Índice general

Capítulo 1

Resumen Descriptivo: Parte 1

1.1. Lectura de Datos

```
data1 <- uno
dim(data1)
```

```
## [1] 25  6
```

```
knitr::kable(head(data1))
```

dradio	radio	dhumero	humero	dcubito	cubito
1.103	1.052	2.139	2.238	0.873	0.872
0.842	0.859	1.873	1.741	0.590	0.744
0.925	0.873	1.887	1.809	0.767	0.713
0.857	0.744	1.739	1.547	0.706	0.674
0.795	0.809	1.734	1.715	0.549	0.654
0.787	0.779	1.509	1.474	0.782	0.571

```
knitr::kable(data1)
```

dradio	radio	dhumero	humero	dcubito	cubito
1.103	1.052	2.139	2.238	0.873	0.872
0.842	0.859	1.873	1.741	0.590	0.744
0.925	0.873	1.887	1.809	0.767	0.713
0.857	0.744	1.739	1.547	0.706	0.674
0.795	0.809	1.734	1.715	0.549	0.654
0.787	0.779	1.509	1.474	0.782	0.571
0.933	0.880	1.695	1.656	0.737	0.803
0.799	0.851	1.740	1.777	0.618	0.682

dradio	radio	dhumero	humero	dcubito	cubito
0.945	0.876	1.811	1.759	0.853	0.777
0.921	0.906	1.954	2.009	0.823	0.765
0.792	0.825	1.624	1.657	0.686	0.668
0.815	0.751	2.204	1.846	0.678	0.546
0.755	0.724	1.508	1.458	0.662	0.595
0.880	0.866	1.786	1.811	0.810	0.819
0.900	0.838	1.902	1.606	0.723	0.677
0.764	0.757	1.743	1.794	0.586	0.541
0.733	0.748	1.863	1.869	0.672	0.752
0.932	0.898	2.028	2.032	0.836	0.805
0.856	0.786	1.390	1.324	0.578	0.610
0.890	0.950	2.187	2.087	0.758	0.718
0.688	0.532	1.650	1.378	0.533	0.482
0.940	0.850	2.334	2.225	0.757	0.731
0.493	0.616	1.037	1.268	0.546	0.615
0.835	0.752	1.509	1.422	0.618	0.664
0.915	0.936	1.971	1.869	0.869	0.868

```
sapply(data1,mean,na.rm=TRUE)
```

```
##  dradio   radio dhumero  humero dcubito  cubito
## 0.84380 0.81832 1.79268 1.73484 0.70440 0.69384
```

```
fivenum(data1[,1])       #estadisticos de tukey:min,q1,median,q3,max
```

```
## [1] 0.493 0.792 0.856 0.921 1.103
```

```
library(Hmisc)
describe(data1[,1])      #fucntion describe de Hmisc
```

```
## data1[, 1]
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##       25        0       25        1   0.8438   0.1216   0.6970   0.7418
##      .25      .50      .75      .90      .95
##   0.7920   0.8560   0.9210   0.9372   0.9440
##
## lowest : 0.493 0.688 0.733 0.755 0.764, highest: 0.932 0.933 0.940 0.945 1.103
```

```
library(pastecs)
stat.desc(data1[,1])     #fucntion stat,desc de pstecs
```

```
##      nbr.val     nbr.null       nbr.na          min          max        range
##  25.00000000   0.00000000   0.00000000   0.49300000   1.10300000   0.61000000
```

```
##          sum       median         mean      SE.mean CI.mean.0.95          var
##  21.09500000   0.85600000   0.84380000   0.02280490   0.04706699   0.01300158
##      std.dev     coef.var
##   0.11402449   0.13513212
```

```
library(psych)
describe(data1[,1:2])      #fucntion describe de psych
```

```
##        vars  n mean   sd median trimmed  mad  min  max range  skew kurtosis
## dradio    1 25 0.84 0.11   0.86    0.85 0.10 0.49 1.10  0.61 -0.75     1.96
## radio     2 25 0.82 0.11   0.84    0.82 0.09 0.53 1.05  0.52 -0.51     0.75
##          se
## dradio 0.02
## radio  0.02
```

```
library(knitr)
library(data.table)
library(pastecs)
tablita<-stat.desc(data1[,1:2])
tablita1<-data.table(Estadisticos=rownames(tablita),tablita)
kable(tablita,caption="Resumen de Datos",digits=3)
```

Tabla 1.3: Resumen de Datos

	dradio	radio
nbr.val	25.000	25.000
nbr.null	0.000	0.000
nbr.na	0.000	0.000
min	0.493	0.532
max	1.103	1.052
range	0.610	0.520
sum	21.095	20.458
median	0.856	0.838
mean	0.844	0.818
SE.mean	0.023	0.021
CI.mean.0.95	0.047	0.044
var	0.013	0.011
std.dev	0.114	0.107
coef.var	0.135	0.131

```
kable(tablita1,caption="Resumen de Datos1",digits=4)
```

Tabla 1.4: Resumen de Datos1

Estadisticos	dradio	radio
nbr.val	25.0000	25.0000
nbr.null	0.0000	0.0000
nbr.na	0.0000	0.0000
min	0.4930	0.5320
max	1.1030	1.0520
range	0.6100	0.5200
sum	21.0950	20.4580
median	0.8560	0.8380
mean	0.8438	0.8183
SE.mean	0.0228	0.0214
CI.mean.0.95	0.0471	0.0441
var	0.0130	0.0114
std.dev	0.1140	0.1069
coef.var	0.1351	0.1306

```
library(stargazer)
stargazer(data1[,1:2],header=FALSE,title = "Resumen de datos-3",digits=4)
```

Tabla 1.5: Resumen de datos-3

Statistic	N	Mean	St. Dev.	Min	Max
dradio	25	0.8438	0.1140	0.4930	1.1030
radio	25	0.8183	0.1069	0.5320	1.0520

```
library(reporttools)                #para cruzar continuasxcategoricas
cate<-factor(rep(c("A","B","C","D"),c(5,5,5,10)))
data2<-cbind(data1,cate)
titulo<-"Resumen de lso datos-5"
vbles<-data2[,1:3]
estadisticas<-list("mean","s",
                   "c$_{\\mathrm{v}}\\%$"=
                     function(x){
                     return((sd(x)/mean(x))*100)
                     },
                   "min","median","max","n", "na"
                     )
#Cruece da vbles_cont con vbles_nominal
tableContinuous(vars = vbles, group = data2[,7], stats = estadisticas,cap=titulo,
                print.pval = "kruskal",longtable=FALSE)
```

% latex table generated in R 4.1.3 by xtable 1.8-4 package % Mon Mar 14 14:50:53 2022

Variable	Levels	$\bar{x}$	s	c_v %	Min	$\tilde{x}$	Max	n	#NA
dradio	A	0.9	0.1	13.3	0.8	0.9	1.1	5	0
	B	0.9	0.1	8.8	0.8	0.9	0.9	5	0
	C	0.8	0.1	7.3	0.8	0.8	0.9	5	0
	D	0.8	0.1	17.3	0.5	0.8	0.9	10	0
$p = 0.45$	all	0.8	0.1	13.5	0.5	0.9	1.1	25	0
radio	A	0.9	0.1	13.2	0.7	0.9	1.1	5	0
	B	0.9	0.0	5.6	0.8	0.9	0.9	5	0
	C	0.8	0.1	7.5	0.7	0.8	0.9	5	0
	D	0.8	0.1	17.2	0.5	0.8	0.9	10	0
$p = 0.40$	all	0.8	0.1	13.1	0.5	0.8	1.1	25	0
dhumero	A	1.9	0.2	8.8	1.7	1.9	2.1	5	0
	B	1.7	0.2	9.4	1.5	1.7	2.0	5	0
	C	1.8	0.3	14.9	1.5	1.8	2.2	5	0
	D	1.8	0.4	22.0	1.0	1.8	2.3	10	0
$p = 0.90$	all	1.8	0.3	15.8	1.0	1.8	2.3	25	0

Tabla 1.6: Resumen de lso datos-5

```
library(reporttools)                #para cruzar continuasxcategoricas
cate<-factor(rep(c("A","B","C","D"),c(5,5,5,10)))
cate1<-factor(rep(c("A","B","C","D"),c(10,5,5,5)))
data2<-cbind(data1,cate,cate1)
# solo tabla de vbles_nominales
titulo2<-"Caracteristicas de las variables nominales"
tableNominal(vars = data2[ ,7:8],cap=titulo2,vertical = FALSE,font.size = "scriptsize
             lab = "tab:nominal1",longtable = FALSE,cumsum = TRUE)
```

% latex table generated in R 4.1.3 by xtable 1.8-4 package % Mon Mar 14 14:50:53 2022

Variable	Levels	n	%	$\sum$ %
cate	A	5	20.0	20.0
	B	5	20.0	40.0
	C	5	20.0	60.0
	D	10	40.0	100.0
	all	25	100.0	
cate1	A	10	40.0	40.0
	B	5	20.0	60.0
	C	5	20.0	80.0
	D	5	20.0	100.0
	all	25	100.0	

