

Índice general

1. Resumen Descriptivo 3 (Ayuda)	3
1.1. Lectura de datos	3
1.2. Resumen descriptivo de una Variable	5
1.3. Resumen para cruce de vbles. cuantitativas con vbles. cualitativas	6
1.4. Cruce de Vbles. Cualitativas (cualixcuali)	7
1.5. Gráficos con ggplot2	8
1.5.1. Box-Plot	8
1.5.2. Mas opciones de Box-Plot	9
1.5.3. Box-Plot (con dos Cualitativas)	10
1.6. Obtención de Resúmenes Descriptivos	10
1.6.1. Vector de Media Muestral	10
1.6.2. Matriz de Varianzas-Covarianzas Muestral	11
1.6.3. Matriz de Correlación Muestral	11
1.7. Gráfica de la Matriz de Correlación (o Matriz de Dispersión)	11
1.7.1. Gráfico de la Matriz de Correlación por default	11
1.7.2. metodo(forma), color-forma, color-etiquetas, inclinación-etiquetas	12
1.7.3. Con: addCoef.col, order	12
1.7.4. Con: addCoef.col,order, Gráfica Triangular	14
1.7.5. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal	15
1.7.6. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con elipses	16
1.7.7. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con círculos	16
1.7.8. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con torta	17
1.8. Para fundir la matriz de correlación (la matriz de corr en forma de un data.frame con dos columnas y una tercera con las corrs)	17
1.9. Distancias	20
1.9.1. Distancia de Mahalanobis	20
1.10. Distancia Euclídeana	20
Bibliografía	20

Capítulo 1

Resumen Descriptivo 3 (Ayuda)

1.1. Lectura de datos

A continuación se tiene un encabezado del conjunto de datos.

Tabla 1.1: Encabezado de Datos

V1	V2	V3	V4	V5	V6
3.9	51	0.20	7.06	12.19	Wilhelm
2.7	49	0.07	7.14	12.23	Wilhelm
2.8	36	0.30	7.00	11.30	Wilhelm
3.1	45	0.08	7.20	13.01	Wilhelm

$$\underline{\mathbf{x}} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

A continuación se tiene el conjunto de datos completos.

Tabla 1.2: Conjunto de Datos Completo

V1	V2	V3	V4	V5	V6
3.9	51.0	0.20	7.06	12.19	Wilhelm
2.7	49.0	0.07	7.14	12.23	Wilhelm
2.8	36.0	0.30	7.00	11.30	Wilhelm
3.1	45.0	0.08	7.20	13.01	Wilhelm
3.5	46.0	0.10	7.81	12.63	Wilhelm
3.9	43.0	0.07	6.25	10.42	Wilhelm
2.7	35.0	0.00	5.11	9.00	Wilhelm
5.0	47.0	0.07	7.06	6.10	SubMuli
3.4	32.0	0.20	5.82	4.69	SubMuli
1.2	12.0	0.00	5.54	3.15	SubMuli

Continúa en la siguiente página

Tabla 1.2: Conjunto de Datos Completo Continuación

V1	V2	V3	V4	V5	V6
8.4	17.0	0.07	6.31	4.55	SubMuli
4.2	36.0	0.50	9.25	4.95	SubMuli
4.2	35.0	0.50	5.69	2.22	SubMuli
3.9	41.0	0.10	5.63	2.94	SubMuli
3.9	36.0	0.07	6.19	2.27	SubMuli
7.3	32.0	0.30	8.02	12.92	SubMuli
4.4	46.0	0.07	7.54	5.76	SubMuli
3.0	30.0	0.00	5.12	10.77	SubMuli
6.3	13.0	0.50	4.24	8.27	Upper
1.7	5.6	1.00	5.69	4.64	Upper
7.3	24.0	0.00	4.34	2.99	Upper
7.8	18.0	0.50	3.92	6.09	Upper
7.8	25.0	0.70	5.39	6.20	Upper
7.8	26.0	1.00	5.02	2.50	Upper
9.5	17.0	0.05	3.52	5.71	Upper
7.7	14.0	0.30	4.65	8.63	Upper
11.0	20.0	0.50	4.27	8.40	Upper
8.0	14.0	0.30	4.32	7.87	Upper
8.4	18.0	0.20	4.38	7.98	Upper
10.0	18.0	0.10	3.06	7.67	Upper
7.3	15.0	0.05	3.76	6.84	Upper
9.5	22.0	0.30	3.98	5.02	Upper
8.4	15.0	0.20	5.02	10.12	Upper
8.4	17.0	0.20	4.42	8.25	Upper
9.5	25.0	0.50	4.44	5.95	Upper
7.2	22.0	1.00	4.70	3.49	Upper
4.0	12.0	0.50	5.71	6.32	Upper
6.7	52.0	0.50	4.80	3.20	Upper
9.0	27.0	0.30	3.69	3.30	Upper
7.8	29.0	1.50	6.72	5.75	Upper
4.5	41.0	0.50	3.33	2.27	Upper
6.2	34.0	0.70	7.56	6.93	Upper
5.6	20.0	0.50	5.07	6.70	Upper
9.0	17.0	0.20	4.39	8.33	Upper
8.4	20.0	0.10	3.74	3.77	Upper
9.5	19.0	0.50	3.72	7.37	Upper
9.0	20.0	0.50	5.97	11.17	Upper
6.2	16.0	0.05	4.23	4.18	Upper
7.3	20.0	0.50	4.39	3.50	Upper
3.6	15.0	0.70	7.00	4.82	Upper

Continúa en la siguiente página

Tabla 1.2: Conjunto de Datos Completo Continuación

V1	V2	V3	V4	V5	V6
6.2	34.0	0.07	4.84	2.37	Upper
7.3	22.0	0.00	4.13	2.70	Upper
4.1	29.0	0.70	5.78	7.76	Upper
5.4	29.0	0.20	4.64	2.65	Upper
5.0	34.0	0.70	4.21	6.50	Upper
6.2	27.0	0.30	3.97	2.97	Upper

1.2. Resumen descriptivo de una Variable

Estadísticos de tukey: *Min, Q1, Median, Q3, Max*

Tabla 1.3: Algunas Estadísticas Resúmenes-1

1.2	3.95	6.25	8.2	11
-----	------	------	-----	----

Fución *describe* del paquete: *Hmisc*

```
## datos[, 1]
##      n missing distinct      Info      Mean      Gmd      .05      .10
##      56       0       31    0.997     6.18    2.799    2.700    3.050
##      .25     .50     .75     .90     .95
##    3.975    6.250    8.100    9.250    9.500
##
## lowest :  1.2  1.7  2.7  2.8  3.0, highest:  8.4  9.0  9.5 10.0 11.0
```

Fución *stat.desc* del paquete: *pastecs*

Tabla 1.4: Algunas Estadísticas Resúmenes-3

	x
nbr.val	56.0000000
nbr.null	0.0000000
nbr.na	0.0000000
min	1.2000000
max	11.0000000
range	9.8000000
sum	346.1000000
median	6.2500000
mean	6.1803571
SE.mean	0.3239611
CI.mean.0.95	0.6492325
var	5.8772435
std.dev	2.4243027
coef.var	0.3922593

Tabla 1.5: Algunas Estadísticas Resúmenes-4 (Varias Vbles.)

