Ejercicio 3.3

Silvia Pineda

Carga de Datos y Librerías

```
library(naniar)
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.1     v stringr   1.5.2
v ggplot2   4.0.0     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

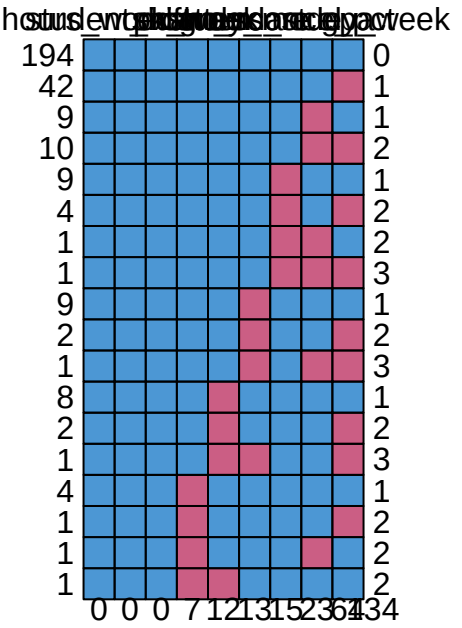
```
library(rio)

data <- read.csv("students_FP.csv",
  na.strings = c("", "NA", "NaN", "NULL"),
  stringsAsFactors = TRUE
)
```

Imputación múltiple

```
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(data)
```



	student_id	hours_work_week	shift	program	exam_score	study_mode	
194	1	1	1	1	1	1	0
42	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1
10	1	1	1	1	1	1	2
9	1	1	1	1	1	1	1
4	1	1	1	1	1	1	2
1	1	1	1	1	1	1	2
	0	0	0	7	12	13	34

```
1              1                 1      1         1            1            1
9              1                 1      1         1            1            0
2              1                 1      1         1            1            0
1              1                 1      1         1            1            0
8              1                 1      1         1            0            1
2              1                 1      1         1            0            1
1              1                 1      1         1            0            0
4              1                 1      1         0            1            1
1              1                 1      1         0            1            1
1              1                 1      1         0            1            1
1              1                 1      1         0            0            1
               0                 0      0         7           12           13
    attendance_pct hours_study_week gpa
194              1                 1   1   0
42               1                 1   0   1
9                1                 0   1   1
10               1                 0   0   2
9                0                 1   1   1
4                0                 1   0   2
1                0                 0   1   2
1                0                 0   0   3
9                1                 1   1   1
2                1                 1   0   2
1                1                 0   0   3
8                1                 1   1   1
2                1                 1   0   2
1                1                 1   0   3
4                1                 1   1   1
1                1                 1   0   2
1                1                 0   1   2
1                1                 1   1   2
                15                23  64 134
```

No hay ninguna observación con todas las variables missing.