Tabla 1.7: Caracteristicas de las variables nominales

1.2. Obtención del vector de media muestral

```
vector_medias = medias<-apply(data1,2,mean);
knitr::kable(round(vector_medias,3))
```

	x
dradio	0.844
radio	0.818
dhumero	1.793
humero	1.735
dcubito	0.704
cubito	0.694

1.3. obtención de la matriz de covarianza muestral

```
var_cov <- cov(data1);
knitr::kable(round(var_cov,3))
```

	dradio	radio	dhumero	humero	dcubito	cubito
dradio	0.013	0.010	0.022	0.020	0.009	0.008
radio	0.010	0.011	0.019	0.021	0.009	0.009
dhumero	0.022	0.019	0.080	0.067	0.017	0.013
humero	0.020	0.021	0.067	0.069	0.018	0.017
dcubito	0.009	0.009	0.017	0.018	0.012	0.008
cubito	0.008	0.009	0.013	0.017	0.008	0.011

1.4. obtención de la matriz de correlación muestral

```
corr <- cor(data1);
knitr::kable(round(corr,3))
```

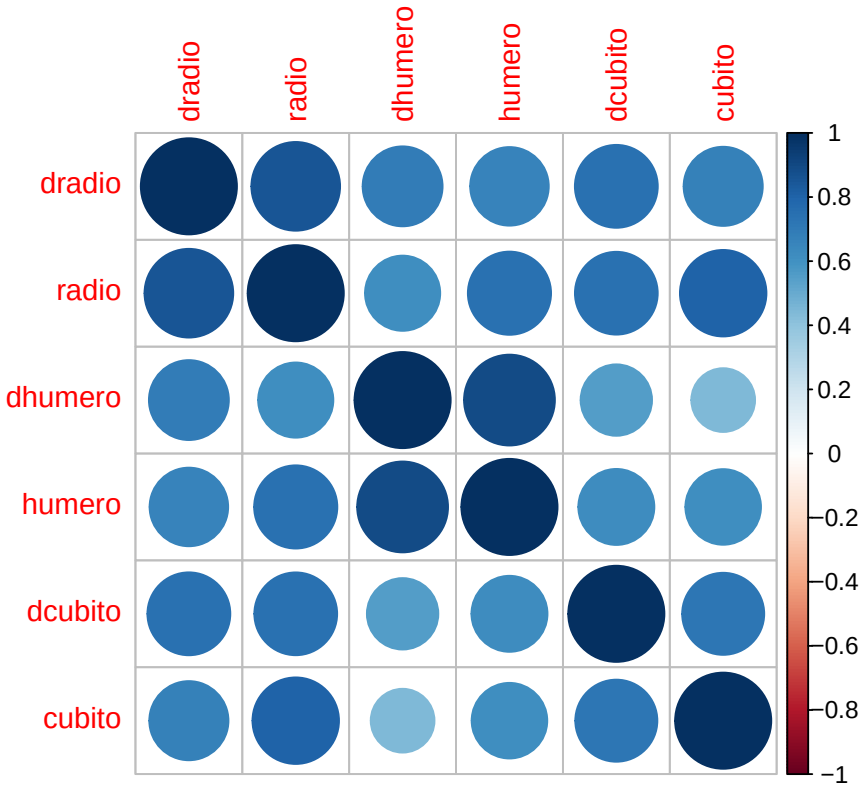
	dradio	radio	dhumero	humero	dcubito	cubito
dradio	1.000	0.852	0.691	0.668	0.744	0.678
radio	0.852	1.000	0.612	0.749	0.742	0.810
dhumero	0.691	0.612	1.000	0.894	0.552	0.440
humero	0.668	0.749	0.894	1.000	0.626	0.619
dcubito	0.744	0.742	0.552	0.626	1.000	0.729
cubito	0.678	0.810	0.440	0.619	0.729	1.000

1.4.1. Gráfica de la matriz de correlación

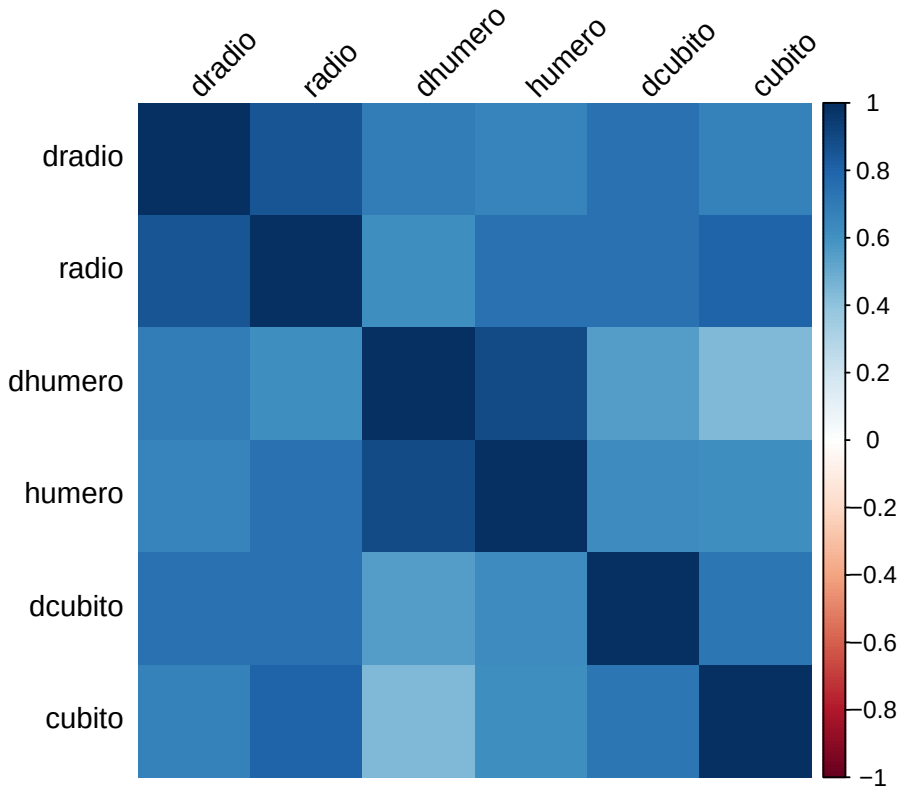
```
library(ggplot2)
library(corrplot)
matcorrs.cor<-cor(data1,method="pearson")
knitr::kable(round(matcorrs.cor,3))
```

	dradio	radio	dhumero	humero	dcubito	cubito
dradio	1.000	0.852	0.691	0.668	0.744	0.678
radio	0.852	1.000	0.612	0.749	0.742	0.810
dhumero	0.691	0.612	1.000	0.894	0.552	0.440
humero	0.668	0.749	0.894	1.000	0.626	0.619
dcubito	0.744	0.742	0.552	0.626	1.000	0.729
cubito	0.678	0.810	0.440	0.619	0.729	1.000

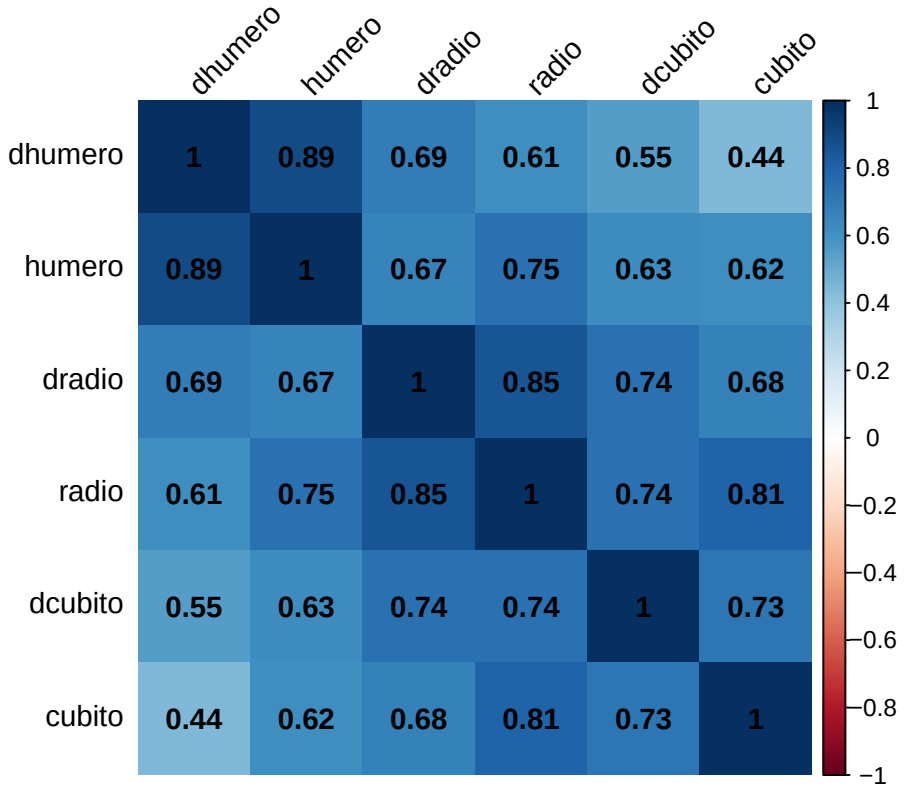
```
# Gráfico de la MAtriz de Correlación por default
corrplot(matcorrs.cor)
```