	V1	V2	V3	V4	V5
nbr.val	56.0000	56.0000	56.0000	56.0000	56.0000
nbr.null	0.0000	0.0000	5.0000	0.0000	0.0000
nbr.na	0.0000	0.0000	0.0000	0.0000	0.0000
min	1.2000	5.6000	0.0000	3.0600	2.2200
max	11.0000	52.0000	1.5000	9.2500	13.0100
range	9.8000	46.4000	1.5000	6.1900	10.7900
sum	346.1000	1514.6000	19.1200	296.7500	360.2800
median	6.2500	25.0000	0.3000	5.0200	6.0950
mean	6.1804	27.0464	0.3414	5.2991	6.4336
SE.mean	0.3240	1.5509	0.0419	0.1854	0.4215
CI.mean.0.95	0.6492	3.1080	0.0841	0.3715	0.8448
var	5.8772	134.6880	0.0985	1.9244	9.9512
std.dev	2.4243	11.6055	0.3139	1.3872	3.1546
coef.var	0.3923	0.4291	0.9194	0.2618	0.4903

Fución *describe* del paquete: *psych*

Fución *stargazer* del paquete: *stargazer*

1.3. Resumen para cruce de vbles. cuantitativas con vbles. cualitativas

En este caso se utilizia el paquete: *reporttools*

Tabla 1.6: Algunas Estadísticas Resúmenes-5 (Varías Vbles.)

Estadísticos	V1	V2	V3	V4	V5
nbr.val	56.0000	56.0000	56.0000	56.0000	56.0000
nbr.null	0.0000	0.0000	5.0000	0.0000	0.0000
nbr.na	0.0000	0.0000	0.0000	0.0000	0.0000
min	1.2000	5.6000	0.0000	3.0600	2.2200
max	11.0000	52.0000	1.5000	9.2500	13.0100
range	9.8000	46.4000	1.5000	6.1900	10.7900
sum	346.1000	1514.6000	19.1200	296.7500	360.2800
median	6.2500	25.0000	0.3000	5.0200	6.0950
mean	6.1804	27.0464	0.3414	5.2991	6.4336
SE.mean	0.3240	1.5509	0.0419	0.1854	0.4215
CI.mean.0.95	0.6492	3.1080	0.0841	0.3715	0.8448
var	5.8772	134.6880	0.0985	1.9244	9.9512
std.dev	2.4243	11.6055	0.3139	1.3872	3.1546
coef.var	0.3923	0.4291	0.9194	0.2618	0.4903

Tabla 1.7: Algunas Estadísticas Resúmenes-6

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	56	6.180357	2.424303	6.25	6.206522	3.18759	1.2	11	9.8	-0.1024264	-1.141537	0.3239611

Tabla 1.8: Resumen de Datos con stargazer

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
V1	56	6.1804	2.4243	1.2000	3.9750	8.1000	11.0000
V2	56	27.0464	11.6055	6	17.8	35	52
V3	56	0.3414	0.3139	0.0000	0.0700	0.5000	1.5000
V4	56	5.2991	1.3872	3.0600	4.2625	6.2050	9.2500
V5	56	6.4336	3.1546	2.2200	3.4975	8.2850	13.0100

1.4. Cruce de Vbles. Cualitativas (cualixcuali)

Variable	Levels	\bar{x}	s	c_v %	Min	\tilde{x}	Max	IQR	n	#NA	skewness	kurtosis
V1	SubMuli	4.4	2.0	44.2	1.2	4.2	8.4	1.1	11	0	0.6	2.6
	Upper	7.2	2.0	27.6	1.7	7.5	11.0	2.2	38	0	-0.6	3.0
	Wilhelm	3.2	0.5	16.7	2.7	3.1	3.9	1.0	7	0	0.2	1.0
$p < 0.0001$	all	6.2	2.4	39.2	1.2	6.2	11.0	4.1	56	0	-0.1	1.9
V2	SubMuli	33.1	10.8	32.5	12.0	35.0	47.0	7.5	11	0	-0.6	2.3
	Upper	22.3	8.8	39.4	5.6	20.0	52.0	9.8	38	0	1.1	4.9
	Wilhelm	43.6	6.1	14.0	35.0	45.0	51.0	8.0	7	0	-0.3	1.3
$p < 0.0001$	all	27.0	11.6	42.9	5.6	25.0	52.0	17.2	56	0	0.5	2.2
V3	SubMuli	0.2	0.2	107.6	0.0	0.1	0.5	0.2	11	0	0.9	2.1
	Upper	0.4	0.3	76.4	0.0	0.5	1.5	0.3	38	0	1.0	4.0
	Wilhelm	0.1	0.1	85.5	0.0	0.1	0.3	0.1	7	0	0.7	1.9
$p = 0.0039$	all	0.3	0.3	91.9	0.0	0.3	1.5	0.4	56	0	1.3	4.8
V4	SubMuli	6.6	1.3	19.3	5.1	6.2	9.2	1.6	11	0	0.8	2.3
	Upper	4.7	1.0	21.3	3.1	4.4	7.6	1.0	38	0	1.1	3.8
	Wilhelm	6.8	0.9	12.8	5.1	7.1	7.8	0.5	7	0	-0.8	2.3
$p < 0.0001$	all	5.3	1.4	26.2	3.1	5.0	9.2	1.9	56	0	0.7	2.7

Tabla 1.9: Resumen de los Datos con reporttools (cruce de cuantixcuali)

Variable	\bar{x}	s	c_v %	Min	\tilde{x}	Max	IQR	n	#NA	skewness	kurtosis
V1	6.2	2.4	39.2	1.2	6.2	11.0	4.1	56	0	-0.1	1.9
V2	27.0	11.6	42.9	5.6	25.0	52.0	17.2	56	0	0.5	2.2
V3	0.3	0.3	91.9	0.0	0.3	1.5	0.4	56	0	1.3	4.8
V4	5.3	1.4	26.2	3.1	5.0	9.2	1.9	56	0	0.7	2.7

