

# PAC4 - Algebra lineal

Josep Andreu Miralles

22/5/2021

1. [5%] Genereu un vector label que contingui la cinquena columna (Month) del data frame `airquality`. De la mateixa manera, genereu una matriu `W` de 153 files i 4 columnes del data frame `airquality`. Per a generar la matriu feu servir, per exemple, `as.matrix`. Assegureu-vos que tant `label` com `W` són matrius i no data frames fent servir la instrucció:

```
1 > class ( label )
2 [1] "integer"
3 > class (W)
4 [1] "matrix" "array"
```

```
nycair <- airquality
nycair
```

##	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6
## 7	23	299	8.6	65	5	7
## 8	19	99	13.8	59	5	8
## 9	8	19	20.1	61	5	9
## 10	NA	194	8.6	69	5	10
## 11	7	NA	6.9	74	5	11
## 12	16	256	9.7	69	5	12
## 13	11	290	9.2	66	5	13
## 14	14	274	10.9	68	5	14
## 15	18	65	13.2	58	5	15
## 16	14	334	11.5	64	5	16
## 17	34	307	12.0	66	5	17
## 18	6	78	18.4	57	5	18
## 19	30	322	11.5	68	5	19
## 20	11	44	9.7	62	5	20
## 21	1	8	9.7	59	5	21
## 22	11	320	16.6	73	5	22
## 23	4	25	9.7	61	5	23
## 24	32	92	12.0	61	5	24
## 25	NA	66	16.6	57	5	25

## 26	NA	266	14.9	58	5	26
## 27	NA	NA	8.0	57	5	27
## 28	23	13	12.0	67	5	28
## 29	45	252	14.9	81	5	29
## 30	115	223	5.7	79	5	30
## 31	37	279	7.4	76	5	31
## 32	NA	286	8.6	78	6	1
## 33	NA	287	9.7	74	6	2
## 34	NA	242	16.1	67	6	3
## 35	NA	186	9.2	84	6	4
## 36	NA	220	8.6	85	6	5
## 37	NA	264	14.3	79	6	6
## 38	29	127	9.7	82	6	7
## 39	NA	273	6.9	87	6	8
## 40	71	291	13.8	90	6	9
## 41	39	323	11.5	87	6	10
## 42	NA	259	10.9	93	6	11
## 43	NA	250	9.2	92	6	12
## 44	23	148	8.0	82	6	13
## 45	NA	332	13.8	80	6	14
## 46	NA	322	11.5	79	6	15
## 47	21	191	14.9	77	6	16
## 48	37	284	20.7	72	6	17
## 49	20	37	9.2	65	6	18
## 50	12	120	11.5	73	6	19
## 51	13	137	10.3	76	6	20
## 52	NA	150	6.3	77	6	21
## 53	NA	59	1.7	76	6	22
## 54	NA	91	4.6	76	6	23
## 55	NA	250	6.3	76	6	24
## 56	NA	135	8.0	75	6	25
## 57	NA	127	8.0	78	6	26
## 58	NA	47	10.3	73	6	27
## 59	NA	98	11.5	80	6	28
## 60	NA	31	14.9	77	6	29
## 61	NA	138	8.0	83	6	30
## 62	135	269	4.1	84	7	1
## 63	49	248	9.2	85	7	2
## 64	32	236	9.2	81	7	3
## 65	NA	101	10.9	84	7	4
## 66	64	175	4.6	83	7	5
## 67	40	314	10.9	83	7	6
## 68	77	276	5.1	88	7	7
## 69	97	267	6.3	92	7	8
## 70	97	272	5.7	92	7	9
## 71	85	175	7.4	89	7	10
## 72	NA	139	8.6	82	7	11
## 73	10	264	14.3	73	7	12
## 74	27	175	14.9	81	7	13
## 75	NA	291	14.9	91	7	14
## 76	7	48	14.3	80	7	15
## 77	48	260	6.9	81	7	16
## 78	35	274	10.3	82	7	17
## 79	61	285	6.3	84	7	18

## 80	79	187	5.1	87	7	19
## 81	63	220	11.5	85	7	20
## 82	16	7	6.9	74	7	21
## 83	NA	258	9.7	81	7	22
## 84	NA	295	11.5	82	7	23
## 85	80	294	8.6	86	7	24
## 86	108	223	8.0	85	7	25
## 87	20	81	8.6	82	7	26
## 88	52	82	12.0	86	7	27
## 89	82	213	7.4	88	7	28
## 90	50	275	7.4	86	7	29
## 91	64	253	7.4	83	7	30
## 92	59	254	9.2	81	7	31
## 93	39	83	6.9	81	8	1
## 94	9	24	13.8	81	8	2
## 95	16	77	7.4	82	8	3
## 96	78	NA	6.9	86	8	4
## 97	35	NA	7.4	85	8	5
## 98	66	NA	4.6	87	8	6
## 99	122	255	4.0	89	8	7
## 100	89	229	10.3	90	8	8
## 101	110	207	8.0	90	8	9
## 102	NA	222	8.6	92	8	10
## 103	NA	137	11.5	86	8	11
## 104	44	192	11.5	86	8	12
## 105	28	273	11.5	82	8	13
## 106	65	157	9.7	80	8	14
## 107	NA	64	11.5	79	8	15
## 108	22	71	10.3	77	8	16
## 109	59	51	6.3	79	8	17
## 110	23	115	7.4	76	8	18
## 111	31	244	10.9	78	8	19
## 112	44	190	10.3	78	8	20
## 113	21	259	15.5	77	8	21
## 114	9	36	14.3	72	8	22
## 115	NA	255	12.6	75	8	23
## 116	45	212	9.7	79	8	24
## 117	168	238	3.4	81	8	25
## 118	73	215	8.0	86	8	26
## 119	NA	153	5.7	88	8	27
## 120	76	203	9.7	97	8	28
## 121	118	225	2.3	94	8	29
## 122	84	237	6.3	96	8	30
## 123	85	188	6.3	94	8	31
## 124	96	167	6.9	91	9	1
## 125	78	197	5.1	92	9	2
## 126	73	183	2.8	93	9	3
## 127	91	189	4.6	93	9	4
## 128	47	95	7.4	87	9	5
## 129	32	92	15.5	84	9	6
## 130	20	252	10.9	80	9	7
## 131	23	220	10.3	78	9	8
## 132	21	230	10.9	75	9	9
## 133	24	259	9.7	73	9	10

```
label <- nycair [, "Month"]
label
```

```
class(label)
```

```
W <- as.matrix(nycair[, 1:4])
W
```

4

##	[15,]	18	65	13.2	58
##	[16,]	14	334	11.5	64
##	[17,]	34	307	12.0	66
##	[18,]	6	78	18.4	57
##	[19,]	30	322	11.5	68
##	[20,]	11	44	9.7	62
##	[21,]	1	8	9.7	59
##	[22,]	11	320	16.6	73
##	[23,]	4	25	9.7	61
##	[24,]	32	92	12.0	61
##	[25,]	NA	66	16.6	57
##	[26,]	NA	266	14.9	58
##	[27,]	NA	NA	8.0	57
##	[28,]	23	13	12.0	67
##	[29,]	45	252	14.9	81
##	[30,]	115	223	5.7	79
##	[31,]	37	279	7.4	76
##	[32,]	NA	286	8.6	78
##	[33,]	NA	287	9.7	74
##	[34,]	NA	242	16.1	67
##	[35,]	NA	186	9.2	84
##	[36,]	NA	220	8.6	85
##	[37,]	NA	264	14.3	79
##	[38,]	29	127	9.7	82
##	[39,]	NA	273	6.9	87
##	[40,]	71	291	13.8	90
##	[41,]	39	323	11.5	87
##	[42,]	NA	259	10.9	93
##	[43,]	NA	250	9.2	92
##	[44,]	23	148	8.0	82
##	[45,]	NA	332	13.8	80
##	[46,]	NA	322	11.5	79
##	[47,]	21	191	14.9	77
##	[48,]	37	284	20.7	72
##	[49,]	20	37	9.2	65
##	[50,]	12	120	11.5	73
##	[51,]	13	137	10.3	76
##	[52,]	NA	150	6.3	77
##	[53,]	NA	59	1.7	76
##	[54,]	NA	91	4.6	76
##	[55,]	NA	250	6.3	76
##	[56,]	NA	135	8.0	75
##	[57,]	NA	127	8.0	78
##	[58,]	NA	47	10.3	73
##	[59,]	NA	98	11.5	80
##	[60,]	NA	31	14.9	77
##	[61,]	NA	138	8.0	83
##	[62,]	135	269	4.1	84
##	[63,]	49	248	9.2	85
##	[64,]	32	236	9.2	81
##	[65,]	NA	101	10.9	84
##	[66,]	64	175	4.6	83
##	[67,]	40	314	10.9	83
##	[68,]	77	276	5.1	88

