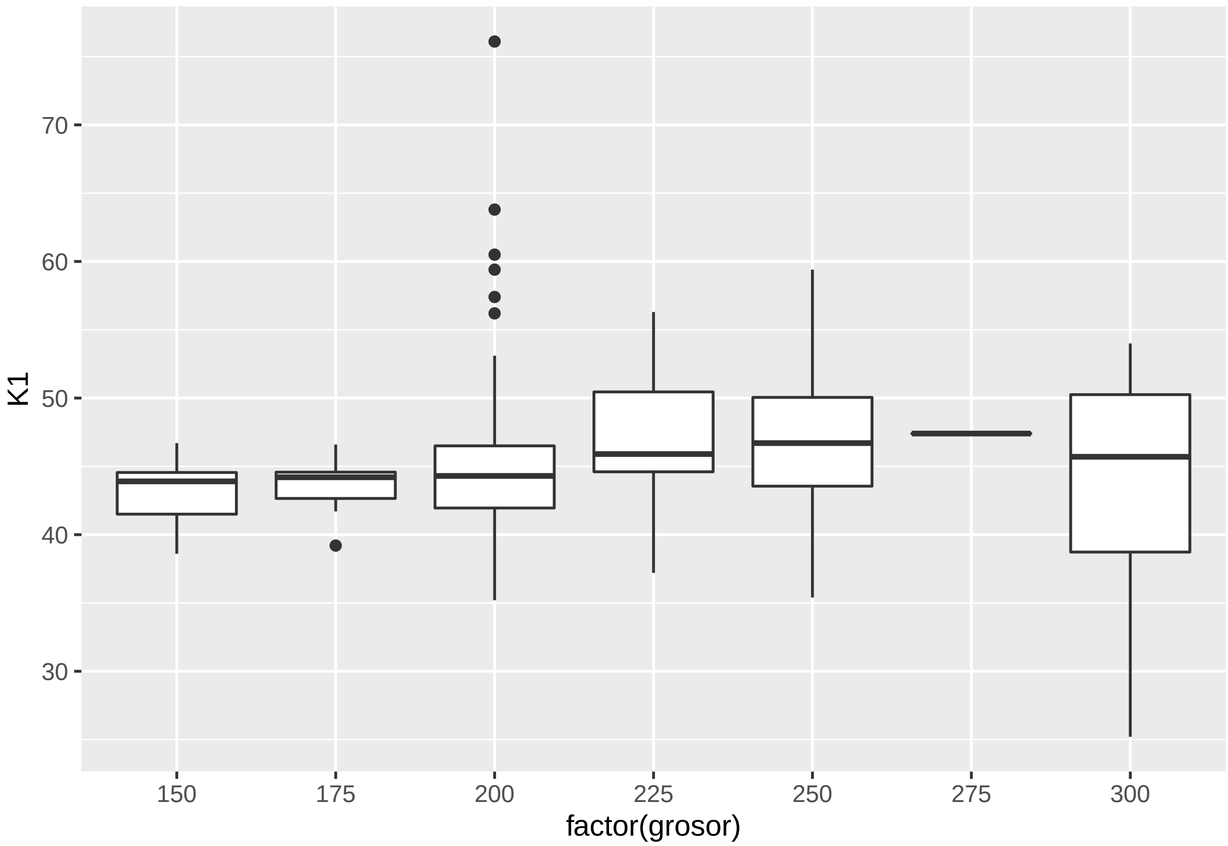
5. Build a scatter plot of the relation between K1 and K2 with "faceting" in terms of the parameters diam and na, by assigning different colours to the points according to the thickness (grosor) of the ring. In order to visualise all points correctly use a transparency of value 1/3.

```
qplot(K1, K2, data = queratocono, colour = factor(grosor), facets = diam ~ na,
      size = I(1/3)) +
      geom_point() +
      scale_shape_manual(values = 0:7) +
      xlab("K1") + ylab("K2")
```