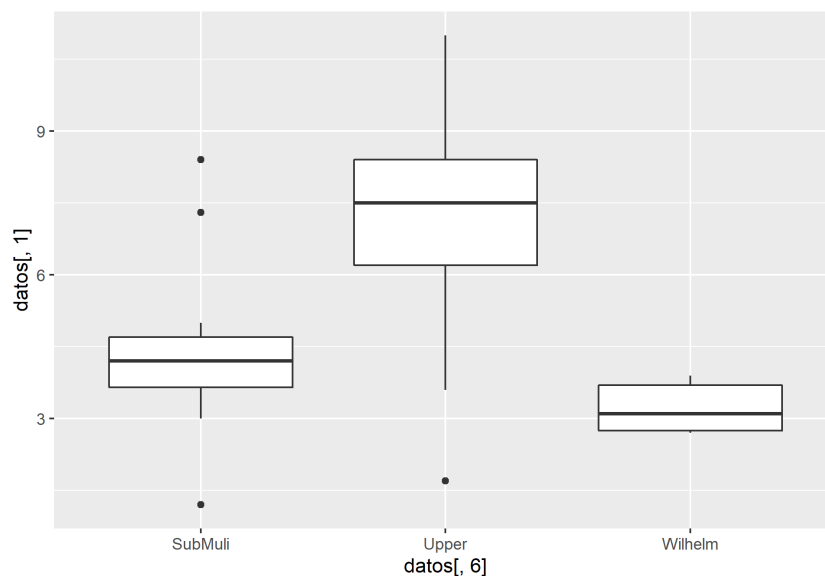
Tabla 1.10: Resumen de los Datos con reporttools (cruce de cuantixcuali)

1.5. Gráficos con ggplot2

1.5.1. Box-Plot

Se realiza el gráfico *box_plot* para las variables 6 y 1 del archivo de datos *datos* y se guarda dicha gráfica en la carpeta: *figuras.R* con el nombre: *graf1.png*.

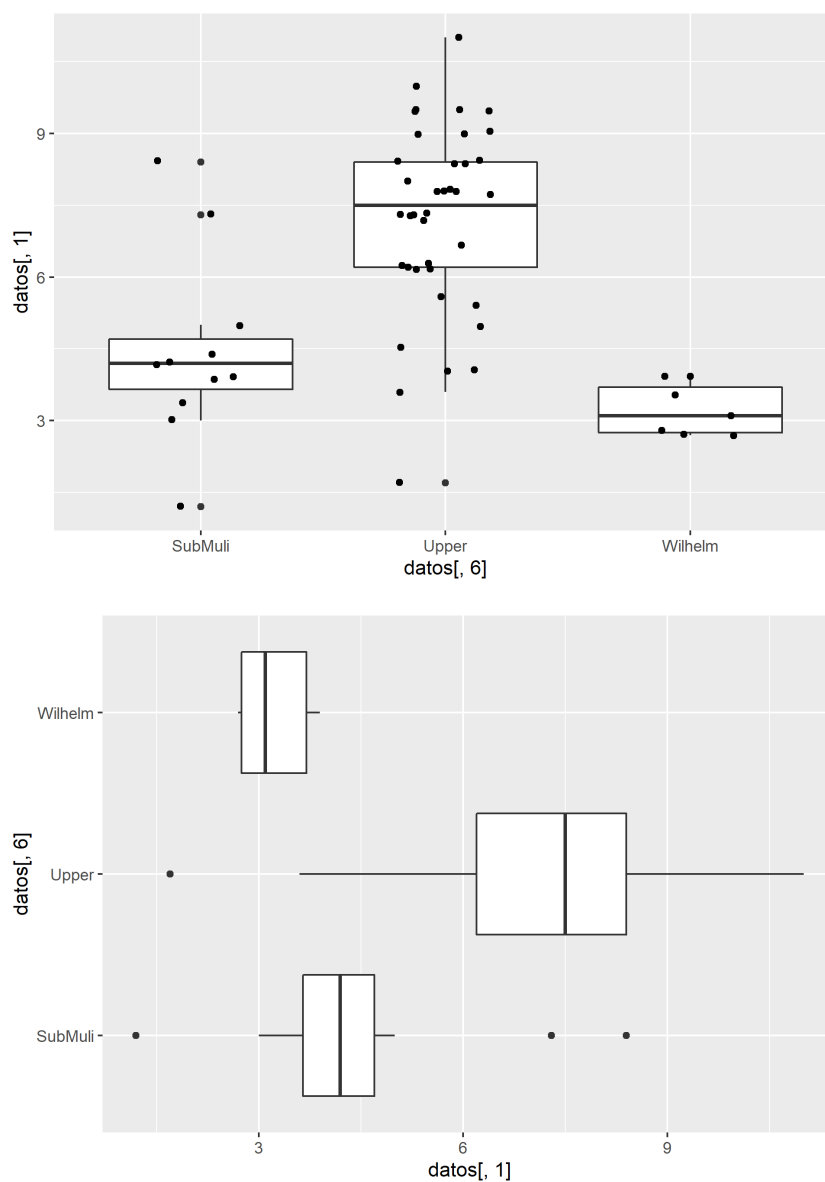
Se carga al documento la gráfica de nombre: *graf1.png* de la carpeta: *figura.R*.