##	[69,]	97	267	6.3	92
##	[70,]	97	272	5.7	92
##	[71,]	85	175	7.4	89
##	[72,]	NA	139	8.6	82
##	[73,]	10	264	14.3	73
##	[74,]	27	175	14.9	81
##	[75,]	NA	291	14.9	91
##	[76,]	7	48	14.3	80
##	[77,]	48	260	6.9	81
##	[78,]	35	274	10.3	82
##	[79,]	61	285	6.3	84
##	[80,]	79	187	5.1	87
##	[81,]	63	220	11.5	85
##	[82,]	16	7	6.9	74
##	[83,]	NA	258	9.7	81
##	[84,]	NA	295	11.5	82
##	[85,]	80	294	8.6	86
##	[86,]	108	223	8.0	85
##	[87,]	20	81	8.6	82
##	[88,]	52	82	12.0	86
##	[89,]	82	213	7.4	88
##	[90,]	50	275	7.4	86
##	[91,]	64	253	7.4	83
##	[92,]	59	254	9.2	81
##	[93,]	39	83	6.9	81
##	[94,]	9	24	13.8	81
##	[95,]	16	77	7.4	82
##	[96,]	78	NA	6.9	86
##	[97,]	35	NA	7.4	85
##	[98,]	66	NA	4.6	87
##	[99,]	122	255	4.0	89
##	[100,]	89	229	10.3	90
##	[101,]	110	207	8.0	90
##	[102,]	NA	222	8.6	92
##	[103,]	NA	137	11.5	86
##	[104,]	44	192	11.5	86
##	[105,]	28	273	11.5	82
##	[106,]	65	157	9.7	80
##	[107,]	NA	64	11.5	79
##	[108,]	22	71	10.3	77
##	[109,]	59	51	6.3	79
##	[110,]	23	115	7.4	76
##	[111,]	31	244	10.9	78
##	[112,]	44	190	10.3	78
##	[113,]	21	259	15.5	77
##	[114,]	9	36	14.3	72
##	[115,]	NA	255	12.6	75
##	[116,]	45	212	9.7	79
##	[117,]	168	238	3.4	81
##	[118,]	73	215	8.0	86
##	[119,]	NA	153	5.7	88
##	[120,]	76	203	9.7	97
##	[121,]	118	225	2.3	94
##	[122,]	84	237	6.3	96

```
## [123,] 85      188 6.3 94
## [124,] 96      167 6.9 91
## [125,] 78      197 5.1 92
## [126,] 73      183 2.8 93
## [127,] 91      189 4.6 93
## [128,] 47       95 7.4 87
## [129,] 32       92 15.5 84
## [130,] 20      252 10.9 80
## [131,] 23      220 10.3 78
## [132,] 21      230 10.9 75
## [133,] 24      259 9.7 73
## [134,] 44      236 14.9 81
## [135,] 21      259 15.5 76
## [136,] 28      238 6.3 77
## [137,] 9       24 10.9 71
## [138,] 13      112 11.5 71
## [139,] 46      237 6.9 78
## [140,] 18      224 13.8 67
## [141,] 13       27 10.3 76
## [142,] 24      238 10.3 68
## [143,] 16      201 8.0 82
## [144,] 13      238 12.6 64
## [145,] 23      14 9.2 71
## [146,] 36      139 10.3 81
## [147,] 7       49 10.3 69
## [148,] 14      20 16.6 63
## [149,] 30      193 6.9 70
## [150,] NA      145 13.2 77
## [151,] 14      191 14.3 75
## [152,] 18      131 8.0 76
## [153,] 20      223 11.5 68
```

```
class(W)
```

```
## [1] "matrix" "array"
```

2. [5%] Una part important del tractament de les dades és el seu preprocessament. El primer pas és comprovar que no hi ha cap observació sense valor en alguna o algunes variables. Una manera de fer-ho és repassar visualment el conjunt de dades. Ara bé, donat que tenim  $153 \times 6 = 612$  valors, val la pena fer-ho de forma automàtica. A R, la instrucció `is.na(W)` retornarà una matriu plena de `TRUE` (equivalent a un 1) o `FALSE` (equivalent a un zero) en funció de si falta alguna dada o no, respectivament. Finalment, si calculem la suma de tots els elements d'aquesta matriu, amb la instrucció `sum` obtindrem un 0 si no hi ha cap valor que falti o un altre nombre natural en funció de les dades que faltin. Què obteniu? Quantes dades falten?

```
is.na(W)
```

```
##      Ozone Solar.R Wind Temp
## [1,] FALSE  FALSE FALSE FALSE
## [2,] FALSE  FALSE FALSE FALSE
```

```

## [3,] FALSE FALSE FALSE FALSE
## [4,] FALSE FALSE FALSE FALSE
## [5,] TRUE TRUE FALSE FALSE
## [6,] FALSE TRUE FALSE FALSE
## [7,] FALSE FALSE FALSE FALSE
## [8,] FALSE FALSE FALSE FALSE
## [9,] FALSE FALSE FALSE FALSE
## [10,] TRUE FALSE FALSE FALSE
## [11,] FALSE TRUE FALSE FALSE
## [12,] FALSE FALSE FALSE FALSE
## [13,] FALSE FALSE FALSE FALSE
## [14,] FALSE FALSE FALSE FALSE
## [15,] FALSE FALSE FALSE FALSE
## [16,] FALSE FALSE FALSE FALSE
## [17,] FALSE FALSE FALSE FALSE
## [18,] FALSE FALSE FALSE FALSE
## [19,] FALSE FALSE FALSE FALSE
## [20,] FALSE FALSE FALSE FALSE
## [21,] FALSE FALSE FALSE FALSE
## [22,] FALSE FALSE FALSE FALSE
## [23,] FALSE FALSE FALSE FALSE
## [24,] FALSE FALSE FALSE FALSE
## [25,] TRUE FALSE FALSE FALSE
## [26,] TRUE FALSE FALSE FALSE
## [27,] TRUE TRUE FALSE FALSE
## [28,] FALSE FALSE FALSE FALSE
## [29,] FALSE FALSE FALSE FALSE
## [30,] FALSE FALSE FALSE FALSE
## [31,] FALSE FALSE FALSE FALSE
## [32,] TRUE FALSE FALSE FALSE
## [33,] TRUE FALSE FALSE FALSE
## [34,] TRUE FALSE FALSE FALSE
## [35,] TRUE FALSE FALSE FALSE
## [36,] TRUE FALSE FALSE FALSE
## [37,] TRUE FALSE FALSE FALSE
## [38,] FALSE FALSE FALSE FALSE
## [39,] TRUE FALSE FALSE FALSE
## [40,] FALSE FALSE FALSE FALSE
## [41,] FALSE FALSE FALSE FALSE
## [42,] TRUE FALSE FALSE FALSE
## [43,] TRUE FALSE FALSE FALSE
## [44,] FALSE FALSE FALSE FALSE
## [45,] TRUE FALSE FALSE FALSE
## [46,] TRUE FALSE FALSE FALSE
## [47,] FALSE FALSE FALSE FALSE
## [48,] FALSE FALSE FALSE FALSE
## [49,] FALSE FALSE FALSE FALSE
## [50,] FALSE FALSE FALSE FALSE
## [51,] FALSE FALSE FALSE FALSE
## [52,] TRUE FALSE FALSE FALSE
## [53,] TRUE FALSE FALSE FALSE
## [54,] TRUE FALSE FALSE FALSE
## [55,] TRUE FALSE FALSE FALSE
## [56,] TRUE FALSE FALSE FALSE

```



```

## [57,] TRUE FALSE FALSE FALSE
## [58,] TRUE FALSE FALSE FALSE
## [59,] TRUE FALSE FALSE FALSE
## [60,] TRUE FALSE FALSE FALSE
## [61,] TRUE FALSE FALSE FALSE
## [62,] FALSE FALSE FALSE FALSE
## [63,] FALSE FALSE FALSE FALSE
## [64,] FALSE FALSE FALSE FALSE
## [65,] TRUE FALSE FALSE FALSE
## [66,] FALSE FALSE FALSE FALSE
## [67,] FALSE FALSE FALSE FALSE
## [68,] FALSE FALSE FALSE FALSE
## [69,] FALSE FALSE FALSE FALSE
## [70,] FALSE FALSE FALSE FALSE
## [71,] FALSE FALSE FALSE FALSE
## [72,] TRUE FALSE FALSE FALSE
## [73,] FALSE FALSE FALSE FALSE
## [74,] FALSE FALSE FALSE FALSE
## [75,] TRUE FALSE FALSE FALSE
## [76,] FALSE FALSE FALSE FALSE
## [77,] FALSE FALSE FALSE FALSE
## [78,] FALSE FALSE FALSE FALSE
## [79,] FALSE FALSE FALSE FALSE
## [80,] FALSE FALSE FALSE FALSE
## [81,] FALSE FALSE FALSE FALSE
## [82,] FALSE FALSE FALSE FALSE
## [83,] TRUE FALSE FALSE FALSE
## [84,] TRUE FALSE FALSE FALSE
## [85,] FALSE FALSE FALSE FALSE
## [86,] FALSE FALSE FALSE FALSE
## [87,] FALSE FALSE FALSE FALSE
## [88,] FALSE FALSE FALSE FALSE
## [89,] FALSE FALSE FALSE FALSE
## [90,] FALSE FALSE FALSE FALSE
## [91,] FALSE FALSE FALSE FALSE
## [92,] FALSE FALSE FALSE FALSE
## [93,] FALSE FALSE FALSE FALSE
## [94,] FALSE FALSE FALSE FALSE
## [95,] FALSE FALSE FALSE FALSE
## [96,] FALSE TRUE FALSE FALSE
## [97,] FALSE TRUE FALSE FALSE
## [98,] FALSE TRUE FALSE FALSE
## [99,] FALSE FALSE FALSE FALSE
## [100,] FALSE FALSE FALSE FALSE
## [101,] FALSE FALSE FALSE FALSE
## [102,] TRUE FALSE FALSE FALSE
## [103,] TRUE FALSE FALSE FALSE
## [104,] FALSE FALSE FALSE FALSE
## [105,] FALSE FALSE FALSE FALSE
## [106,] FALSE FALSE FALSE FALSE
## [107,] TRUE FALSE FALSE FALSE
## [108,] FALSE FALSE FALSE FALSE
## [109,] FALSE FALSE FALSE FALSE
## [110,] FALSE FALSE FALSE FALSE