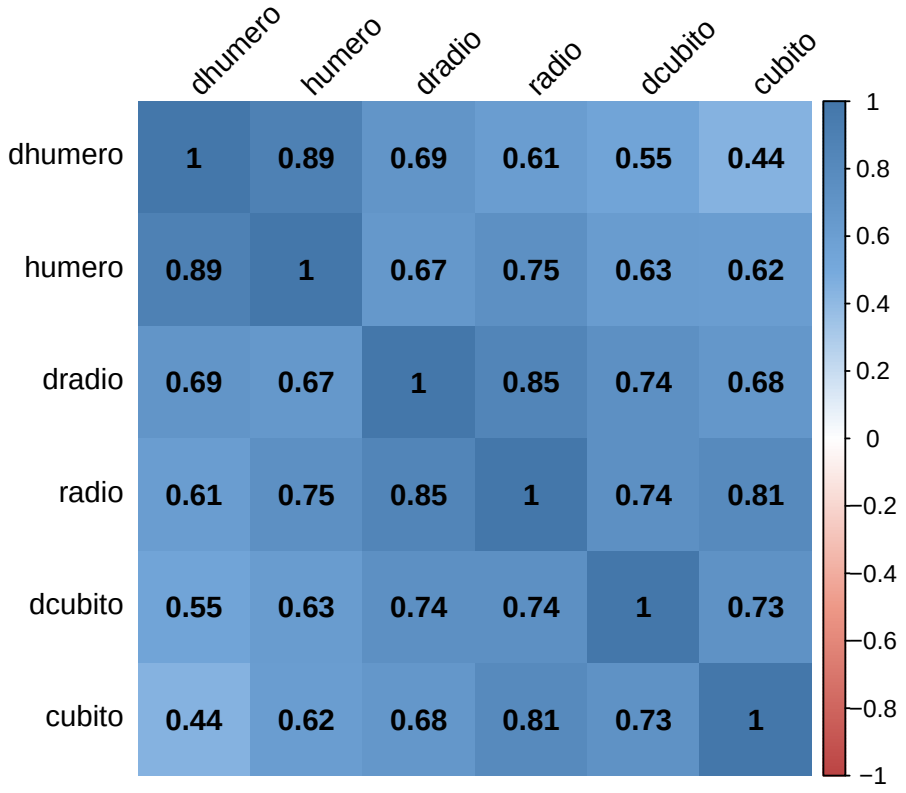

```
# Gráfica de la matriz de correlación
# metodo(forma),color-forma,color-etiquetas,inclinación-etiquetas
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45 )
```



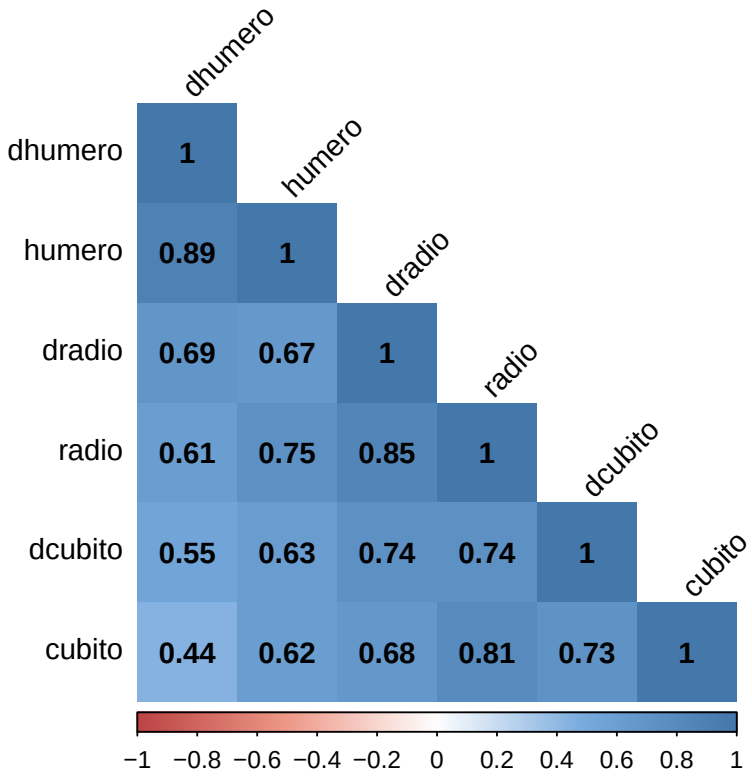

```
# Gráfica de la matriz de correlación
# addCoef.col,order
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",
         tl.srt=45,addCoef.col="black",addcolorlabel="no",order="AOE" )
```



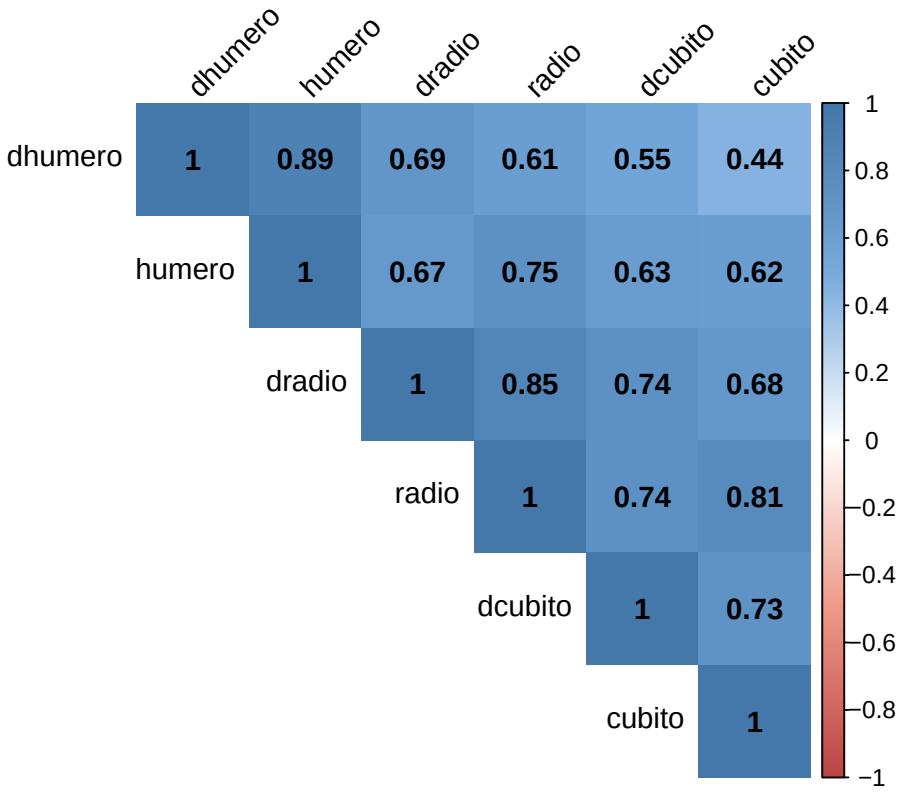
```
# Definir una paleta de colores
col<-colorRampPalette(c("#BB4444","#EE9988","#FFFFFF","#77AADD","#4477AA"))

# Gráfica de la matriz de correlación
# addCoef.col,order
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE" )
```



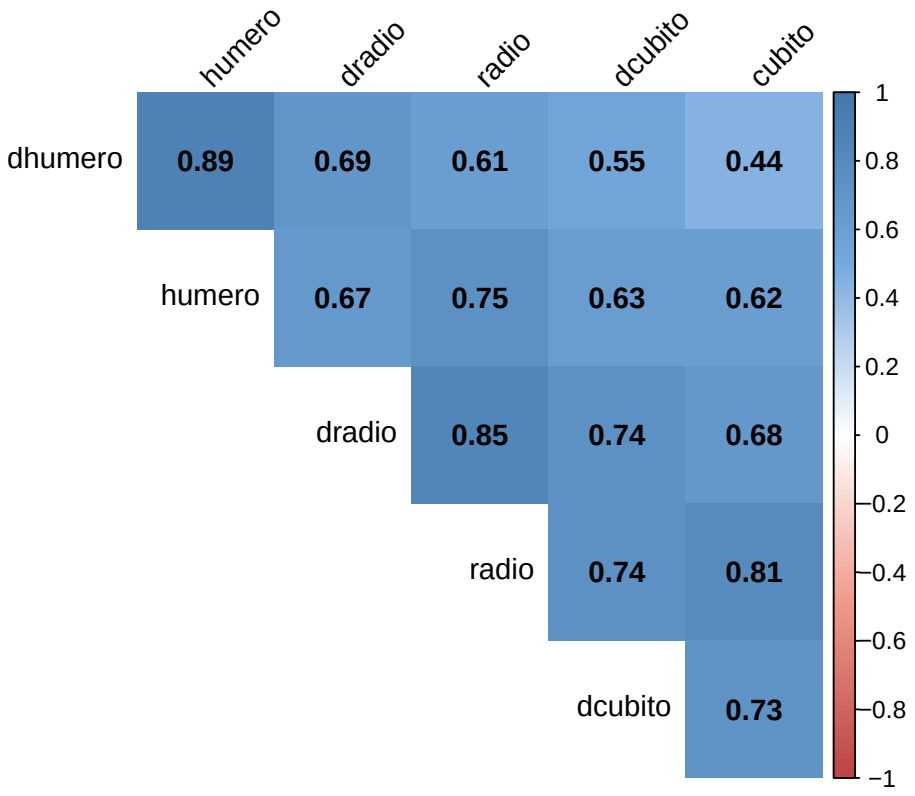

```
# Gráfica Triagular
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="lower" )
```