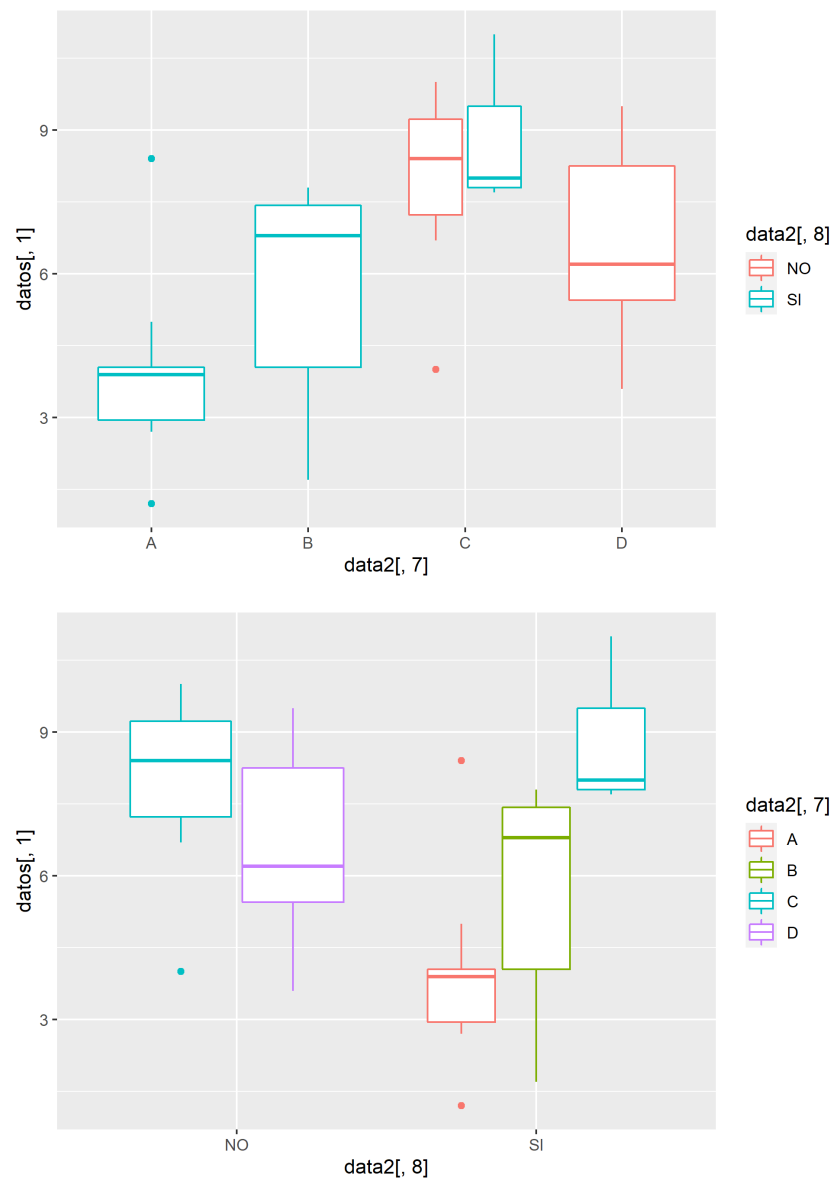
Variable	Levels	n	%	Σ %
V6	SubMuli	11	19.6	19.6
	Upper	38	67.9	87.5
	Wilhelm	7	12.5	100.0
	all	56	100.0	
cate	A	15	26.8	26.8
	B	8	14.3	41.1
	C	15	26.8	67.9
	D	18	32.1	100.0
	all	56	100.0	

Tabla 1.11: Características de las variables nominales

1.5.2. Mas opciones de Box-Plot



1.5.3. Box-Plot (con dos Cualitativas)



1.6. Obtención de Resúmenes Descriptivos

1.6.1. Vector de Media Muestral

Tabla 1.12: Vector de Medias Muestrales

	x
V1	6.1804
V2	27.0464
V3	0.3414
V4	5.2991
V5	6.4336

1.6.2. Matriz de Varianzas-Covarianzas Muestral

Tabla 1.13: Matriz de Var-Cov Muestrales

	V1	V2	V3	V4	V5
V1	5.8772	-12.4027	0.0844	-1.8475	-0.6213
V2	-12.4027	134.6880	-0.6125	8.3181	6.5687
V3	0.0844	-0.6125	0.0985	0.0311	-0.1622
V4	-1.8475	8.3181	0.0311	1.9244	1.6316
V5	-0.6213	6.5687	-0.1622	1.6316	9.9512

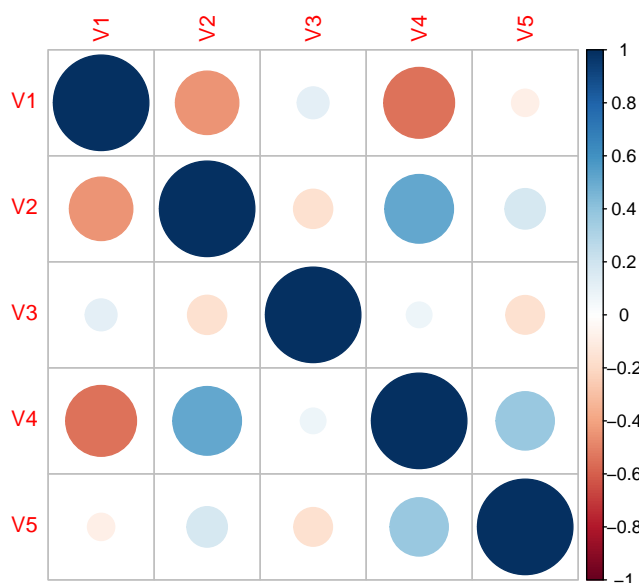
1.6.3. Matriz de Correlación Muestral

Tabla 1.14: Matriz de Correlación Muestrales

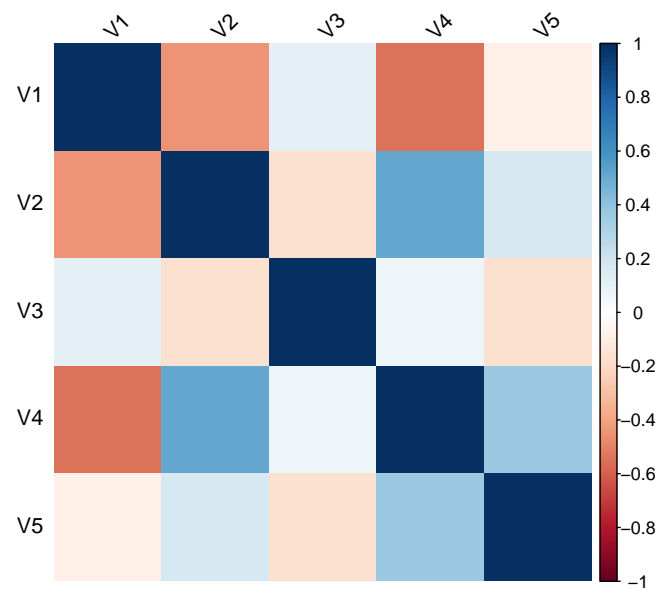
	V1	V2	V3	V4	V5
V1	1.0000	-0.4408	0.1110	-0.5493	-0.0812
V2	-0.4408	1.0000	-0.1681	0.5167	0.1794
V3	0.1110	-0.1681	1.0000	0.0715	-0.1638
V4	-0.5493	0.5167	0.0715	1.0000	0.3728
V5	-0.0812	0.1794	-0.1638	0.3728	1.0000