```

```
## [111,] FALSE FALSE FALSE FALSE
## [112,] FALSE FALSE FALSE FALSE
## [113,] FALSE FALSE FALSE FALSE
## [114,] FALSE FALSE FALSE FALSE
## [115,] TRUE FALSE FALSE FALSE
## [116,] FALSE FALSE FALSE FALSE
## [117,] FALSE FALSE FALSE FALSE
## [118,] FALSE FALSE FALSE FALSE
## [119,] TRUE FALSE FALSE FALSE
## [120,] FALSE FALSE FALSE FALSE
## [121,] FALSE FALSE FALSE FALSE
## [122,] FALSE FALSE FALSE FALSE
## [123,] FALSE FALSE FALSE FALSE
## [124,] FALSE FALSE FALSE FALSE
## [125,] FALSE FALSE FALSE FALSE
## [126,] FALSE FALSE FALSE FALSE
## [127,] FALSE FALSE FALSE FALSE
## [128,] FALSE FALSE FALSE FALSE
## [129,] FALSE FALSE FALSE FALSE
## [130,] FALSE FALSE FALSE FALSE
## [131,] FALSE FALSE FALSE FALSE
## [132,] FALSE FALSE FALSE FALSE
## [133,] FALSE FALSE FALSE FALSE
## [134,] FALSE FALSE FALSE FALSE
## [135,] FALSE FALSE FALSE FALSE
## [136,] FALSE FALSE FALSE FALSE
## [137,] FALSE FALSE FALSE FALSE
## [138,] FALSE FALSE FALSE FALSE
## [139,] FALSE FALSE FALSE FALSE
## [140,] FALSE FALSE FALSE FALSE
## [141,] FALSE FALSE FALSE FALSE
## [142,] FALSE FALSE FALSE FALSE
## [143,] FALSE FALSE FALSE FALSE
## [144,] FALSE FALSE FALSE FALSE
## [145,] FALSE FALSE FALSE FALSE
## [146,] FALSE FALSE FALSE FALSE
## [147,] FALSE FALSE FALSE FALSE
## [148,] FALSE FALSE FALSE FALSE
## [149,] FALSE FALSE FALSE FALSE
## [150,] TRUE FALSE FALSE FALSE
## [151,] FALSE FALSE FALSE FALSE
## [152,] FALSE FALSE FALSE FALSE
## [153,] FALSE FALSE FALSE FALSE
```

```
sum(is.na(W))
```

```
## [1] 44
```

```
sum(is.na(W[,1]))
```

```
## [1] 37
```

```
sum(is.na(W[,2]))
```

```
## [1] 7
```

```
sum(is.na(W[,3]))
```

```
## [1] 0
```

```
sum(is.na(W[,4]))
```

```
## [1] 0
```

**Falten 44 dades en total.**

3. [5%] Com haureu pogut veure en la pregunta anterior, falten una quantitat important de dades. El paquet *mice* proporciona una funció interessant, *md.pattern* que ens permet veure, amb un cop d'ull, el patró de dades que falten. Si carregueu aquesta llibreria i executeu les instruccions:

```
1 > library ( mice )
```

```
2 > md.pattern (W)
```

obtindreu una matriu similar a aquesta (atenció! no obtindreu la mateixa matriu):

```
1 Wind Temp Solar .R Ozone 2 98 1 1 1 1 0 3 45 1 1 1 0 1 4 6 1 1 0 1 1 5 4 1 1 0 0 2 6 0 0 10 49 59
```

Aquesta matriu s'ha d'interpretar de la següent manera:

- hi ha 98 observacions on no hi falta cap dada; - hi ha 45 observacions on hi falta la dada de la variable *Ozone*; - hi ha 6 observacions on hi falta la dada de la variable *Solar.R*; i - hi ha 4 observacions on hi falten les dades de les variables *Solar.R* i *Ozone*.

En total, faltarien 59 dades.

Per a les dades d'aquesta pràctica, quina matriu obteniu?

```
library(mice)
```

```
##
```

```
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

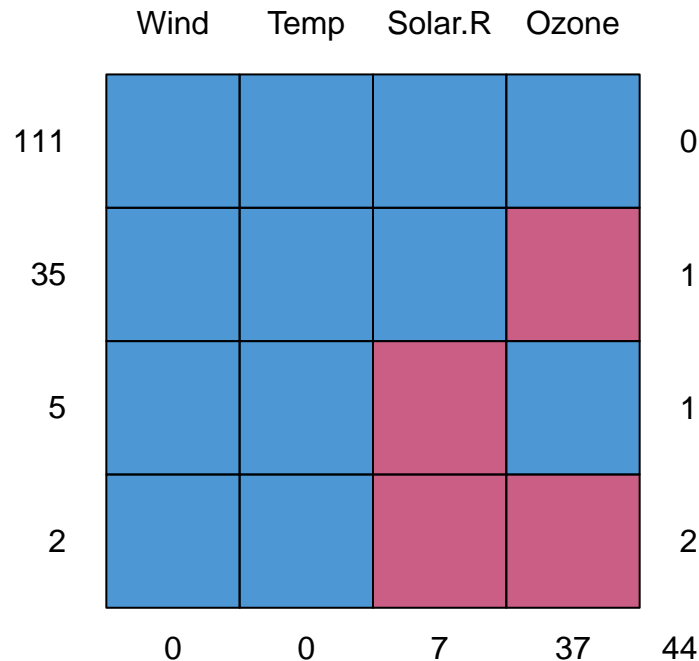
```
##      filter
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      cbind, rbind
```

```
md.pattern(W)
```



```
##      Wind Temp Solar.R Ozone
## 111    1    1      1     1  0
## 35     1    1      1     0  1
## 5      1    1      0     1  1
## 2      1    1      0     0  2
##       0    0      7    37 44
```

Falten 44 dades en total, 7 en la columna Solar.R i 37 en la columna Ozone.

4. [5%] Abans de poder continuar amb la nostra anàlisi, cal procedir a completar les dades no disponibles. Aquest procés rep el nom de *data imputation* (imputació de dades). Existeix una gran diversitat de maneres diferents d'imputar les dades que falten, com ara la mitjana aritmètica o la mediana, entre d'altres. En aquest cas, farem servir el mètode anomenat *predictive mean matching*<sup>1</sup>. Però no us preocupeu, ja que no haureu de programar aquest mètode! La funció *mice* s'encarregarà de tot, si escriviu:

```
1 > nycairi <- mice ( data = W, m = 5, method = "pmm", maxit = 50, seed = 500)
2 > W <- complete ( nycairi, 1)
```

*Noteu que hem redefinit la matriu W, que ara ja no conté dades que falten.*