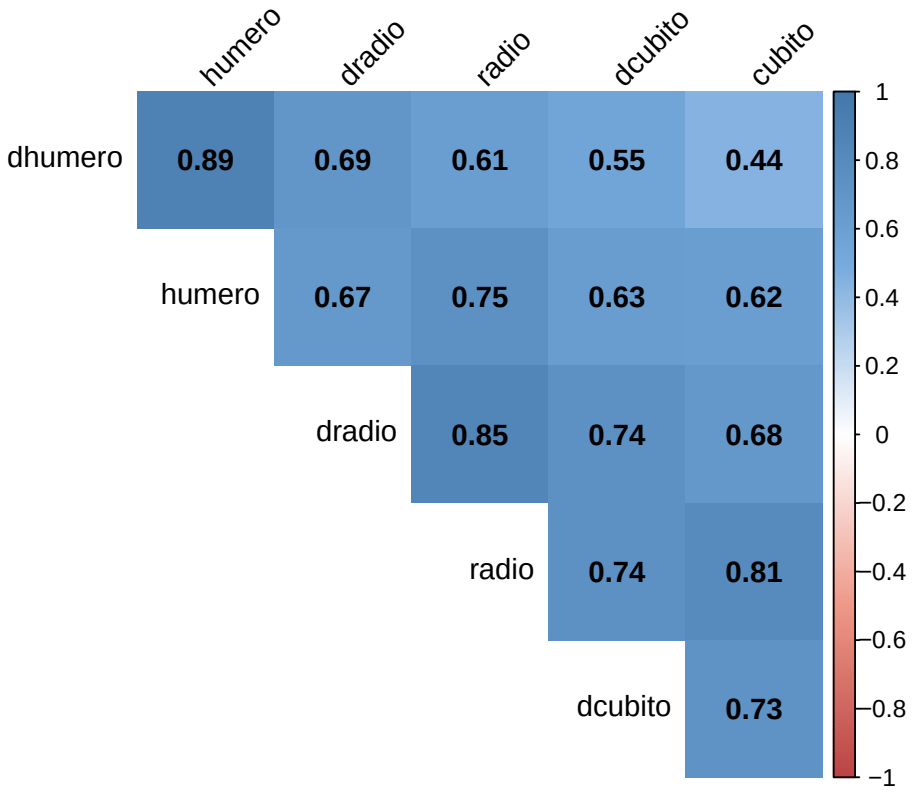

```
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper" )
```



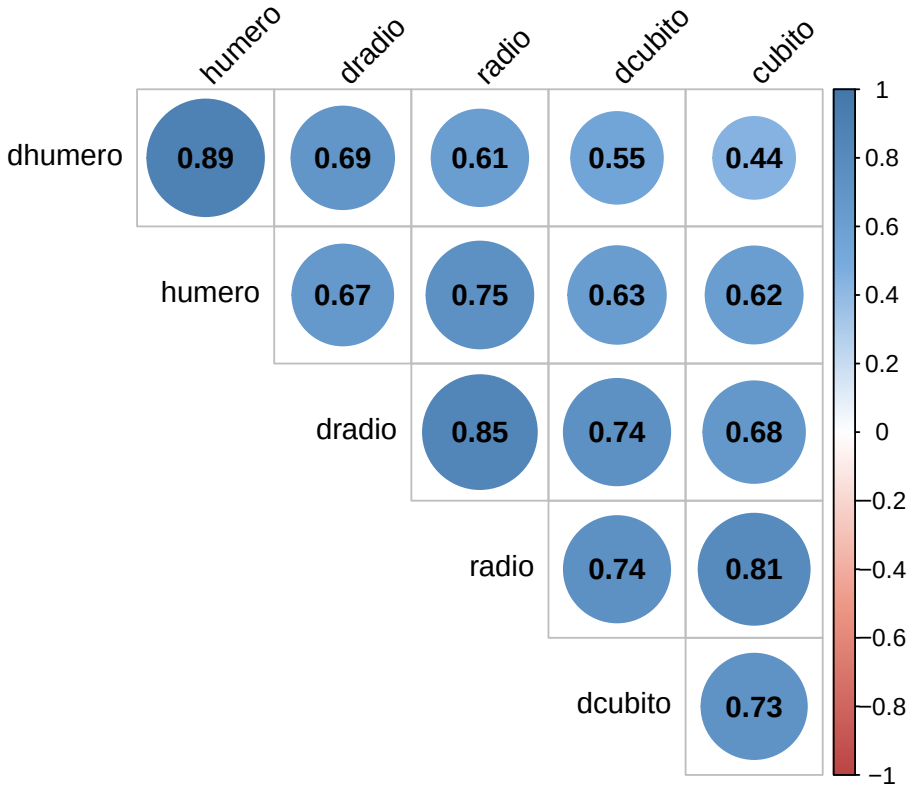
```
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F )
```



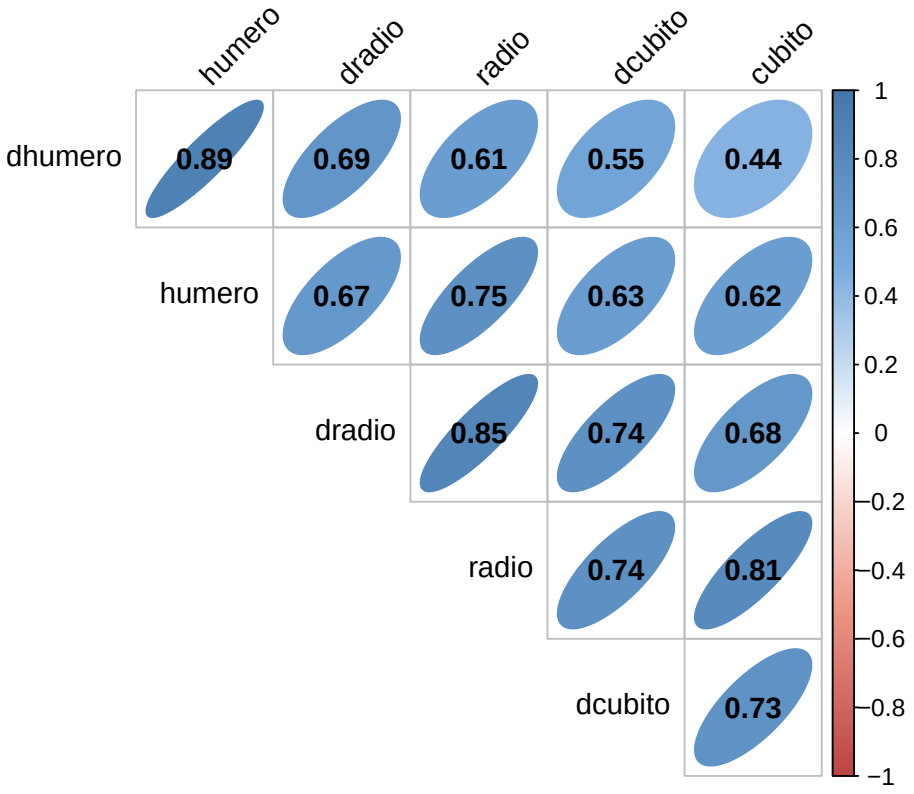

```
corrplot(matcorrs.cor,method="shade",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```



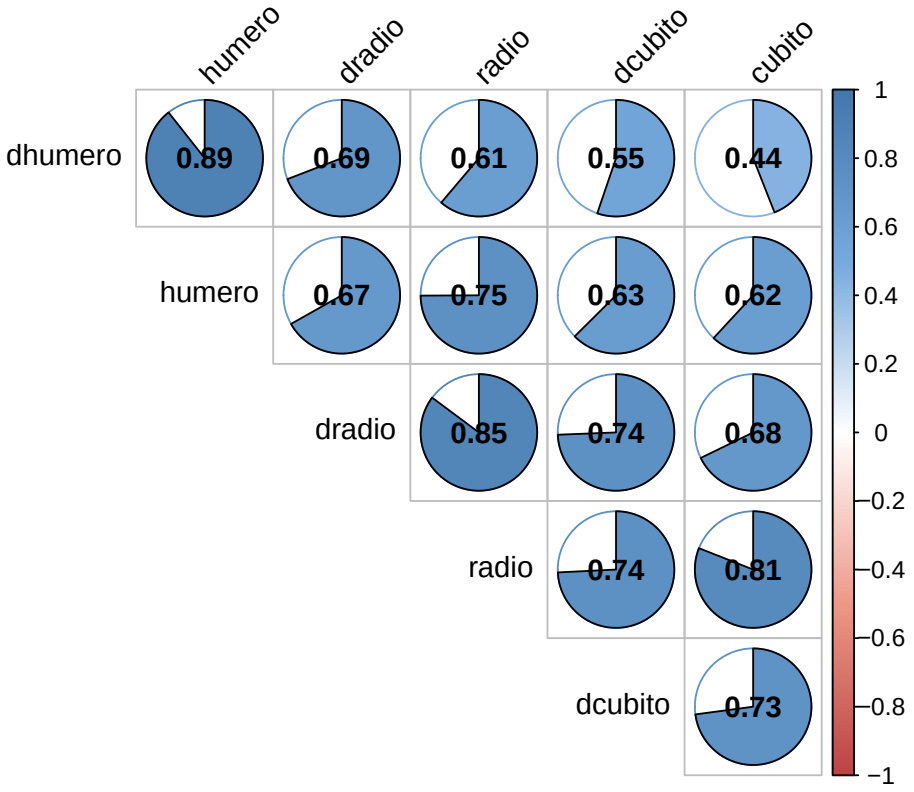