```
impData <- mice(select(data,-student_id),m=5,maxit=50,seed=500)
```

```
 iter imp variable
  1   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  1   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  1   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
```

1	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
1	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
2	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
2	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
2	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
2	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
2	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
3	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
3	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
3	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
3	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
3	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
4	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
4	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
4	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
4	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
4	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
5	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
5	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
5	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
5	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
5	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
6	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
6	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
6	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
6	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
6	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
7	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
7	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
7	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
7	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
7	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
8	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
8	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
8	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
8	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
8	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
9	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
9	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
9	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
9	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
9	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
10	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode

10	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
10	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
10	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
10	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
11	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
11	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
11	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
11	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
11	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
12	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
12	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
12	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
12	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
12	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
13	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
13	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
13	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
13	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
13	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
14	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
14	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
14	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
14	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
14	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
15	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
15	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
15	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
15	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
15	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
16	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
16	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
16	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
16	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
16	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
17	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
17	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
17	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
17	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
17	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
18	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
18	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
18	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
18	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode

18	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
19	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
19	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
19	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
19	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
19	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
20	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
20	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
20	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
20	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
20	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
21	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
21	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
21	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
21	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
21	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
22	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
22	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
22	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
22	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
22	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
23	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
23	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
23	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
23	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
23	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
24	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
24	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
24	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
24	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
24	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
25	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
25	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
25	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
25	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
25	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
26	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
26	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
26	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
26	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
26	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
27	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
27	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode

27	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
27	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
27	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
28	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
28	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
28	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
28	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
28	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
29	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
29	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
29	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
29	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
29	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
30	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
30	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
30	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
30	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
30	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
31	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
31	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
31	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
31	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
31	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
32	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
32	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
32	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
32	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
32	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
33	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
33	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
33	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
33	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
33	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
34	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
34	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
34	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
34	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
34	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
35	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
35	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
35	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
35	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
35	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode

36	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
36	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
36	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
36	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
36	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
37	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
37	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
37	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
37	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
37	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
38	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
38	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
38	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
38	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
38	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
39	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
39	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
39	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
39	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
39	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
40	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
40	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
40	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
40	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
40	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
41	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
41	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
41	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
41	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
41	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
42	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
42	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
42	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
42	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
42	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
43	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
43	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
43	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
43	4	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
43	5	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
44	1	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
44	2	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode
44	3	hours_study_week	attendance_pct	gpa	exam_score	program	study_mode

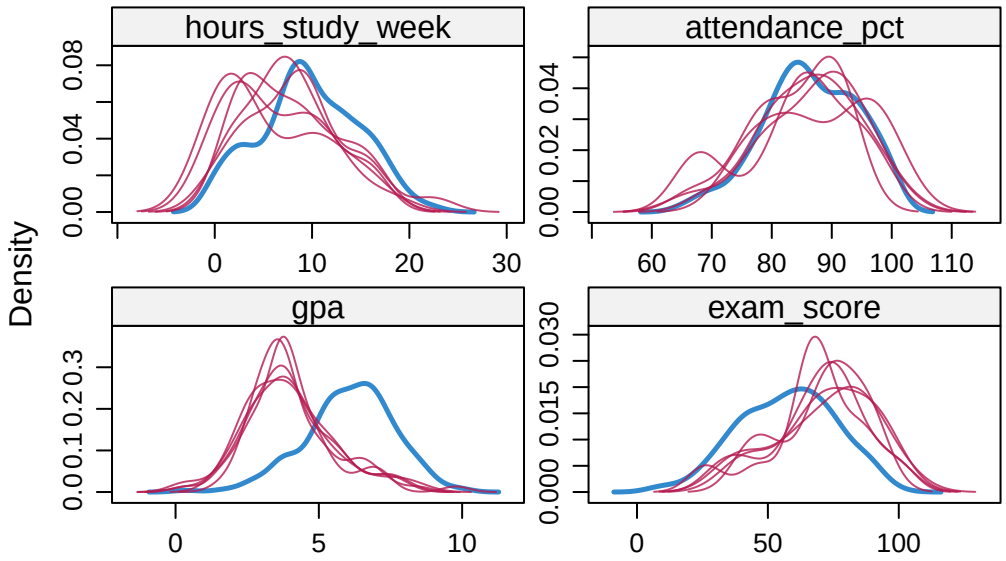
```
  44   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  44   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  45   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  45   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  45   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  45   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  45   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  46   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  46   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  46   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  46   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  46   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  47   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  47   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  47   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  47   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  47   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  48   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  48   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  48   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  48   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  48   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  49   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  49   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  49   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  49   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  49   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  50   1  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  50   2  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  50   3  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  50   4  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
  50   5  hours_study_week  attendance_pct  gpa  exam_score  program  study_mode
```

```
summary(impData)
```

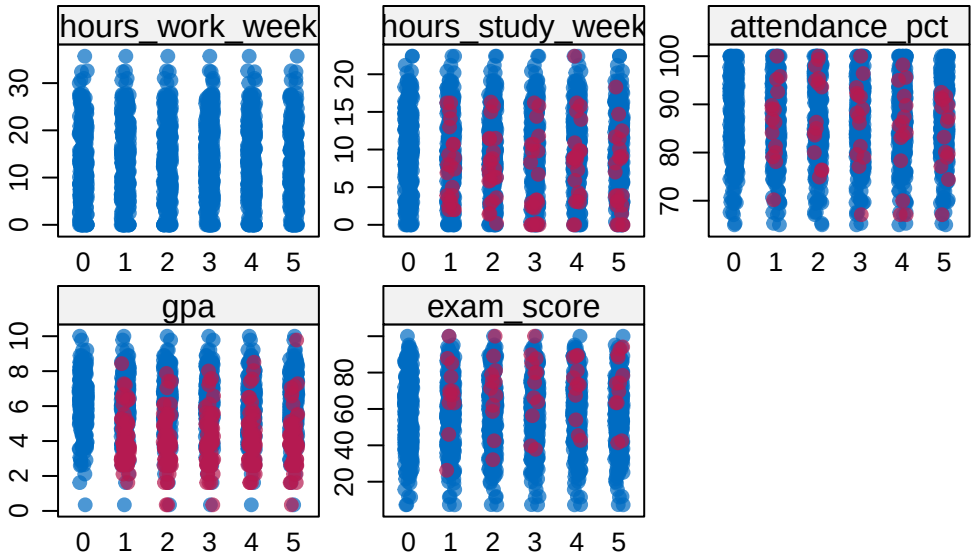
```
Class: mids
Number of multiple imputations:  5
Imputation methods:
 hours_work_week hours_study_week   attendance_pct              gpa
              ""            "pmm"            "pmm"           "pmm"
      exam_score          program       study_mode            shift
           "pmm"        "polyreg"        "polyreg"               ""
```