1.7. Gráfica de la Matriz de Correlación (o Matriz de Dispersión)

1.7.1. Gráfico de la Matriz de Correlación por default



1.7.2. metodo(forma), color-forma, color-etiquetas, inclinación-etiquetas



1.7.3. Con: addCoef.col, order

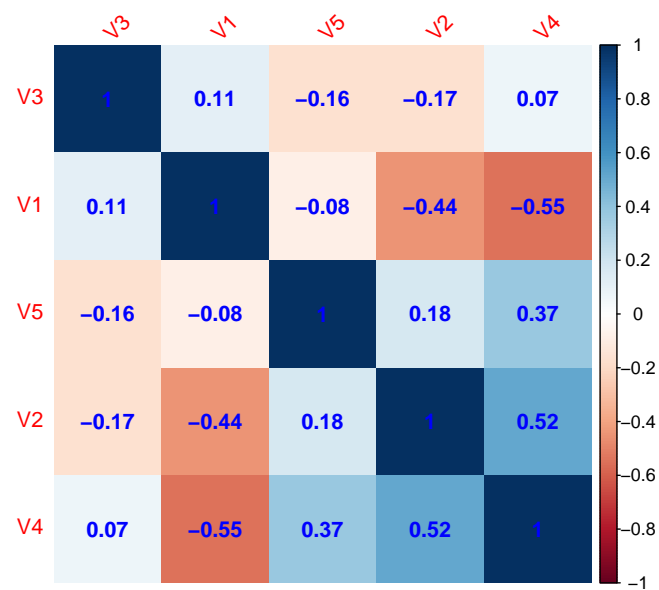
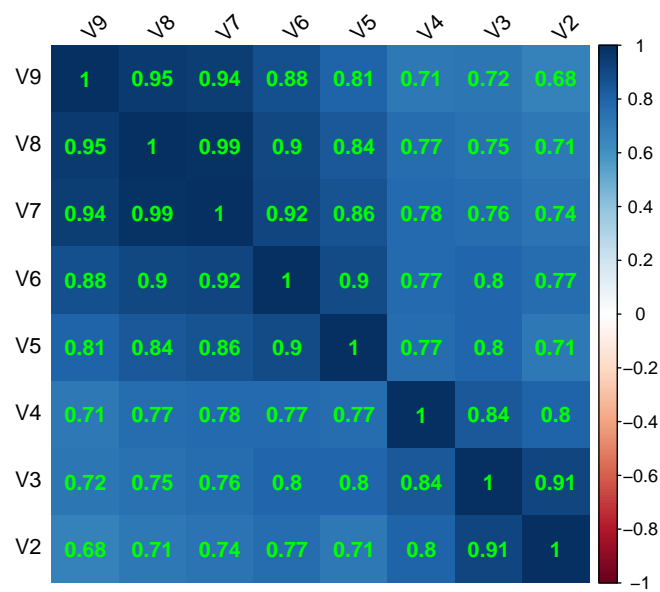
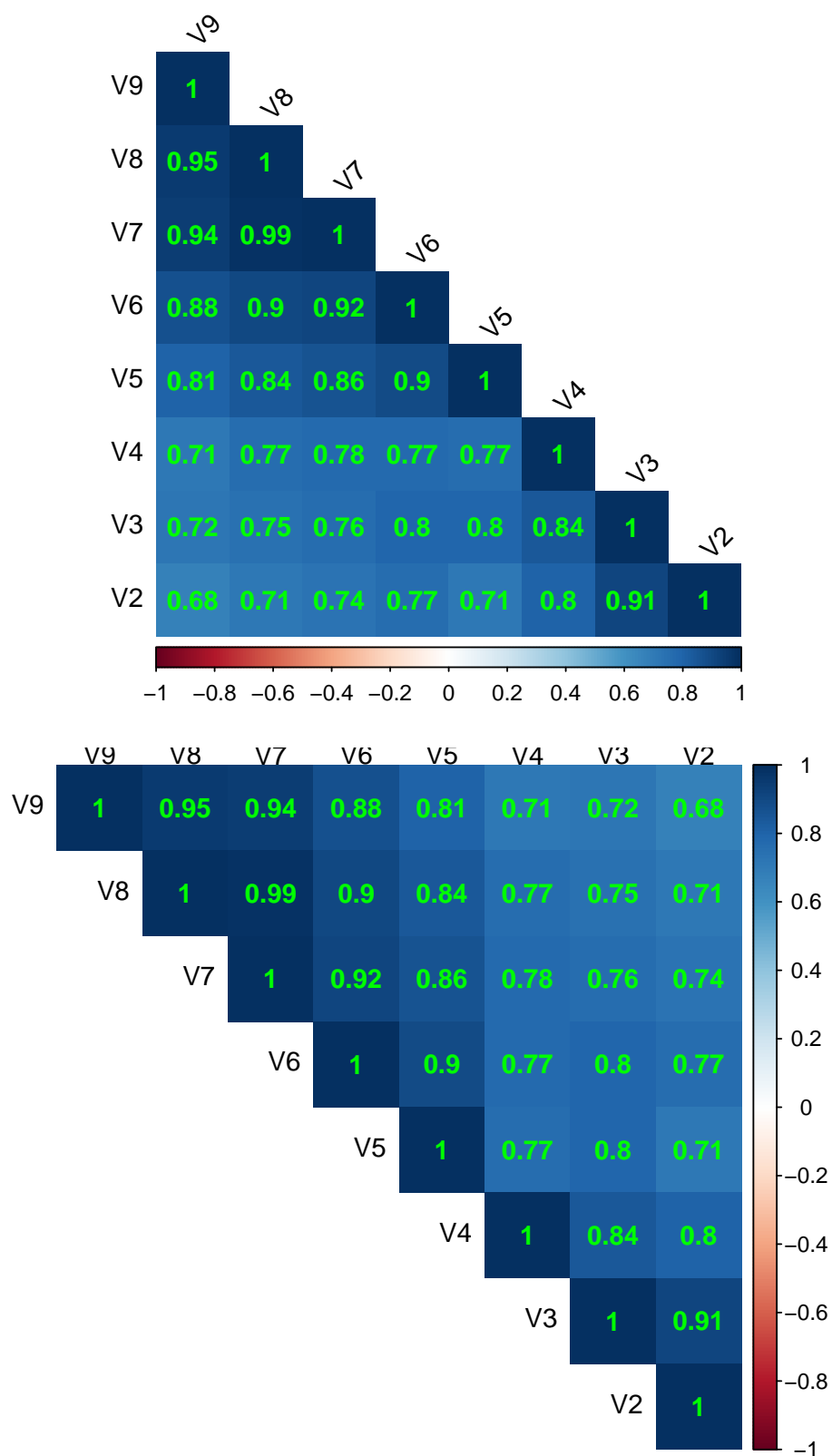