```
corrplot(matcorrs.cor,method="circle",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```



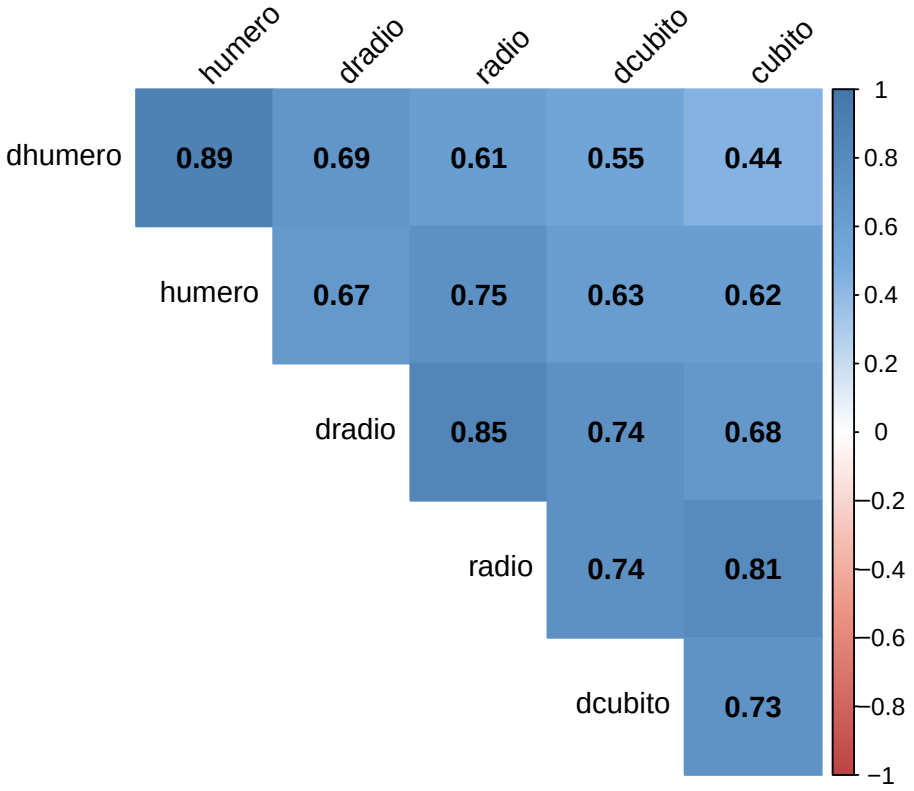
```
corrplot(matcorrs.cor,method="ellipse",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```



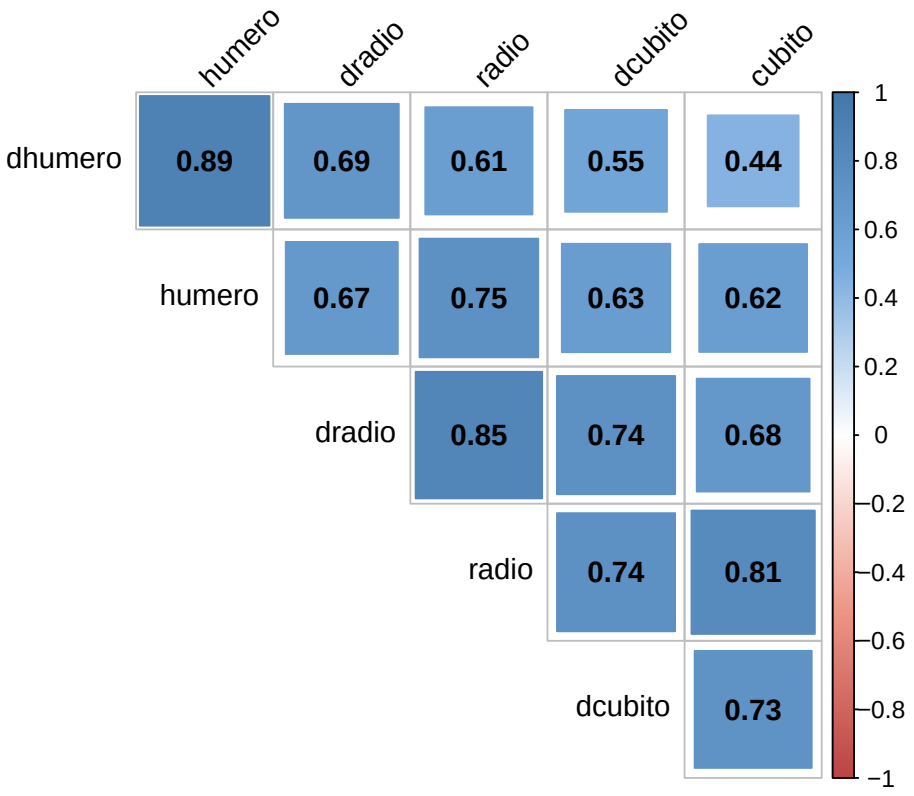

```
corrplot(matcorrs.cor,method="pie",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```



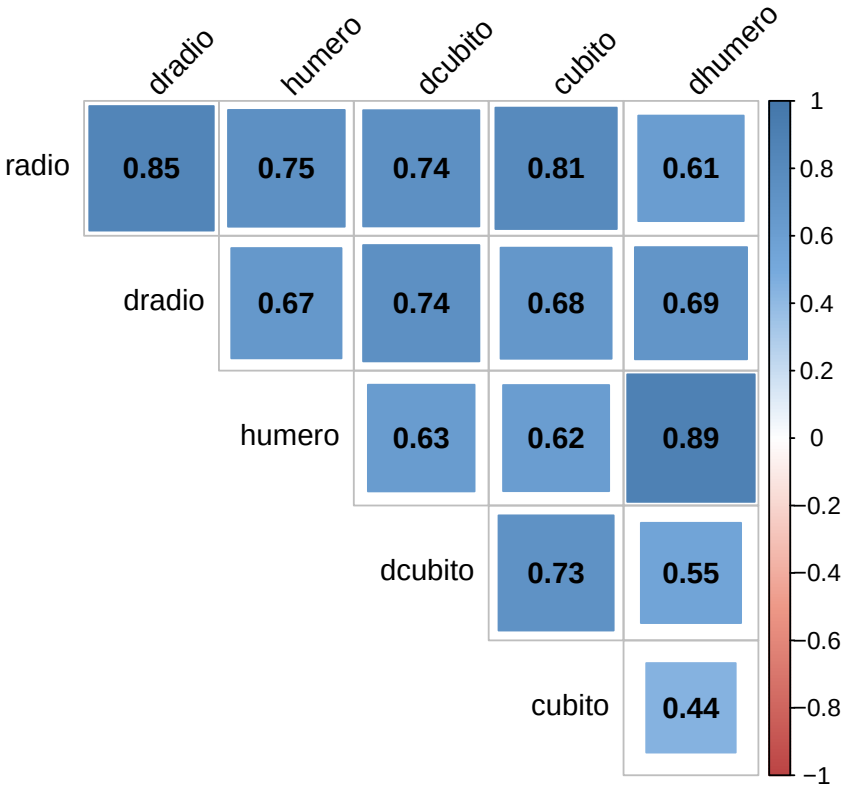

```
corrplot(matcorrs.cor,method="color",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```



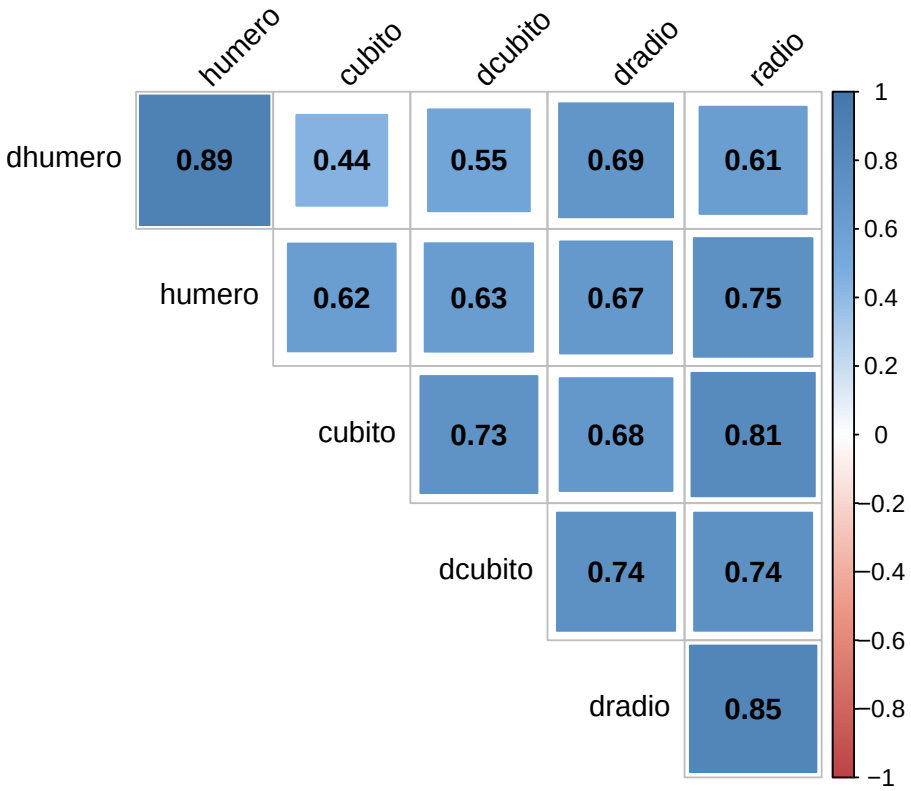
```
corrplot(matcorrs.cor,method="square",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="AOE",type="upper",
diag=F,addshade="all" )
```