```
nycairi <- mice(data = W, m = 5, method = "pmm", maxit = 50, seed = 500)
```

```
##
## iter imp variable
## 1 1 Ozone Solar.R
## 1 2 Ozone Solar.R
## 1 3 Ozone Solar.R
## 1 4 Ozone Solar.R
## 1 5 Ozone Solar.R
## 2 1 Ozone Solar.R
## 2 2 Ozone Solar.R
## 2 3 Ozone Solar.R
## 2 4 Ozone Solar.R
## 2 5 Ozone Solar.R
## 3 1 Ozone Solar.R
## 3 2 Ozone Solar.R
## 3 3 Ozone Solar.R
## 3 4 Ozone Solar.R
## 3 5 Ozone Solar.R
## 4 1 Ozone Solar.R
## 4 2 Ozone Solar.R
## 4 3 Ozone Solar.R
## 4 4 Ozone Solar.R
## 4 5 Ozone Solar.R
## 5 1 Ozone Solar.R
## 5 2 Ozone Solar.R
## 5 3 Ozone Solar.R
## 5 4 Ozone Solar.R
## 5 5 Ozone Solar.R
## 6 1 Ozone Solar.R
## 6 2 Ozone Solar.R
## 6 3 Ozone Solar.R
## 6 4 Ozone Solar.R
## 6 5 Ozone Solar.R
## 7 1 Ozone Solar.R
## 7 2 Ozone Solar.R
## 7 3 Ozone Solar.R
## 7 4 Ozone Solar.R
## 7 5 Ozone Solar.R
## 8 1 Ozone Solar.R
## 8 2 Ozone Solar.R
## 8 3 Ozone Solar.R
## 8 4 Ozone Solar.R
## 8 5 Ozone Solar.R
## 9 1 Ozone Solar.R
## 9 2 Ozone Solar.R
## 9 3 Ozone Solar.R
## 9 4 Ozone Solar.R
## 9 5 Ozone Solar.R
## 10 1 Ozone Solar.R
## 10 2 Ozone Solar.R
## 10 3 Ozone Solar.R
```

##	10	4	Ozone	Solar.R
##	10	5	Ozone	Solar.R
##	11	1	Ozone	Solar.R
##	11	2	Ozone	Solar.R
##	11	3	Ozone	Solar.R
##	11	4	Ozone	Solar.R
##	11	5	Ozone	Solar.R
##	12	1	Ozone	Solar.R
##	12	2	Ozone	Solar.R
##	12	3	Ozone	Solar.R
##	12	4	Ozone	Solar.R
##	12	5	Ozone	Solar.R
##	13	1	Ozone	Solar.R
##	13	2	Ozone	Solar.R
##	13	3	Ozone	Solar.R
##	13	4	Ozone	Solar.R
##	13	5	Ozone	Solar.R
##	14	1	Ozone	Solar.R
##	14	2	Ozone	Solar.R
##	14	3	Ozone	Solar.R
##	14	4	Ozone	Solar.R
##	14	5	Ozone	Solar.R
##	15	1	Ozone	Solar.R
##	15	2	Ozone	Solar.R
##	15	3	Ozone	Solar.R
##	15	4	Ozone	Solar.R
##	15	5	Ozone	Solar.R
##	16	1	Ozone	Solar.R
##	16	2	Ozone	Solar.R
##	16	3	Ozone	Solar.R
##	16	4	Ozone	Solar.R
##	16	5	Ozone	Solar.R
##	17	1	Ozone	Solar.R
##	17	2	Ozone	Solar.R
##	17	3	Ozone	Solar.R
##	17	4	Ozone	Solar.R
##	17	5	Ozone	Solar.R
##	18	1	Ozone	Solar.R
##	18	2	Ozone	Solar.R
##	18	3	Ozone	Solar.R
##	18	4	Ozone	Solar.R
##	18	5	Ozone	Solar.R
##	19	1	Ozone	Solar.R
##	19	2	Ozone	Solar.R
##	19	3	Ozone	Solar.R
##	19	4	Ozone	Solar.R
##	19	5	Ozone	Solar.R
##	20	1	Ozone	Solar.R
##	20	2	Ozone	Solar.R
##	20	3	Ozone	Solar.R
##	20	4	Ozone	Solar.R
##	20	5	Ozone	Solar.R
##	21	1	Ozone	Solar.R
##	21	2	Ozone	Solar.R

##	21	3	Ozone	Solar.R
##	21	4	Ozone	Solar.R
##	21	5	Ozone	Solar.R
##	22	1	Ozone	Solar.R
##	22	2	Ozone	Solar.R
##	22	3	Ozone	Solar.R
##	22	4	Ozone	Solar.R
##	22	5	Ozone	Solar.R
##	23	1	Ozone	Solar.R
##	23	2	Ozone	Solar.R
##	23	3	Ozone	Solar.R
##	23	4	Ozone	Solar.R
##	23	5	Ozone	Solar.R
##	24	1	Ozone	Solar.R
##	24	2	Ozone	Solar.R
##	24	3	Ozone	Solar.R
##	24	4	Ozone	Solar.R
##	24	5	Ozone	Solar.R
##	25	1	Ozone	Solar.R
##	25	2	Ozone	Solar.R
##	25	3	Ozone	Solar.R
##	25	4	Ozone	Solar.R
##	25	5	Ozone	Solar.R
##	26	1	Ozone	Solar.R
##	26	2	Ozone	Solar.R
##	26	3	Ozone	Solar.R
##	26	4	Ozone	Solar.R
##	26	5	Ozone	Solar.R
##	27	1	Ozone	Solar.R
##	27	2	Ozone	Solar.R
##	27	3	Ozone	Solar.R
##	27	4	Ozone	Solar.R
##	27	5	Ozone	Solar.R
##	28	1	Ozone	Solar.R
##	28	2	Ozone	Solar.R
##	28	3	Ozone	Solar.R
##	28	4	Ozone	Solar.R
##	28	5	Ozone	Solar.R
##	29	1	Ozone	Solar.R
##	29	2	Ozone	Solar.R
##	29	3	Ozone	Solar.R
##	29	4	Ozone	Solar.R
##	29	5	Ozone	Solar.R
##	30	1	Ozone	Solar.R
##	30	2	Ozone	Solar.R
##	30	3	Ozone	Solar.R
##	30	4	Ozone	Solar.R
##	30	5	Ozone	Solar.R
##	31	1	Ozone	Solar.R
##	31	2	Ozone	Solar.R
##	31	3	Ozone	Solar.R
##	31	4	Ozone	Solar.R
##	31	5	Ozone	Solar.R
##	32	1	Ozone	Solar.R

##	32	2	Ozone	Solar.R
##	32	3	Ozone	Solar.R
##	32	4	Ozone	Solar.R
##	32	5	Ozone	Solar.R
##	33	1	Ozone	Solar.R
##	33	2	Ozone	Solar.R
##	33	3	Ozone	Solar.R
##	33	4	Ozone	Solar.R
##	33	5	Ozone	Solar.R
##	34	1	Ozone	Solar.R
##	34	2	Ozone	Solar.R
##	34	3	Ozone	Solar.R
##	34	4	Ozone	Solar.R
##	34	5	Ozone	Solar.R
##	35	1	Ozone	Solar.R
##	35	2	Ozone	Solar.R
##	35	3	Ozone	Solar.R
##	35	4	Ozone	Solar.R
##	35	5	Ozone	Solar.R
##	36	1	Ozone	Solar.R
##	36	2	Ozone	Solar.R
##	36	3	Ozone	Solar.R
##	36	4	Ozone	Solar.R
##	36	5	Ozone	Solar.R
##	37	1	Ozone	Solar.R
##	37	2	Ozone	Solar.R
##	37	3	Ozone	Solar.R
##	37	4	Ozone	Solar.R
##	37	5	Ozone	Solar.R
##	38	1	Ozone	Solar.R
##	38	2	Ozone	Solar.R
##	38	3	Ozone	Solar.R
##	38	4	Ozone	Solar.R
##	38	5	Ozone	Solar.R
##	39	1	Ozone	Solar.R
##	39	2	Ozone	Solar.R
##	39	3	Ozone	Solar.R
##	39	4	Ozone	Solar.R
##	39	5	Ozone	Solar.R
##	40	1	Ozone	Solar.R
##	40	2	Ozone	Solar.R
##	40	3	Ozone	Solar.R
##	40	4	Ozone	Solar.R
##	40	5	Ozone	Solar.R
##	41	1	Ozone	Solar.R
##	41	2	Ozone	Solar.R
##	41	3	Ozone	Solar.R
##	41	4	Ozone	Solar.R
##	41	5	Ozone	Solar.R
##	42	1	Ozone	Solar.R
##	42	2	Ozone	Solar.R
##	42	3	Ozone	Solar.R
##	42	4	Ozone	Solar.R
##	42	5	Ozone	Solar.R



```
## 43 1 Ozone Solar.R
## 43 2 Ozone Solar.R
## 43 3 Ozone Solar.R
## 43 4 Ozone Solar.R
## 43 5 Ozone Solar.R
## 44 1 Ozone Solar.R
## 44 2 Ozone Solar.R
## 44 3 Ozone Solar.R
## 44 4 Ozone Solar.R
## 44 5 Ozone Solar.R
## 45 1 Ozone Solar.R
## 45 2 Ozone Solar.R
## 45 3 Ozone Solar.R
## 45 4 Ozone Solar.R
## 45 5 Ozone Solar.R
## 46 1 Ozone Solar.R
## 46 2 Ozone Solar.R
## 46 3 Ozone Solar.R
## 46 4 Ozone Solar.R
## 46 5 Ozone Solar.R
## 47 1 Ozone Solar.R
## 47 2 Ozone Solar.R
## 47 3 Ozone Solar.R
## 47 4 Ozone Solar.R
## 47 5 Ozone Solar.R
## 48 1 Ozone Solar.R
## 48 2 Ozone Solar.R
## 48 3 Ozone Solar.R
## 48 4 Ozone Solar.R
## 48 5 Ozone Solar.R
## 49 1 Ozone Solar.R
## 49 2 Ozone Solar.R
## 49 3 Ozone Solar.R
## 49 4 Ozone Solar.R
## 49 5 Ozone Solar.R
## 50 1 Ozone Solar.R
## 50 2 Ozone Solar.R
## 50 3 Ozone Solar.R
## 50 4 Ozone Solar.R
## 50 5 Ozone Solar.R
```