Tabla 1.15: Matriz de Correlación Muestrales

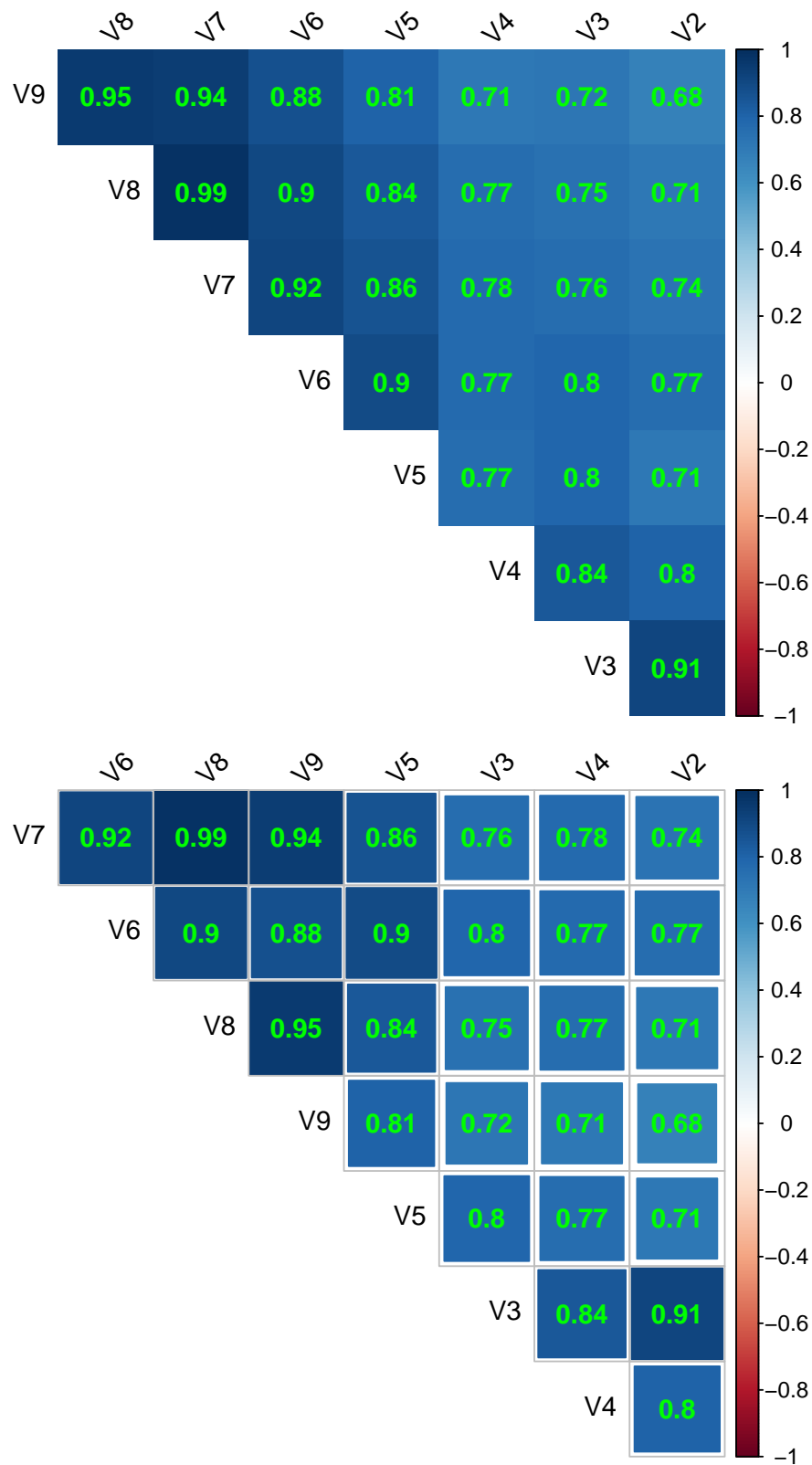
	V2	V3	V4	V5	V6	V7	V8	V9
V2	1.0000	0.9148	0.8041	0.7119	0.7658	0.7399	0.7148	0.6765
V3	0.9148	1.0000	0.8449	0.7969	0.7951	0.7613	0.7480	0.7211
V4	0.8041	0.8449	1.0000	0.7677	0.7716	0.7797	0.7657	0.7127
V5	0.7119	0.7969	0.7677	1.0000	0.8958	0.8607	0.8431	0.8070
V6	0.7658	0.7951	0.7716	0.8958	1.0000	0.9165	0.9013	0.8778
V7	0.7399	0.7613	0.7797	0.8607	0.9165	1.0000	0.9882	0.9441
V8	0.7148	0.7480	0.7657	0.8431	0.9013	0.9882	1.0000	0.9542
V9	0.6765	0.7211	0.7127	0.8070	0.8778	0.9441	0.9542	1.0000



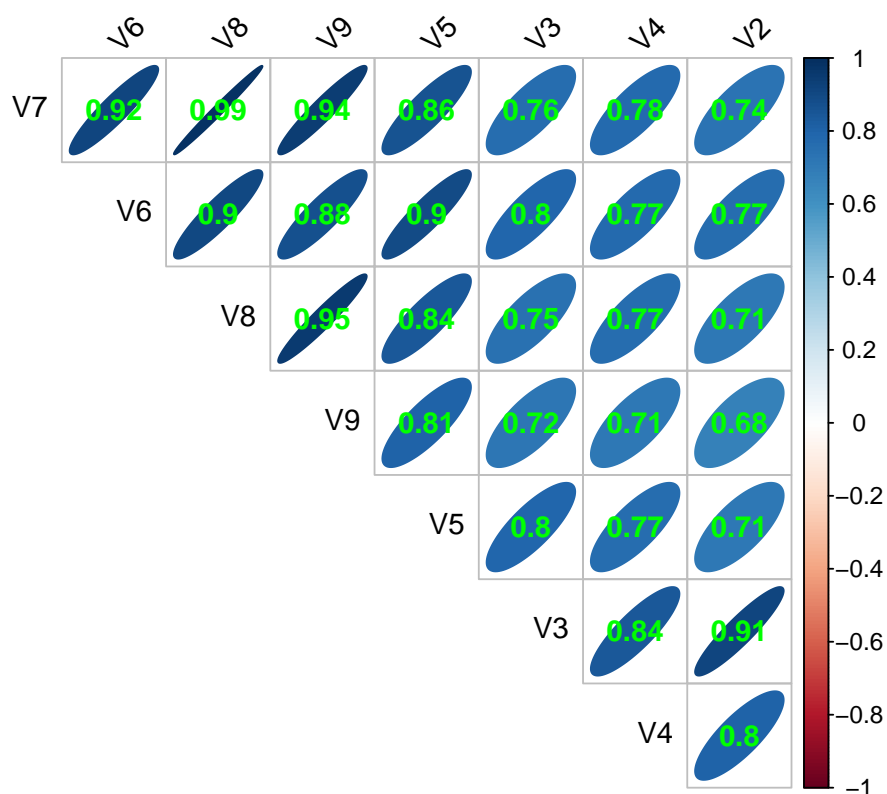
1.7.4. Con: addCoef.col,order, Gráfica Triangular