```
PredictorMatrix:
                 hours_work_week hours_study_week attendance_pct gpa exam_score
hours_work_week                0                1              1   1          1
hours_study_week               1                0              1   1          1
attendance_pct                 1                1              0   1          1
gpa                            1                1              1   0          1
exam_score                     1                1              1   1          0
program                        1                1              1   1          1
                 program study_mode shift
hours_work_week        1          1     1
hours_study_week       1          1     1
attendance_pct         1          1     1
gpa                    1          1     1
exam_score             1          1     1
program                0          1     1
```

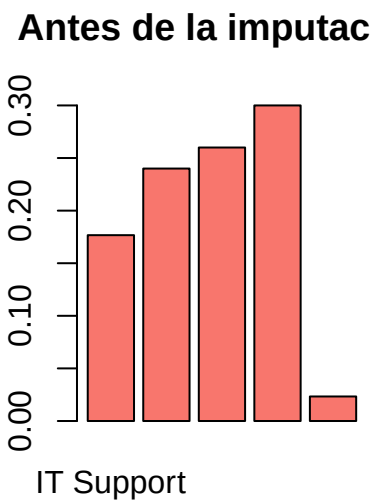
```r
###Visualizar las variables cuantitativas
densityplot(impData)
```

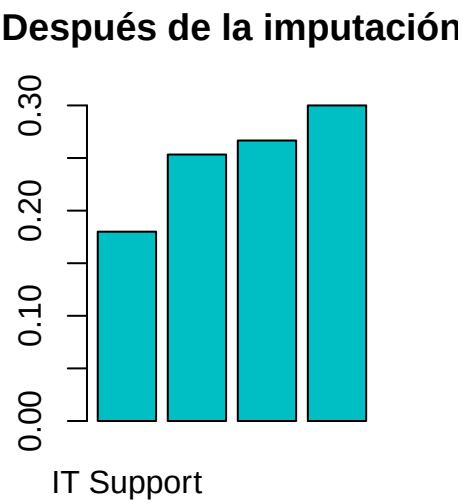



```r
stripplot(impData, pch = 20, cex = 1.2)
```

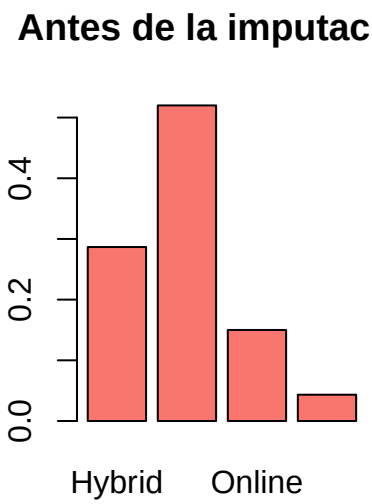


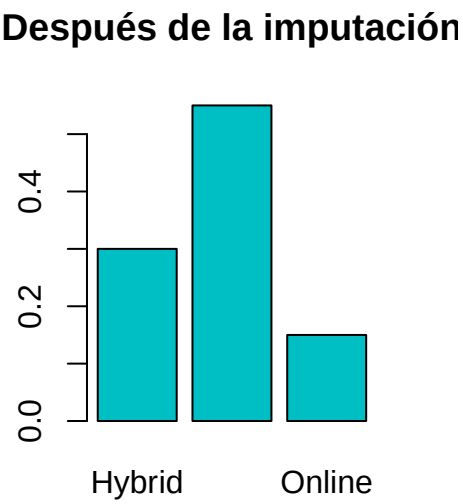

```r
###Visualizar las variables cualitativas
completedData <- complete(impData,1)
par(mfrow = c(1, 2))
barplot(prop.table(table(data$program, useNA = "ifany")),
        main = "Antes de la imputación", col = "#F8766D" )
barplot(prop.table(table(completedData$program)),
        main = "Después de la imputación",col = "#00BFC4")
```