```
corrplot(matcorrs.cor,method="square",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="FPC",type="upper",
diag=F,addshade="all" )
```



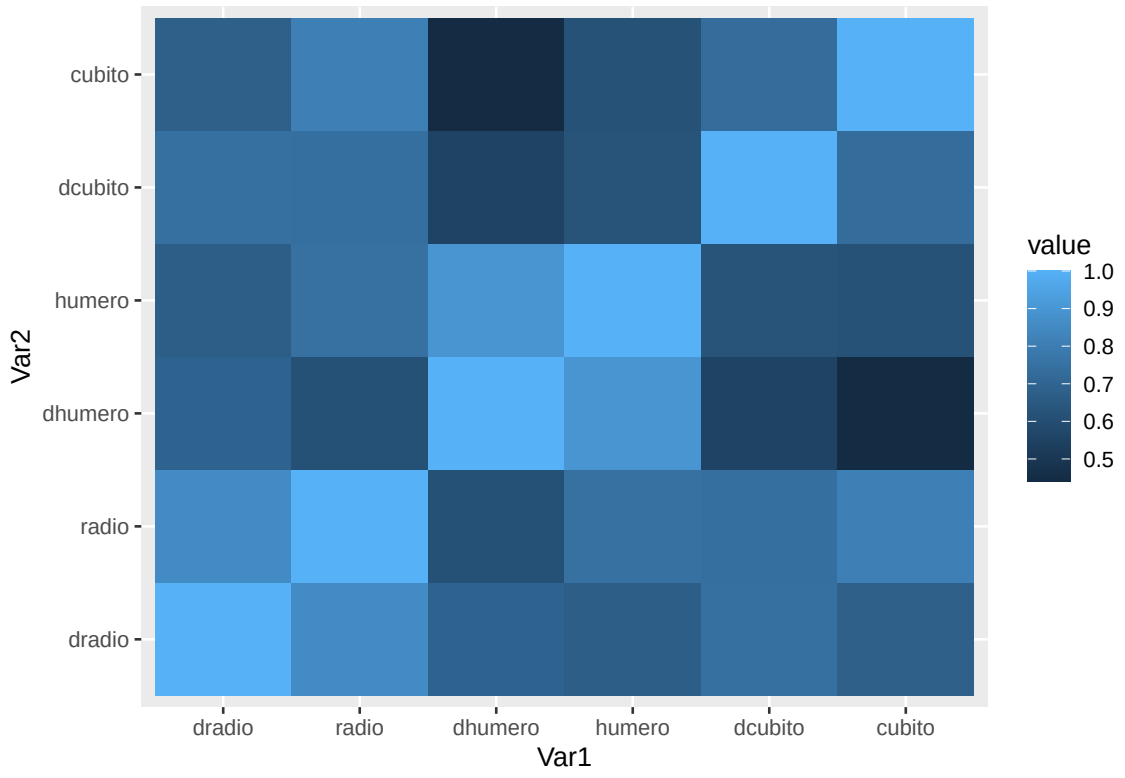

```
corrplot(matcorrs.cor,method="square",shade.col=NA,tl.col="black",tl.srt=45,
col=col(200),addCoef.col="black",addcolorlabel="no",order="hclust",type="upper",diag=
```



```
## Para fundir la matriz de correlación (la matriz de corr en forma de un data.frma

library(reshape2)
##head(matcorrs.cor)
matcorrs.melt<-melt(matcorrs.cor)
head(matcorrs.melt)
```

```
##      Var1   Var2     value
## 1  dradio dradio 1.0000000
## 2   radio dradio 0.8518067
## 3 dhumero dradio 0.6914590
## 4  humero dradio 0.6682584
## 5 dcubito dradio 0.7436926
## 6  cubito dradio 0.6778941
```