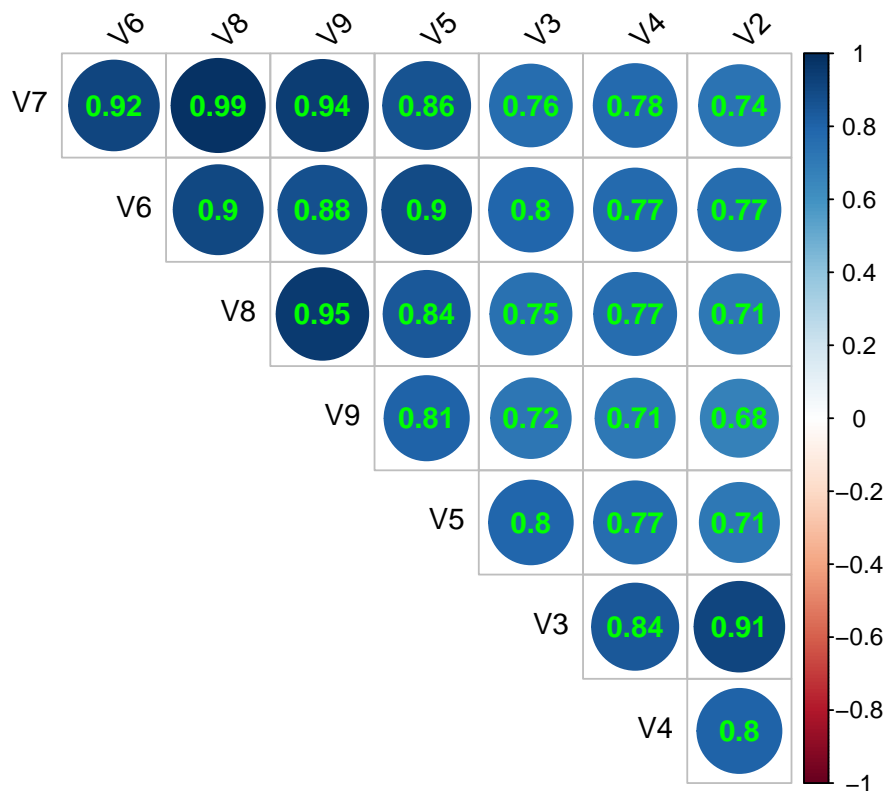
1.7.5. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal



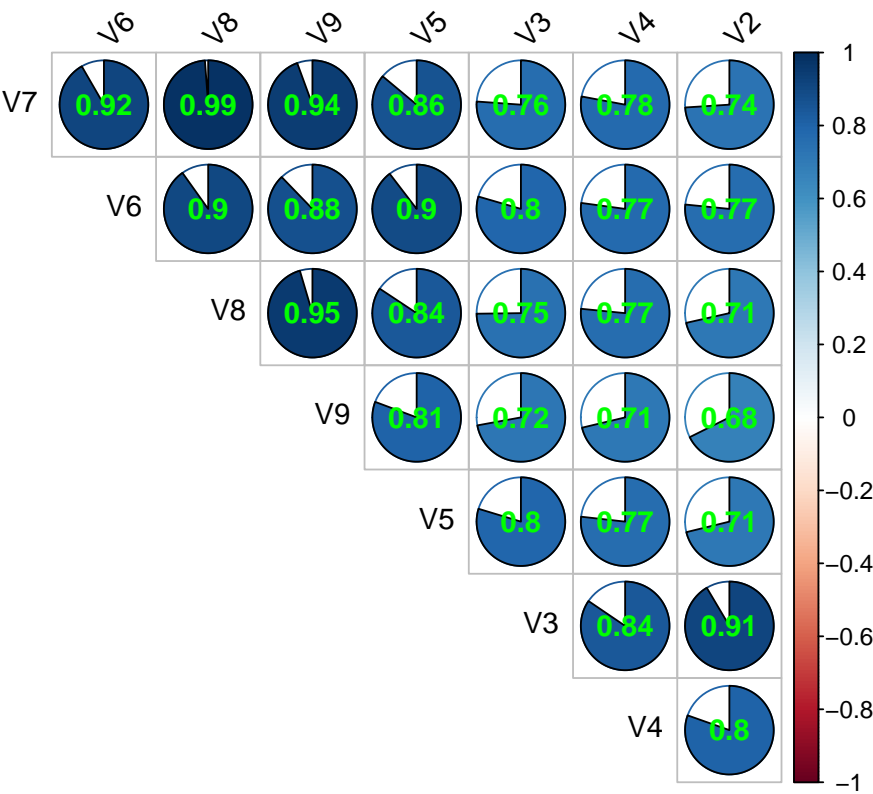
1.7.6. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con elipses



1.7.7. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con círculos



1.7.8. Con: addCoef.col,order, Gráfica Triangular, sin Diagonal, con torta



1.8. Para fundir la matriz de correlación (la matriz de corr en forma de un data.frmae con dos columnas y una tercera con las corrs)

Tabla 1.16: Correlaciones por Pares

Var1	Var2	value
V2	V2	1.0000
V3	V2	0.9148
V4	V2	0.8041
V5	V2	0.7119
V6	V2	0.7658
V7	V2	0.7399

Tabla 1.17: Correlaciones Por Pares2

Var1	Var2	value
V2	V2	1.0000
V3	V2	0.9148

Continúa en la siguiente página

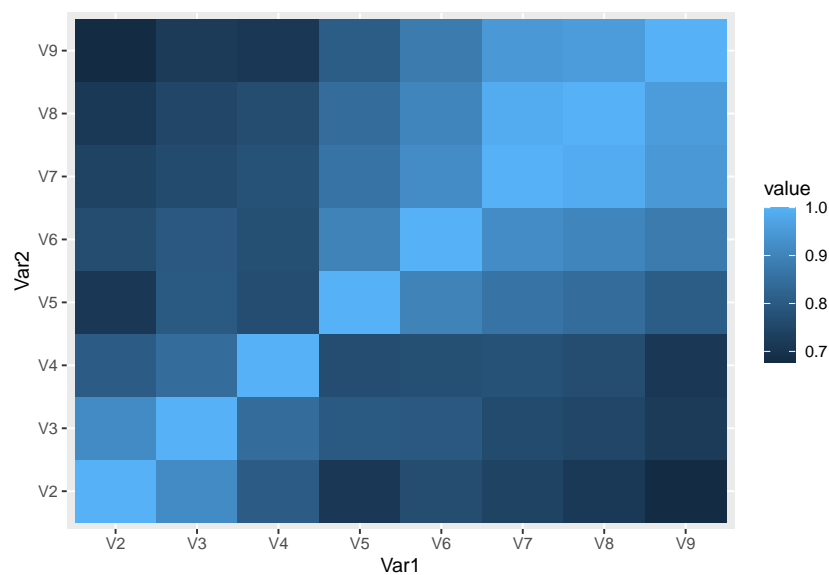
Tabla 1.17: Correlaciones Por Pares2 Continuación