```
W <- complete (nycairi, 1)
W
```

```
##      Ozone Solar.R Wind Temp
## 1      41      190  7.4   67
## 2      36      118  8.0   72
## 3      12      149 12.6   74
## 4      18      313 11.5   62
## 5      14       27 14.3   56
## 6      28      150 14.9   66
## 7      23      299  8.6   65
## 8      19       99 13.8   59
## 9       8       19 20.1   61
```

## 10	24	194	8.6	69
## 11	7	7	6.9	74
## 12	16	256	9.7	69
## 13	11	290	9.2	66
## 14	14	274	10.9	68
## 15	18	65	13.2	58
## 16	14	334	11.5	64
## 17	34	307	12.0	66
## 18	6	78	18.4	57
## 19	30	322	11.5	68
## 20	11	44	9.7	62
## 21	1	8	9.7	59
## 22	11	320	16.6	73
## 23	4	25	9.7	61
## 24	32	92	12.0	61
## 25	14	66	16.6	57
## 26	4	266	14.9	58
## 27	21	49	8.0	57
## 28	23	13	12.0	67
## 29	45	252	14.9	81
## 30	115	223	5.7	79
## 31	37	279	7.4	76
## 32	40	286	8.6	78
## 33	20	287	9.7	74
## 34	9	242	16.1	67
## 35	16	186	9.2	84
## 36	48	220	8.6	85
## 37	52	264	14.3	79
## 38	29	127	9.7	82
## 39	135	273	6.9	87
## 40	71	291	13.8	90
## 41	39	323	11.5	87
## 42	168	259	10.9	93
## 43	79	250	9.2	92
## 44	23	148	8.0	82
## 45	28	332	13.8	80
## 46	32	322	11.5	79
## 47	21	191	14.9	77
## 48	37	284	20.7	72
## 49	20	37	9.2	65
## 50	12	120	11.5	73
## 51	13	137	10.3	76
## 52	59	150	6.3	77
## 53	49	59	1.7	76
## 54	32	91	4.6	76
## 55	71	250	6.3	76
## 56	29	135	8.0	75
## 57	45	127	8.0	78
## 58	32	47	10.3	73
## 59	21	98	11.5	80
## 60	13	31	14.9	77
## 61	16	138	8.0	83
## 62	135	269	4.1	84
## 63	49	248	9.2	85

## 64	32	236	9.2	81
## 65	45	101	10.9	84
## 66	64	175	4.6	83
## 67	40	314	10.9	83
## 68	77	276	5.1	88
## 69	97	267	6.3	92
## 70	97	272	5.7	92
## 71	85	175	7.4	89
## 72	32	139	8.6	82
## 73	10	264	14.3	73
## 74	27	175	14.9	81
## 75	115	291	14.9	91
## 76	7	48	14.3	80
## 77	48	260	6.9	81
## 78	35	274	10.3	82
## 79	61	285	6.3	84
## 80	79	187	5.1	87
## 81	63	220	11.5	85
## 82	16	7	6.9	74
## 83	71	258	9.7	81
## 84	71	295	11.5	82
## 85	80	294	8.6	86
## 86	108	223	8.0	85
## 87	20	81	8.6	82
## 88	52	82	12.0	86
## 89	82	213	7.4	88
## 90	50	275	7.4	86
## 91	64	253	7.4	83
## 92	59	254	9.2	81
## 93	39	83	6.9	81
## 94	9	24	13.8	81
## 95	16	77	7.4	82
## 96	78	189	6.9	86
## 97	35	47	7.4	85
## 98	66	145	4.6	87
## 99	122	255	4.0	89
## 100	89	229	10.3	90
## 101	110	207	8.0	90
## 102	135	222	8.6	92
## 103	16	137	11.5	86
## 104	44	192	11.5	86
## 105	28	273	11.5	82
## 106	65	157	9.7	80
## 107	36	64	11.5	79
## 108	22	71	10.3	77
## 109	59	51	6.3	79
## 110	23	115	7.4	76
## 111	31	244	10.9	78
## 112	44	190	10.3	78
## 113	21	259	15.5	77
## 114	9	36	14.3	72
## 115	24	255	12.6	75
## 116	45	212	9.7	79
## 117	168	238	3.4	81

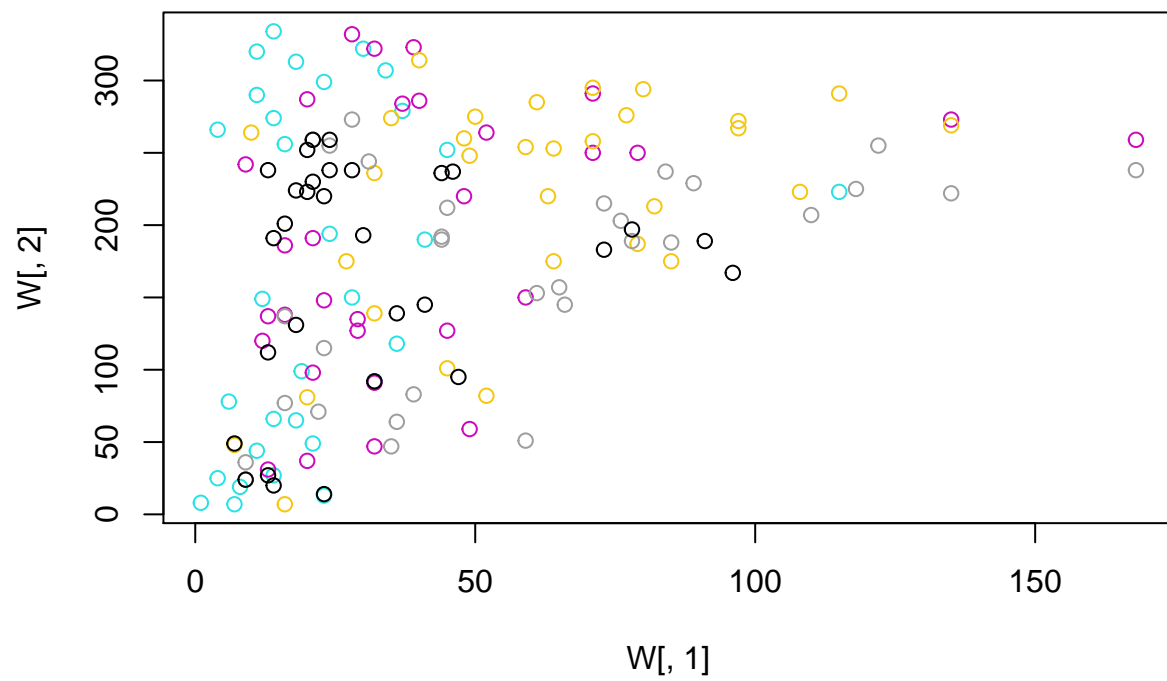
```
## 118    73    215  8.0  86
## 119    61    153  5.7  88
## 120    76    203  9.7  97
## 121   118    225  2.3  94
## 122    84    237  6.3  96
## 123    85    188  6.3  94
## 124    96    167  6.9  91
## 125    78    197  5.1  92
## 126    73    183  2.8  93
## 127    91    189  4.6  93
## 128    47     95  7.4  87
## 129    32     92 15.5  84
## 130    20    252 10.9  80
## 131    23    220 10.3  78
## 132    21    230 10.9  75
## 133    24    259  9.7  73
## 134    44    236 14.9  81
## 135    21    259 15.5  76
## 136    28    238  6.3  77
## 137     9     24 10.9  71
## 138    13    112 11.5  71
## 139    46    237  6.9  78
## 140    18    224 13.8  67
## 141    13     27 10.3  76
## 142    24    238 10.3  68
## 143    16    201  8.0  82
## 144    13    238 12.6  64
## 145    23     14  9.2  71
## 146    36    139 10.3  81
## 147     7     49 10.3  69
## 148    14     20 16.6  63
## 149    30    193  6.9  70
## 150    41    145 13.2  77
## 151    14    191 14.3  75
## 152    18    131  8.0  76
## 153    20    223 11.5  68
```

```
W[5,2]
```

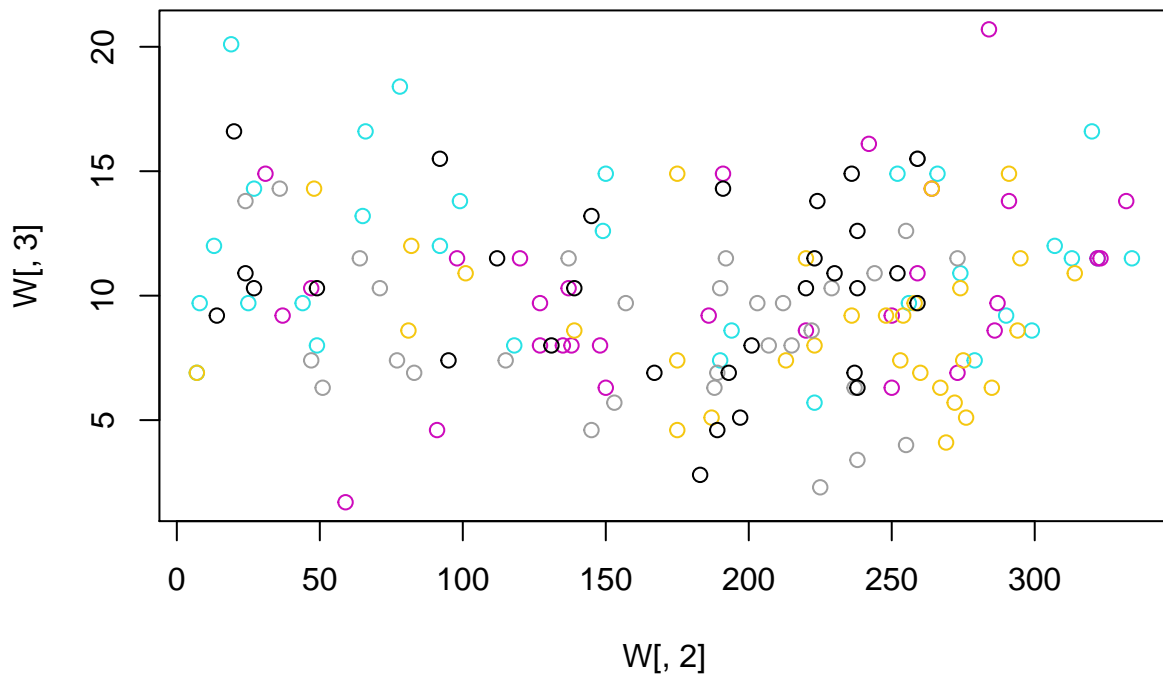
```
## [1] 27
```

5. [5%] Representeu en un diagrama de dispersió, a mode d'exemple, la primera variable (Ozone) respecte de la segona (Solar.R). Representeu també la segona variable (Solar.R) respecte de la tercera (Wind). Feu servir la instrucció `plot` amb les opcions `type = "p"` (punts) i `col = label`. L'opció `col = label` us permetrà veure, de color diferent, les observacions dels diferents mesos. Què observeu? Les classes se superposen o es poden separar clarament?

```
plot (W[,1], W[,2], type = "p", col = label)
```



```
plot (W[,2], W[,3], type = "p", col = label)
```



6. [5%] Escaleu les dades de la matriu  $W$  tal com s'explica a la Secció 2.1 del mòdul. Anomeneu a la matriu resultant  $Ws$ . Comproveu que la mitjana aritmètica de les quatre columnes de la matriu  $Ws$  és zero (de fet, obtindreu un valor molt petit, proper al zero de màquina) i que la desviació tipus de les quatre columnes és 1. Podeu fer servir les instruccions `mean` i `sd`. Quina d'aquestes dues instruccions us ajudarà a resoldre aquesta pregunta:

- Aquesta? `apply(Ws,1,mean)` - O aquesta? `apply(Ws,2,mean)`

Justifiqueu la resposta.