```
ggplot(data=matcorrs.melt, aes(x=Var1 , y=Var2, fill=value))+
   geom_tile()
```



1.5. obtención del coeficiente de asimetría muestral

x
-0.511

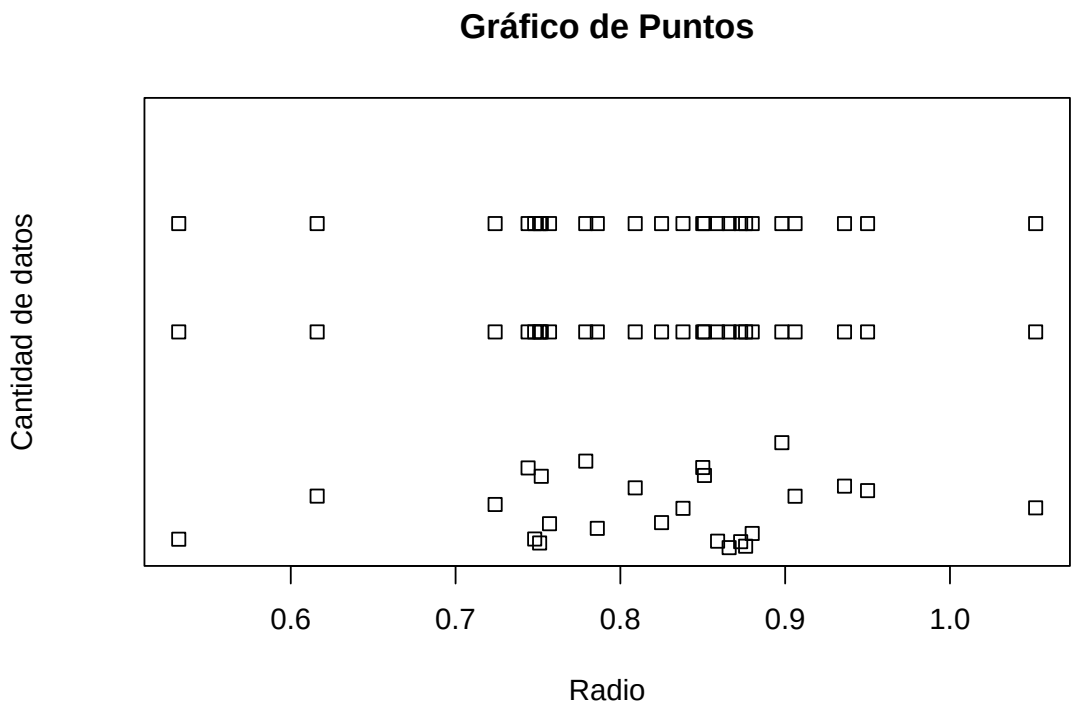
1.6. obtención del coeficiente de curtosis muestral

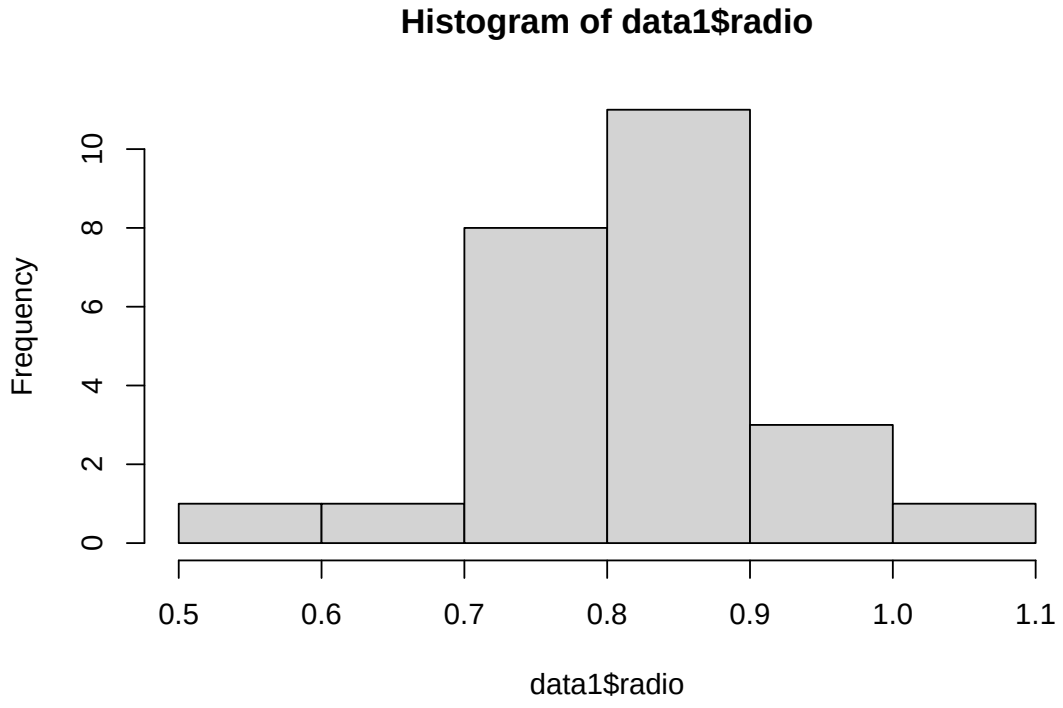
x
3.745

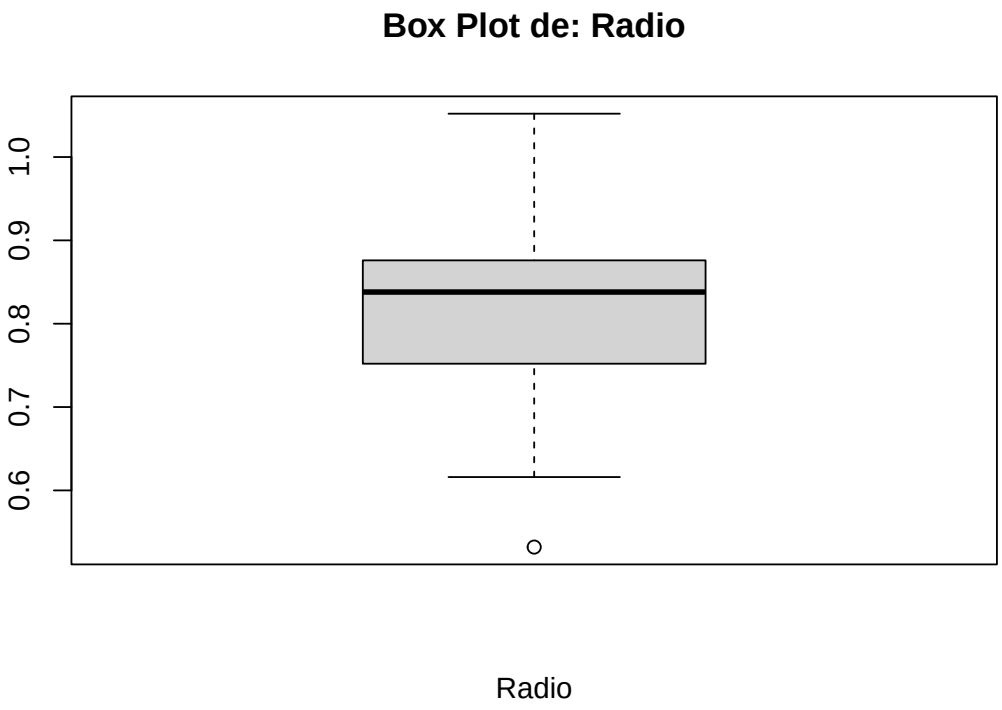
```
## [1] 1.0000000 0.8183200 0.6806088 0.5742721

##
##  D'Agostino skewness test
##
## data:  radio
## skew = -0.54306, z = -1.28329, p-value = 0.1994
## alternative hypothesis: data have a skewness

##
##  Anscombe-Glynn kurtosis test
##
## data:  radio
## kurt = 4.0637, z = 1.5936, p-value = 0.111
## alternative hypothesis: kurtosis is not equal to 3