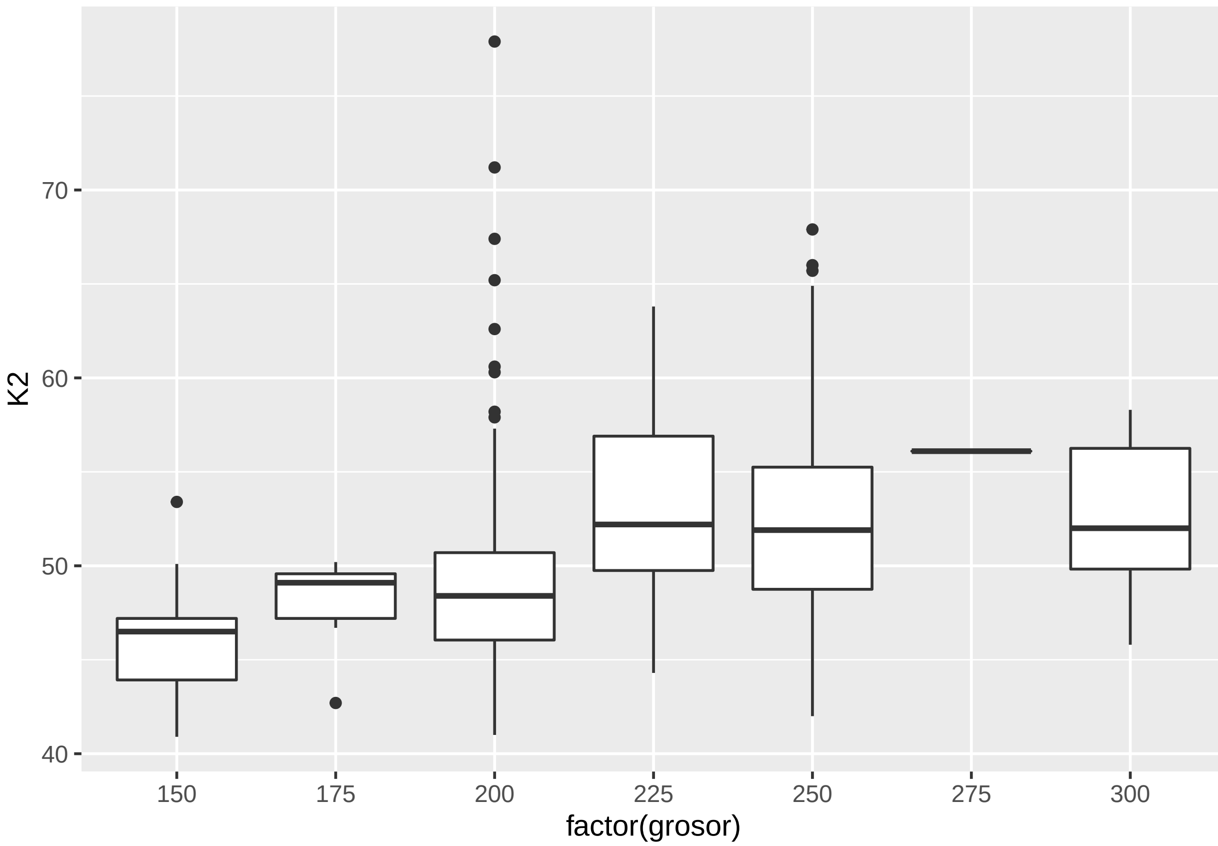


6. Create two boxplots that show a summary of the distributions of K1 and K2 (separately) with respect to the thickness (grosor).

```
qplot(factor(grosor), K1, data = queratocono, geom = "boxplot") +
  xlab("factor(grosor)") + ylab("K1")
```




```
qplot(factor(grosor), K2, data = queratocono, geom = "boxplot") +
  xlab("factor(grosor)") + ylab("K2")
```



K2
70
60
50
40
150
175
200
225
250
275
300
factor(grosor)