```
Ws <- scale (W, center = TRUE, scale = TRUE)
Ws
```

```
##           Ozone      Solar.R      Wind      Temp
## [1,] -0.06341481  0.09362416 -0.72594816 -1.14971398
## [2,] -0.20953628 -0.69367642 -0.55563883 -0.62146702
## [3,] -0.91091932 -0.35469978  0.75006604 -0.41016823
## [4,] -0.73557356  1.43859599  0.43783226 -1.67796094
## [5,] -0.85247074 -1.68873688  1.23260914 -2.31185730
## [6,] -0.44333063 -0.34376505  1.40291847 -1.25536337
## [7,] -0.58945209  1.28550977 -0.38532950 -1.36101276
## [8,] -0.70634927 -0.90143630  1.09068470 -1.99490912
## [9,] -1.02781650 -1.77621473  2.87893266 -1.78361034
```

```

## [10,] -0.56022780  0.13736308 -0.38532950 -0.93841519
## [11,] -1.05704079 -1.90743149 -0.86787260 -0.41016823
## [12,] -0.79402215  0.81531636 -0.07309573 -0.93841519
## [13,] -0.94014362  1.18709720 -0.21502017 -1.25536337
## [14,] -0.85247074  1.01214151  0.26752293 -1.04406459
## [15,] -0.73557356 -1.27321713  0.92037537 -2.10055851
## [16,] -0.85247074  1.66822533  0.43783226 -1.46666216
## [17,] -0.26798486  1.37298761  0.57975671 -1.25536337
## [18,] -1.08626509 -1.13106564  2.39638956 -2.20620791
## [19,] -0.38488204  1.53700857  0.43783226 -1.04406459
## [20,] -0.94014362 -1.50284647 -0.07309573 -1.67796094
## [21,] -1.23238655 -1.89649676 -0.07309573 -1.99490912
## [22,] -0.94014362  1.51513911  1.88546157 -0.51581762
## [23,] -1.14471367 -1.71060634 -0.07309573 -1.78361034
## [24,] -0.32643345 -0.97797941  0.57975671 -1.78361034
## [25,] -0.85247074 -1.26228240  1.88546157 -2.20620791
## [26,] -1.14471367  0.92466367  1.40291847 -2.10055851
## [27,] -0.64790068 -1.44817282 -0.55563883 -2.20620791
## [28,] -0.58945209 -1.84182311  0.57975671 -1.14971398
## [29,]  0.05348237  0.77157744  1.40291847  0.32937752
## [30,]  2.09918293  0.45447026 -1.20849126  0.11807873
## [31,] -0.18031198  1.06681516 -0.72594816 -0.19886945
## [32,] -0.09263910  1.14335827 -0.38532950  0.01242934
## [33,] -0.67712497  1.15429300 -0.07309573 -0.41016823
## [34,] -0.99859221  0.66223014  1.74353713 -1.14971398
## [35,] -0.79402215  0.04988524 -0.21502017  0.64632570
## [36,]  0.14115525  0.42166607 -0.38532950  0.75197509
## [37,]  0.25805242  0.90279421  1.23260914  0.11807873
## [38,] -0.41410633 -0.59526385 -0.07309573  0.43502691
## [39,]  2.68366880  1.00120678 -0.86787260  0.96327387
## [40,]  0.81331400  1.19803193  1.09068470  1.28022205
## [41,] -0.12186339  1.54794330  0.43783226  0.96327387
## [42,]  3.64807049  0.84812055  0.26752293  1.59717023
## [43,]  1.04710835  0.74970798 -0.21502017  1.49152084
## [44,] -0.58945209 -0.36563451 -0.55563883  0.43502691
## [45,] -0.44333063  1.64635587  1.09068470  0.22372813
## [46,] -0.32643345  1.53700857  0.43783226  0.11807873
## [47,] -0.64790068  0.10455889  1.40291847 -0.09322005
## [48,] -0.18031198  1.12148881  3.04924199 -0.62146702
## [49,] -0.67712497 -1.57938958 -0.21502017 -1.36101276
## [50,] -0.91091932 -0.67180696  0.43783226 -0.51581762
## [51,] -0.88169503 -0.48591655  0.09721360 -0.19886945
## [52,]  0.46262248 -0.34376505 -1.03818193 -0.09322005
## [53,]  0.17037954 -1.33882551 -2.34388679 -0.19886945
## [54,] -0.32643345 -0.98891414 -1.52072503 -0.19886945
## [55,]  0.81331400  0.74970798 -1.03818193 -0.19886945
## [56,] -0.41410633 -0.50778601 -0.55563883 -0.30451884
## [57,]  0.05348237 -0.59526385 -0.55563883  0.01242934
## [58,] -0.32643345 -1.47004228  0.09721360 -0.51581762
## [59,] -0.64790068 -0.91237103  0.43783226  0.22372813
## [60,] -0.88169503 -1.64499796  1.40291847 -0.09322005
## [61,] -0.79402215 -0.47498182 -0.55563883  0.54067630
## [62,]  2.68366880  0.95746786 -1.66264947  0.64632570
## [63,]  0.17037954  0.72783852 -0.21502017  0.75197509

```

```

## [64,] -0.32643345  0.59662176 -0.21502017  0.32937752
## [65,]  0.05348237 -0.87956684  0.26752293  0.64632570
## [66,]  0.60874395 -0.07039679 -1.52072503  0.54067630
## [67,] -0.09263910  1.44953072  0.26752293  0.54067630
## [68,]  0.98865977  1.03401097 -1.37880059  1.06892327
## [69,]  1.57314564  0.93559840 -1.03818193  1.49152084
## [70,]  1.57314564  0.99027205 -1.20849126  1.49152084
## [71,]  1.22245412 -0.07039679 -0.72594816  1.17457266
## [72,] -0.32643345 -0.46404709 -0.38532950  0.43502691
## [73,] -0.96936791  0.90279421  1.23260914 -0.51581762
## [74,] -0.47255492 -0.07039679  1.40291847  0.32937752
## [75,]  2.09918293  1.19803193  1.40291847  1.38587145
## [76,] -1.05704079 -1.45910755  1.23260914  0.22372813
## [77,]  0.14115525  0.85905529 -0.86787260  0.32937752
## [78,] -0.23876057  1.01214151  0.09721360  0.43502691
## [79,]  0.52107107  1.13242354 -1.03818193  0.64632570
## [80,]  1.04710835  0.06081997 -1.37880059  0.96327387
## [81,]  0.57951965  0.42166607  0.43783226  0.75197509
## [82,] -0.79402215 -1.90743149 -0.86787260 -0.41016823
## [83,]  0.81331400  0.83718582 -0.07309573  0.32937752
## [84,]  0.81331400  1.24177085  0.43783226  0.43502691
## [85,]  1.07633265  1.23083612 -0.38532950  0.85762448
## [86,]  1.89461287  0.45447026 -0.55563883  0.75197509
## [87,] -0.67712497 -1.09826144 -0.38532950  0.43502691
## [88,]  0.25805242 -1.08732671  0.57975671  0.85762448
## [89,]  1.13478123  0.34512296 -0.72594816  1.06892327
## [90,]  0.19960384  1.02307624 -0.72594816  0.85762448
## [91,]  0.60874395  0.78251217 -0.72594816  0.54067630
## [92,]  0.46262248  0.79344690 -0.21502017  0.32937752
## [93,] -0.12186339 -1.07639198 -0.86787260  0.32937752
## [94,] -0.99859221 -1.72154107  1.09068470  0.32937752
## [95,] -0.79402215 -1.14200037 -0.72594816  0.43502691
## [96,]  1.01788406  0.08268943 -0.86787260  0.85762448
## [97,] -0.23876057 -1.47004228 -0.72594816  0.75197509
## [98,]  0.66719253 -0.39843870 -1.52072503  0.96327387
## [99,]  2.30375298  0.80438163 -1.69103436  1.17457266
## [100,] 1.33935129  0.52007864  0.09721360  1.28022205
## [101,] 1.95306146  0.27951458 -0.55563883  1.28022205
## [102,] 2.68366880  0.44353553 -0.38532950  1.49152084
## [103,] -0.79402215 -0.48591655  0.43783226  0.85762448
## [104,]  0.02425807  0.11549362  0.43783226  0.85762448
## [105,] -0.44333063  1.00120678  0.43783226  0.43502691
## [106,]  0.63796824 -0.26722194 -0.07309573  0.22372813
## [107,] -0.20953628 -1.28415186  0.43783226  0.11807873
## [108,] -0.61867639 -1.20760875  0.09721360 -0.09322005
## [109,]  0.46262248 -1.42630335 -1.03818193  0.11807873
## [110,] -0.58945209 -0.72648061 -0.72594816 -0.19886945
## [111,] -0.35565774  0.68409960  0.26752293  0.01242934
## [112,]  0.02425807  0.09362416  0.09721360  0.01242934
## [113,] -0.64790068  0.84812055  1.57322780 -0.09322005
## [114,] -0.99859221 -1.59032431  1.23260914 -0.62146702
## [115,] -0.56022780  0.80438163  0.75006604 -0.30451884
## [116,]  0.05348237  0.33418823 -0.07309573  0.11807873
## [117,]  3.64807049  0.61849122 -1.86134369  0.32937752

```