##
##  Jarque-Bera Normality Test
##
## data:  radio
## JB = 2.4075, p-value = 0.3001
## alternative hypothesis: greater

## [1] Resultados:  Curtosis =  3.74512405032383 Asimetría =  -0.510805193333147
```

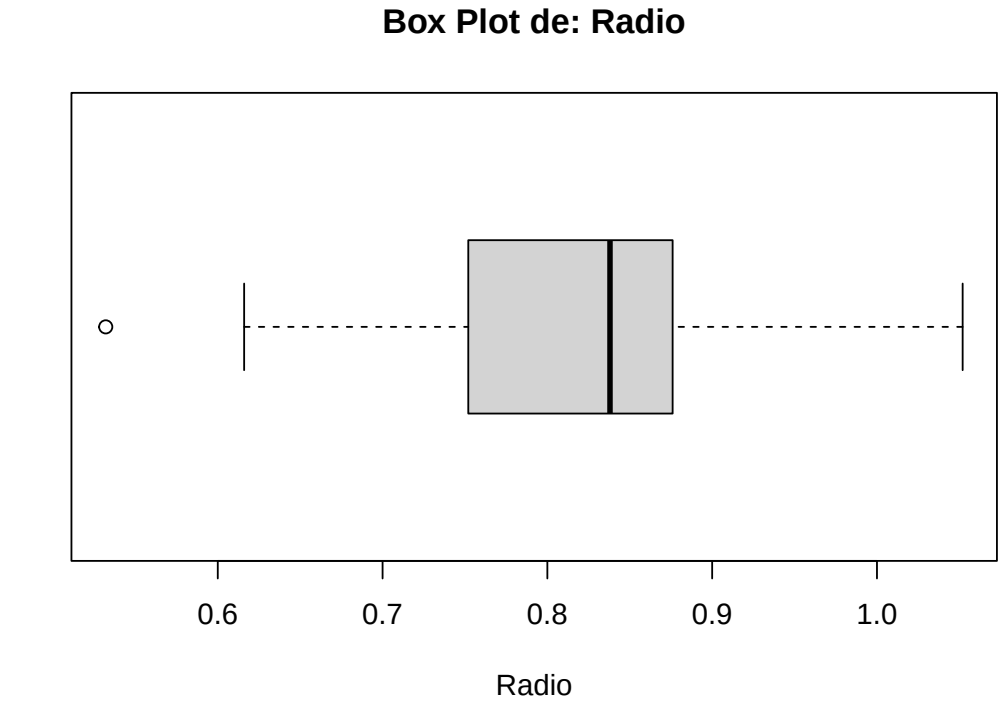
1.7. Gráficos Varios

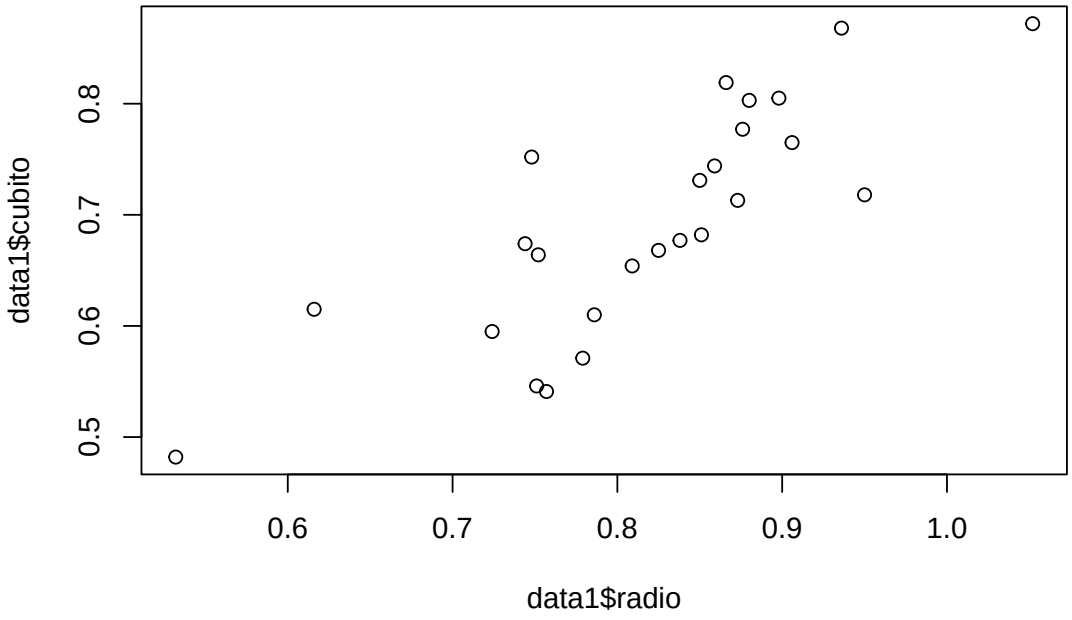
1.7.1. Gráfico de Puntos

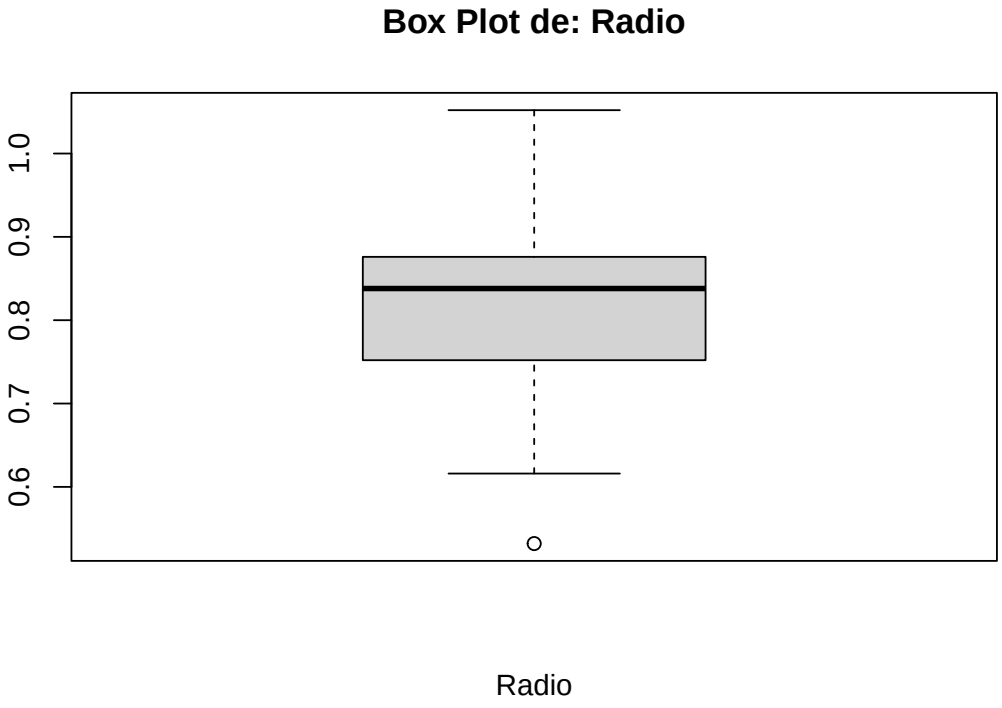




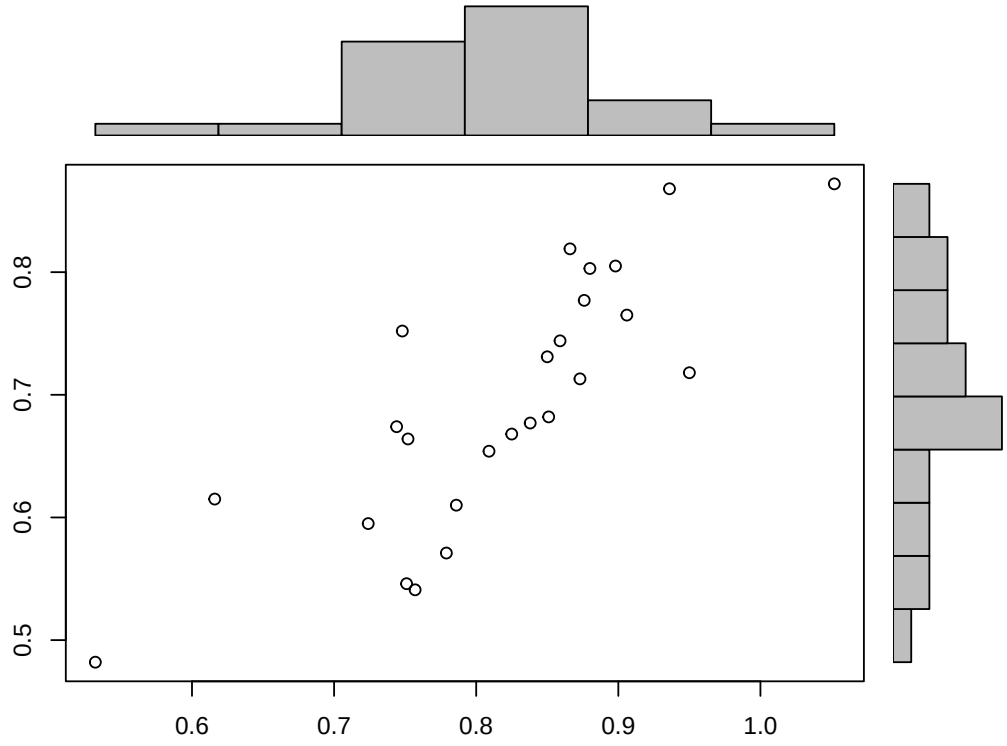

Box Plot de: Radio
1.0
0.9
0.8
0.7
0.6
Radio


Box Plot de: Radio
0.6
0.7
0.8
0.9
1.0
Radio

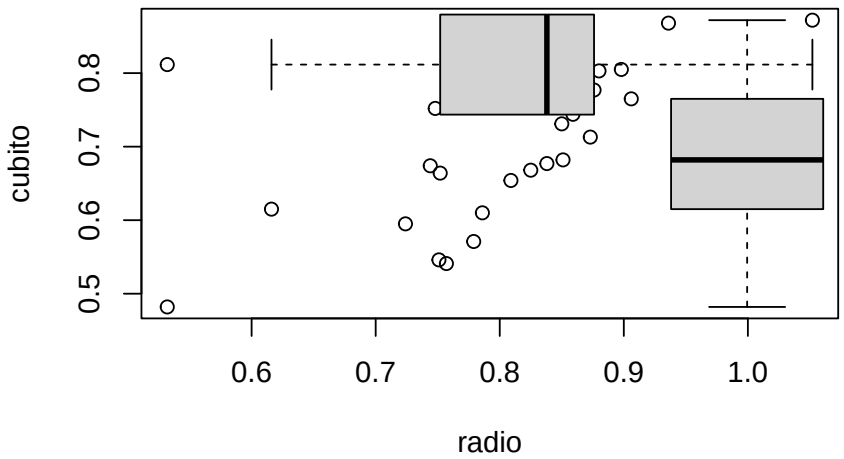

data1$cubito
0.5
0.6
0.7
0.8
0.6
0.7
0.8
0.9
1.0
data1$radio


Box Plot de: Radio
0.6
0.7
0.8
0.9
1.0
Radio

1.7.2. Scatterplot con Histograma lateral



1.7.3. Scatterplot con box-plot



1.8. CRUCE DE VBLES CUALITATIVAS vs UNA CUALITATIVA DICOTÓMICA

```
library(reporttools)   #para cruzar continuasxcategoricas
datos <- read.table(file("datos/T11-7.DAT"))
cate<-factor(rep(c("A","B","C","D"),c(15,8,15,18)))
cate2<-factor(rep(c("SI","NO"),c(28,28)))
data2<-cbind(datos,cate,cate2)
# solo tabla de vbles_nominales
titulo3<-"Caracteristicas de las variables nominales"
tableNominal(vars = data2[,6:7], group=data2[,8], cap=titulo3,vertical = FALSE,
             font.size = "scriptsize", print.pval="chi2", fisher.B="fisher.test",
             lab = "tab:nominal1",longtable = FALSE,cumsum = TRUE)
```

% latex table generated in R 4.1.3 by xtable 1.8-4 package % Mon Mar 14 14:50:54 2022

Variable	**Levels**	$\mathbf{n}_{NO}$	$\%_{NO}$	$\sum \%_{NO}$	$\mathbf{n}_{SI}$	$\%_{SI}$	$\sum \%_{SI}$	$\mathbf{n}_{all}$	$\%_{all}$	$\sum \%_{all}$
V6	SubMuli	0	0.0	0.0	11	39.3	39.3	11	19.6	19.6
	Upper	28	100.0	100.0	10	35.7	75.0	38	67.9	87.5
	Wilhelm	0	0.0	100.0	7	25.0	100.0	7	12.5	100.0
$p < 0.0001$	all	28	100.0		28	100.0		56	100.0	
cate	A	0	0.0	0.0	15	53.6	53.6	15	26.8	26.8
	B	0	0.0	0.0	8	28.6	82.1	8	14.3	41.1
	C	10	35.7	35.7	5	17.9	100.0	15	26.8	67.9
	D	18	64.3	100.0	0	0.0	100.0	18	32.1	100.0
$p < 0.0001$	all	28	100.0		28	100.0		56	100.0	

Tabla 1.14: Caracteristicas de las variables nominales