Var1	Var2	value
V4	V2	0.8041
V5	V2	0.7119
V6	V2	0.7658
V7	V2	0.7399
V8	V2	0.7148
V9	V2	0.6765
V2	V3	0.9148
V3	V3	1.0000
V4	V3	0.8449
V5	V3	0.7969
V6	V3	0.7951
V7	V3	0.7613
V8	V3	0.7480
V9	V3	0.7211
V2	V4	0.8041
V3	V4	0.8449
V4	V4	1.0000
V5	V4	0.7677
V6	V4	0.7716
V7	V4	0.7797
V8	V4	0.7657
V9	V4	0.7127
V2	V5	0.7119
V3	V5	0.7969
V4	V5	0.7677
V5	V5	1.0000
V6	V5	0.8958
V7	V5	0.8607
V8	V5	0.8431
V9	V5	0.8070
V2	V6	0.7658
V3	V6	0.7951
V4	V6	0.7716
V5	V6	0.8958
V6	V6	1.0000
V7	V6	0.9165
V8	V6	0.9013
V9	V6	0.8778
V2	V7	0.7399
V3	V7	0.7613

Continúa en la siguiente página

Tabla 1.17: Correlaciones Por Pares2 Continuación

Var1	Var2	value
V4	V7	0.7797
V5	V7	0.8607
V6	V7	0.9165
V7	V7	1.0000
V8	V7	0.9882
V9	V7	0.9441
V2	V8	0.7148
V3	V8	0.7480
V4	V8	0.7657
V5	V8	0.8431
V6	V8	0.9013
V7	V8	0.9882
V8	V8	1.0000
V9	V8	0.9542
V2	V9	0.6765
V3	V9	0.7211
V4	V9	0.7127
V5	V9	0.8070
V6	V9	0.8778
V7	V9	0.9441
V8	V9	0.9542
V9	V9	1.0000



1.9. Distancias

1.9.1. Distancia de Mahalanobis

Tabla 1.18: Datos

x1	11.00	12.5	13.00	13.70	14	5.00	15.00	8.00	8	5.00
x2	18.95	19.0	17.95	15.54	14	12.95	8.94	7.49	6	3.99

Tabla 1.19: Medias

x	
x1	10.52
x2	12.48

Tabla 1.20: Matriz de Var-Cov

	x1	x2
x1	14.0262	11.3241
x2	11.3241	30.8544

Tabla 1.21: Dist. Mahalanobis (Al cuadrado)

1.72	1.395	0.992	0.732	0.94	3.29	3.791	0.855	1.363	2.922
------	-------	-------	-------	------	------	-------	-------	-------	-------

Tabla 1.22: Dist. Mahalanobis Ordenadas

0.732	0.855	0.94	0.992	1.363	1.395	1.72	2.922	3.29	3.791
-------	-------	------	-------	-------	-------	------	-------	------	-------

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

```
## [1] 5
```

1.10. Distancia Euclideana

Tabla 1.23: Dist. Euclideanas

1	2	3	4	5	6	7	8	9	10
0.000	1.501	2.236	4.349	5.788	8.485	10.780	11.846	13.293	16.118
1.501	0.000	1.163	3.662	5.220	9.636	10.366	12.358	13.757	16.779
2.236	1.163	0.000	2.510	4.075	9.434	9.229	11.594	12.954	16.090
4.349	3.662	2.510	0.000	1.569	9.077	6.727	9.864	11.113	14.460
5.788	5.220	4.075	1.569	0.000	9.061	5.158	8.853	10.000	13.461
8.485	9.636	9.434	9.077	9.061	0.000	10.774	6.230	7.570	8.960
10.780	10.366	9.229	6.727	5.158	10.774	0.000	7.149	7.592	11.158
11.846	12.358	11.594	9.864	8.853	6.230	7.149	0.000	1.490	4.610
13.293	13.757	12.954	11.113	10.000	7.570	7.592	1.490	0.000	3.611
16.118	16.779	16.090	14.460	13.461	8.960	11.158	4.610	3.611	0.000

Tabla 1.24: Dist. Euclideanas a la Media

1	2	3	4	5	6	7	8	9	10	11
0.000	1.501	2.236	4.349	5.788	8.485	10.780	11.846	13.293	16.118	6.487
1.501	0.000	1.163	3.662	5.220	9.636	10.366	12.358	13.757	16.779	6.813
2.236	1.163	0.000	2.510	4.075	9.434	9.229	11.594	12.954	16.090	6.005
4.349	3.662	2.510	0.000	1.569	9.077	6.727	9.864	11.113	14.460	4.412
5.788	5.220	4.075	1.569	0.000	9.061	5.158	8.853	10.000	13.461	3.797
8.485	9.636	9.434	9.077	9.061	0.000	10.774	6.230	7.570	8.960	5.540
10.780	10.366	9.229	6.727	5.158	10.774	0.000	7.149	7.592	11.158	5.710
11.846	12.358	11.594	9.864	8.853	6.230	7.149	0.000	1.490	4.610	5.591
13.293	13.757	12.954	11.113	10.000	7.570	7.592	1.490	0.000	3.611	6.954
16.118	16.779	16.090	14.460	13.461	8.960	11.158	4.610	3.611	0.000	10.128
6.487	6.813	6.005	4.412	3.797	5.540	5.710	5.591	6.954	10.128	0.000

Tabla 1.25: Dist. Euclideanas a la Media2

1	2	3	4	5	6	7	8	9	10	11
6.487	6.813	6.005	4.412	3.797	5.54	5.71	5.591	6.954	10.128	0

Tabla 1.26: Dist. Euclideanas y de Mahalanobis a la Media

dist_eucli_media_vector	distancias2
6.487	1.720
6.813	1.395
6.005	0.992
4.412	0.732
3.797	0.940
5.540	3.290
5.710	3.791
5.591	0.855
6.954	1.363
10.128	2.922
0.000	1.720