```
## [118,] 0.87176259 0.36699242 -0.55563883 0.85762448
## [119,] 0.52107107 -0.31096086 -1.20849126 1.06892327
## [120,] 0.95943547 0.23577566 -0.07309573 2.01976780
## [121,] 2.18685581 0.47633972 -2.17357746 1.70281962
## [122,] 1.19322982 0.60755649 -1.03818193 1.91411841
## [123,] 1.22245412 0.07175470 -1.03818193 1.70281962
## [124,] 1.54392135 -0.15787464 -0.86787260 1.38587145
## [125,] 1.01788406 0.17016727 -1.37880059 1.49152084
## [126,] 0.87176259 0.01708105 -2.03165302 1.59717023
## [127,] 1.39779988 0.08268943 -1.52072503 1.59717023
## [128,] 0.11193095 -0.94517522 -0.72594816 0.96327387
## [129,] -0.32643345 -0.97797941 1.57322780 0.64632570
## [130,] -0.67712497 0.77157744 0.26752293 0.22372813
## [131,] -0.58945209 0.42166607 0.09721360 0.01242934
## [132,] -0.64790068 0.53101338 0.26752293 -0.30451884
## [133,] -0.56022780 0.84812055 -0.07309573 -0.51581762
## [134,] 0.02425807 0.59662176 1.40291847 0.32937752
## [135,] -0.64790068 0.84812055 1.57322780 -0.19886945
## [136,] -0.44333063 0.61849122 -1.03818193 -0.09322005
## [137,] -0.99859221 -1.72154107 0.26752293 -0.72711641
## [138,] -0.88169503 -0.75928480 0.43783226 -0.72711641
## [139,] 0.08270666 0.60755649 -0.86787260 0.01242934
## [140,] -0.73557356 0.46540499 1.09068470 -1.14971398
## [141,] -0.88169503 -1.68873688 0.09721360 -0.19886945
## [142,] -0.56022780 0.61849122 0.09721360 -1.04406459
## [143,] -0.79402215 0.21390620 -0.55563883 0.43502691
## [144,] -0.88169503 0.61849122 0.75006604 -1.46666216
## [145,] -0.58945209 -1.83088838 -0.21502017 -0.72711641
## [146,] -0.20953628 -0.46404709 0.09721360 0.32937752
## [147,] -1.05704079 -1.44817282 0.09721360 -0.93841519
## [148,] -0.85247074 -1.76528000 1.88546157 -1.57231155
## [149,] -0.38488204 0.12642835 -0.86787260 -0.83276580
## [150,] -0.06341481 -0.39843870 0.92037537 -0.09322005
## [151,] -0.85247074 0.10455889 1.23260914 -0.30451884
## [152,] -0.73557356 -0.55152493 -0.55563883 -0.19886945
## [153,] -0.67712497 0.45447026 0.43783226 -1.04406459
## attr("scaled:center")
##      Ozone      Solar.R      Wind      Temp
## 43.169935 181.437908   9.957516 77.882353
## attr("scaled:scale")
##      Ozone      Solar.R      Wind      Temp
## 34.218107 91.451729   3.523001 9.465270
```

```
apply (Ws, 2, mean)
```

```
##      Ozone      Solar.R      Wind      Temp
## 4.535933e-18 -9.082957e-17 -3.732756e-17 6.973588e-16
```

```
apply (Ws, 2, sd)
```

```
##      Ozone Solar.R      Wind      Temp
##      1      1      1      1
```

```
Ws[3,2]
```

```
##      Solar.R  
## -0.3546998
```

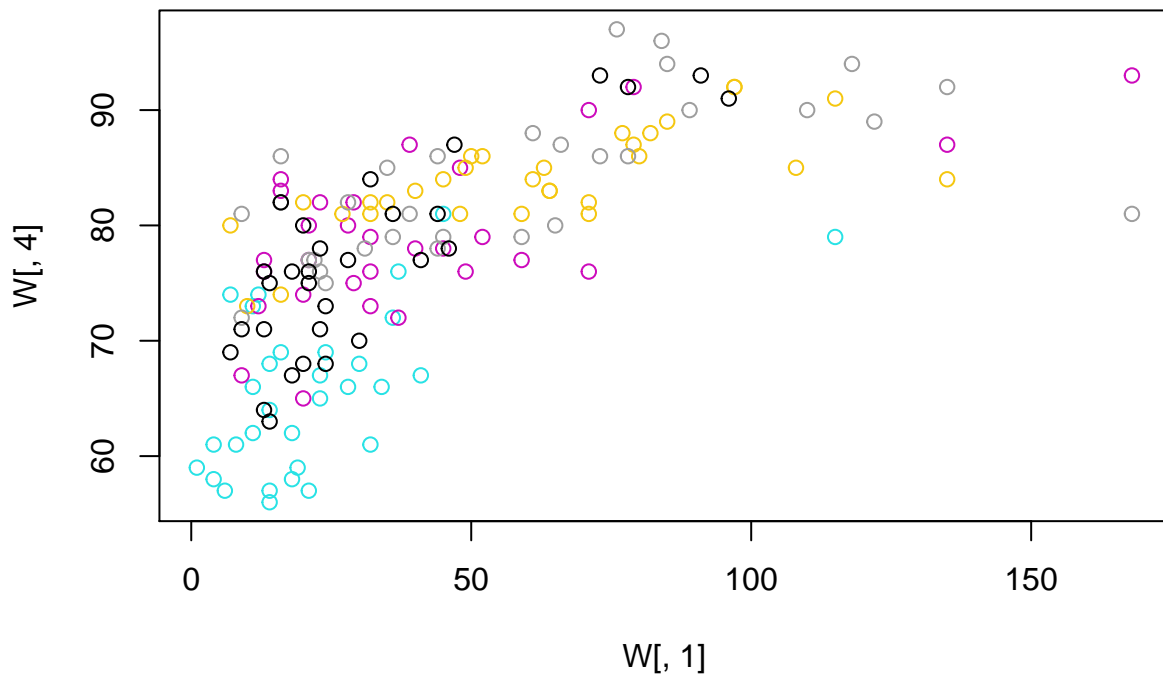
Les instruccions adequades són “`apply (Ws, 2, mean)`” i “`apply (Ws, 2, sd)`”, ja que la variable 2 es refereix a realitzar el càlcul sobre les columnes de la matriu.

7. [10%] Calculeu la matriu de covariàncies de *Ws* tal i com s’explica a la Secció 2.2 del mòdul. Anomeneu *CWs* a la matriu de covariàncies. Quin parell de variables (diferents) presenten la covariància més gran, en valor absolut? Representeu en un diagrama de dispersió aquestes dues variables, amb les mateixes opcions gràfiques que en les preguntes anteriors.

```
CWs <- t(Ws)%*%Ws/152  
CWs
```

```
##           Ozone      Solar.R      Wind      Temp  
## Ozone      1.0000000  0.36117111 -0.50019441  0.6827107  
## Solar.R    0.3611711  1.00000000 -0.04266829  0.2990714  
## Wind      -0.5001944 -0.04266829  1.00000000 -0.4579879  
## Temp       0.6827107  0.29907143 -0.45798788  1.0000000
```

```
plot(W[,1], W[,4], type = "p", col = label)
```



```
sum(CWs)
```

```
## [1] 4.684205
```

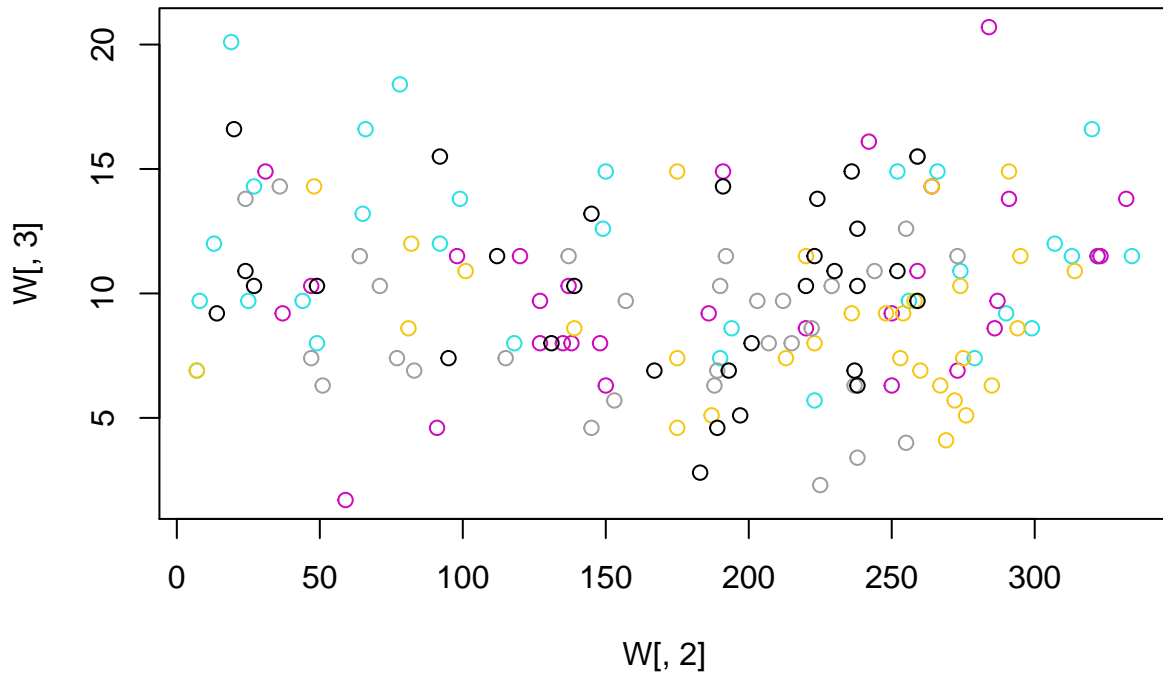
El parell de variables que presenten una covariància més gran són la 1 “Ozone” i la 4 “Temp”.

8. [10%] Quin parell de variables (diferents) presenten la covariància més petita, en valor absolut? Representeu en un diagrama de dispersió aquestes dues variables, amb les mateixes opcions gràfiques que en les preguntes anteriors.

```
CWs
```

```
##           Ozone      Solar.R      Wind      Temp
## Ozone      1.0000000  0.3611711 -0.5001944  0.6827107
## Solar.R    0.3611711  1.0000000 -0.04266829 0.2990714
## Wind      -0.5001944 -0.04266829  1.00000000 -0.4579879
## Temp       0.6827107  0.29907143 -0.45798788  1.0000000
```

```
plot(W[,2], W[,3], type = "p", col = label)
```



El parell de variables que presenten una covariància més petita són la 2 “Solar R.” i la 3 “Wind”.

9. [10%] Calculeu, sense fer servir la comanda `prcomp`, els valors i els vectors propis de la matriu de covariàncies *CWs*. Anomenau *P* a la matriu que conté, per columnes, els vectors propis. Aneu al qüestionari associat a la pràctica i mireu quin és el valor de *N* que us ha estat assignat. Quina és la variabilitat retinguda per les primeres *N* components principals?

```
vv_prop <- eigen(CWs)
vv_prop
```

```
## eigen() decomposition
## $values
## [1] 2.2451176 0.9610280 0.4871210 0.3067334
##
```

```
## $vectors
##           [,1]           [,2]           [,3]           [,4]
## [1,] -0.5939832 -0.00645786 -0.2276217  0.7715767
## [2,] -0.3258124 -0.83372749  0.4257567 -0.1321965
## [3,]  0.4606955 -0.55083922 -0.6796878  0.1495334
## [4,] -0.5734020  0.03785390 -0.5522173 -0.6040143
```

```
P <- vv_prop$vectors
P
```

```
##           [,1]           [,2]           [,3]           [,4]
## [1,] -0.5939832 -0.00645786 -0.2276217  0.7715767
## [2,] -0.3258124 -0.83372749  0.4257567 -0.1321965
## [3,]  0.4606955 -0.55083922 -0.6796878  0.1495334
## [4,] -0.5734020  0.03785390 -0.5522173 -0.6040143
```

```
class (P)
```

```
## [1] "matrix" "array"
```

```
variabilitat_N1 <- (vv_prop$values[1]/4)*100
variabilitat_N2 <- (vv_prop$values[2]/4)*100
variabilitat_N3 <- (vv_prop$values[3]/4)*100
variabilitat_N1_3 <- variabilitat_N1 + variabilitat_N2 + variabilitat_N3

variabilitat_N1_3
```

```
## [1] 92.33167
```

**Les 3 primeres components principals retenen una variabilitat del 92.33167%.**

10. [10%] La primera component principal, PC1, és un vector format per la combinació lineal de les quatre variables originals:

$$PC1 = a_1 \times \text{Ozone} + a_2 \times \text{Solar.R} + a_3 \times \text{Wind} + a_4 \times \text{Temp}$$

Els coeficients  $a_i$ ;  $i = 1; 2; 3; 4$ , en valor absolut, són també una mesura de la importància de cada variable original amb l'objectiu de la classificació de les dades originals. Des d'aquest punt de vista, quines són les dues variables més importants en la primera component principal? I en la segona component principal? Justifiqueu la vostra resposta.

```
PC1 <- P[,1]
PC2 <- P[,2]
PC1
```

```
## [1] -0.5939832 -0.3258124  0.4606955 -0.5734020
```

PC2

```
## [1] -0.00645786 -0.83372749 -0.55083922  0.03785390
```

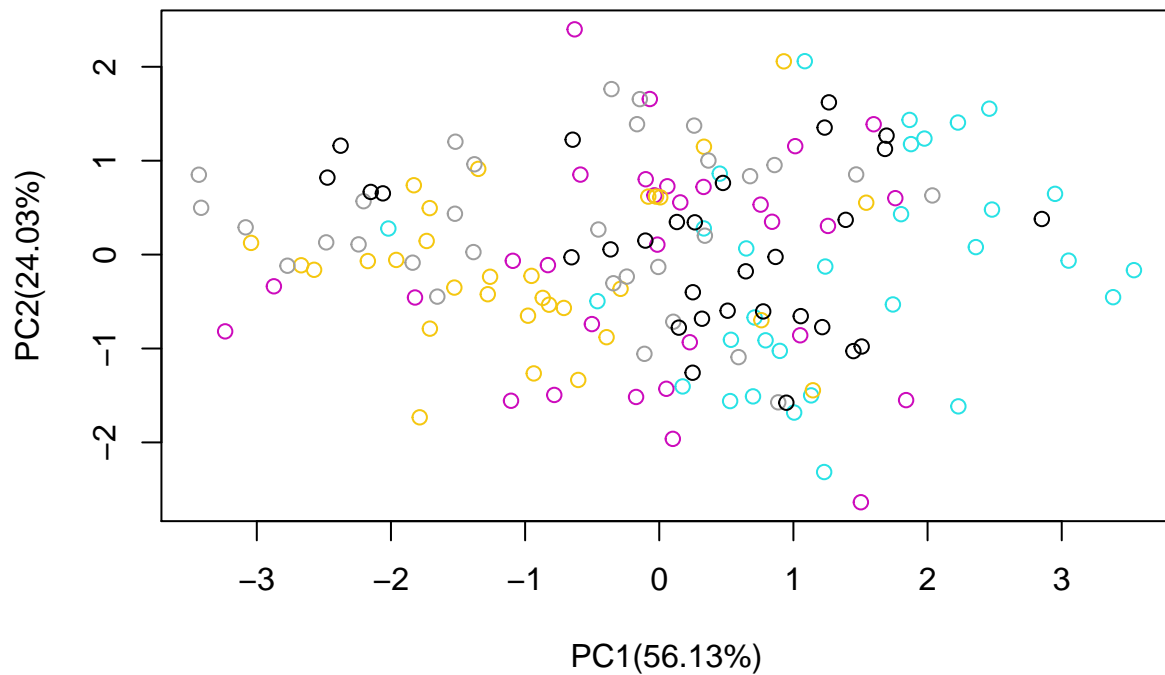
Les dues variables més importants en la primera component principal són, la [u1] “Ozone” i la [u4] “Temp”. Pel que respecta a la segona component principal són, la [u2] “Solar R.” i la [u3] “Wind”.

11. [10%] Representeu gràficament els 153 punts formats per la projecció de les dades originals normalitzades sobre la primera i la segona components principals. Com abans, feu servir la instrucció `plot` amb les opcions `type = "p"` (punts) i `col = label`. Què podeu dir ara mateix sobre les classes (mesos de l'any)? Se superposen o es poden separar clarament? És possible que, tot i haver fet l'anàlisi de components principals, encara no se separin els punts.

Aquest fet està explicat a l'article:

Pozo, Francesc; Vidal, Yolanda. 2016. “Wind Turbine Fault Detection Through Principal Component Analysis and Statistical Hypothesis Testing”. *Energies* 9, no. 1: 3. <https://doi.org/10.3390/en9010003> on es demostra que PCA no és suficient per separar i cal fer un pas més que, en aquest cas, és l'ús del contrast d'hipòtesi.

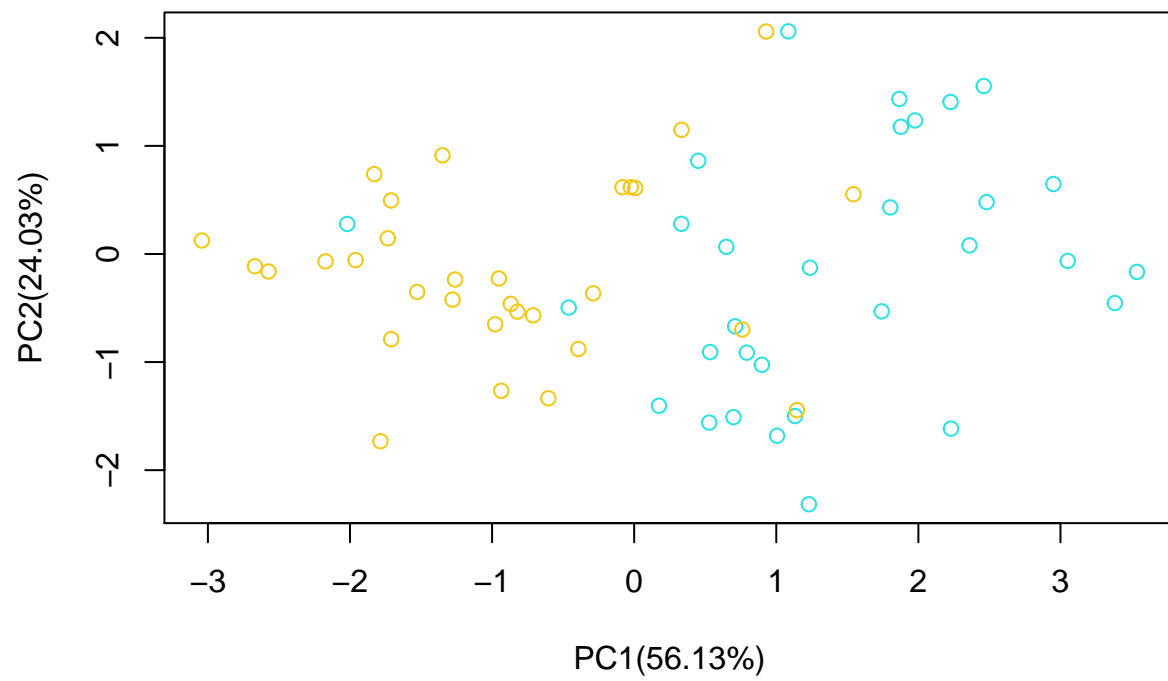
```
T <- Ws%*%P
plot(T[,1], T[,2], type = "p", col = label, xlab = "PC1(56.13%)", ylab = "PC2(24.03%)")
```



12. [10%] Representeu gràficament només els 31 punts corresponents al mes de maig i els 31 punts corresponents al mes de juliol formats per la projecció de les dades originals normalitzades sobre la primera i la segona components principals. Com abans, feu servir la instrucció `plot` amb les opcions `type = "p"` (punts) i `col = label[c(1:31,62:92)]`. Què podeu dir ara?

*Som capaços de distingir el mes de maig i el mes de juliol?*

```
plot(T[c(1:31,62:92), 1], T[c(1:31,62:92),2], type = "p", col = label[c(1:31,62:92)],
      xlab = "PC1(56.13%)", ylab = "PC2(24.03%)")
```



Les dades corresponents al mes de maig i al mes de juliol són perfectament distingibles.