```
par(mfrow = c(1, 2))
barplot(prop.table(table(data$study_mode, useNA = "ifany")),
        main = "Antes de la imputación", col = "#F8766D" )
barplot(prop.table(table(completedData$study_mode)),
        main = "Después de la imputación",col = "#00BFC4")
```





Excepto la variable attendance_pct, ninguna variable se ha imputado demasiado bien, pero la que se ha imputado muy mal es la de gpa que todos los valores se han imputado por debajo del resto de la distribución, esto no necesariamente está mal, precisamente una hipótesis era que los datos faltantes de gpa era MNAR porque podrías estar relacionados con valores bajode de gpa, pero no explicado por el resto de variables. En este caso, al tener en cuenta de forma multivariante toda la base de datos, ha imputado los valores de la variable como bajos desplazando así toda la distribución.

Por otro lado las dos variables cualitativas parece que se imputa de forma correcta.

Imputación con missForest

```
library(missForest)

set.seed(123)
data_missForest<-select(data,-student_id)
imp <- missForest(data_missForest)
imp$OOBerror
```

```
    NRMSE       PFC 
0.1441779 0.4063217 
```

```r
set.seed(123)
imp <- missForest(data_missForest,variablewise = TRUE)
imp$OOBerror
```

```
       MSE        MSE        MSE        MSE        MSE        PFC        PFC
 0.0000000 13.1874025 52.6990576  0.4157744 51.1884986  0.6928328  0.5261324
       PFC
 0.0000000
```

```r
# Calcular la Standard Deviation para normalizar solo en las cuantitativas
num_vars <- names(data_missForest)[sapply(data_missForest, is.numeric)]
id<-match(num_vars,names(data_missForest))

mse_num <- imp$OOBerror[id]
sd_num <- sapply(data_missForest[id], sd, na.rm = TRUE)

# Calcular el NMRSE
NRMSE <- sqrt(mse_num) / sd_num
names(NRMSE)<-colnames(data_missForest[id])
NRMSE
```

```
 hours_work_week hours_study_week   attendance_pct              gpa
       0.0000000        0.7111160        0.9260997        0.4147127
      exam_score
       0.3855255
```

De forma global los dos errores tanto para las cuantitativas como para las cualitativas, los errores son pequeños:

NRMSE = 0.14

PFC = 0.41

Si lo sacamos de forma individual por variable, para las cualitativas tenemos:

program (PFC = 0.69)

study_mode (PFC = 0.53)

hours_study_week (NRMSE = 0.71)

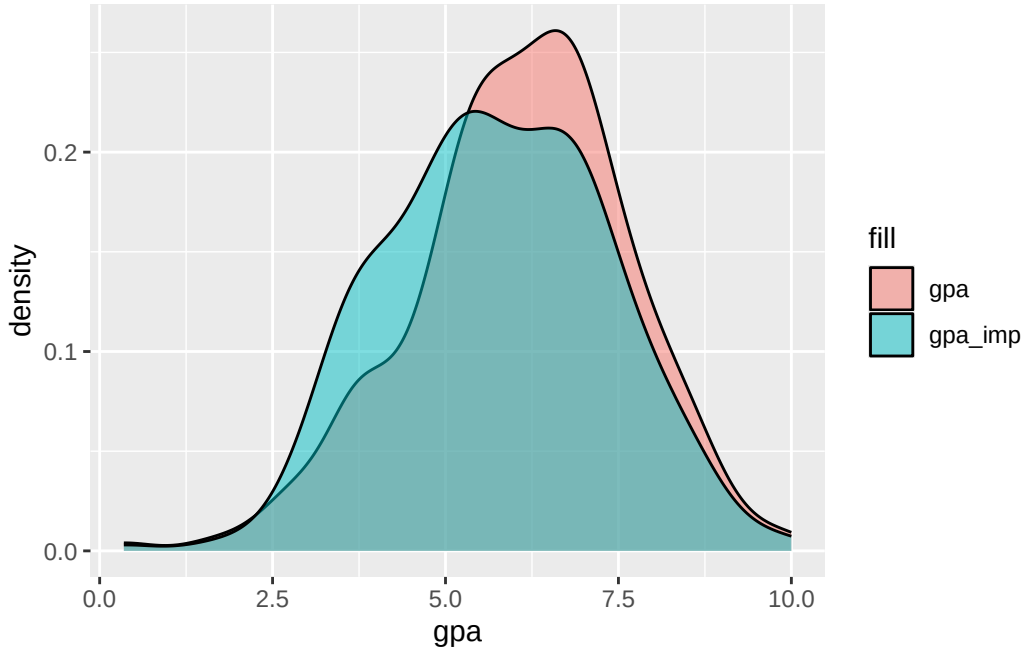
attendance_pct (NRMSE = 0.93)

gpa (NRMSE = 0.41)

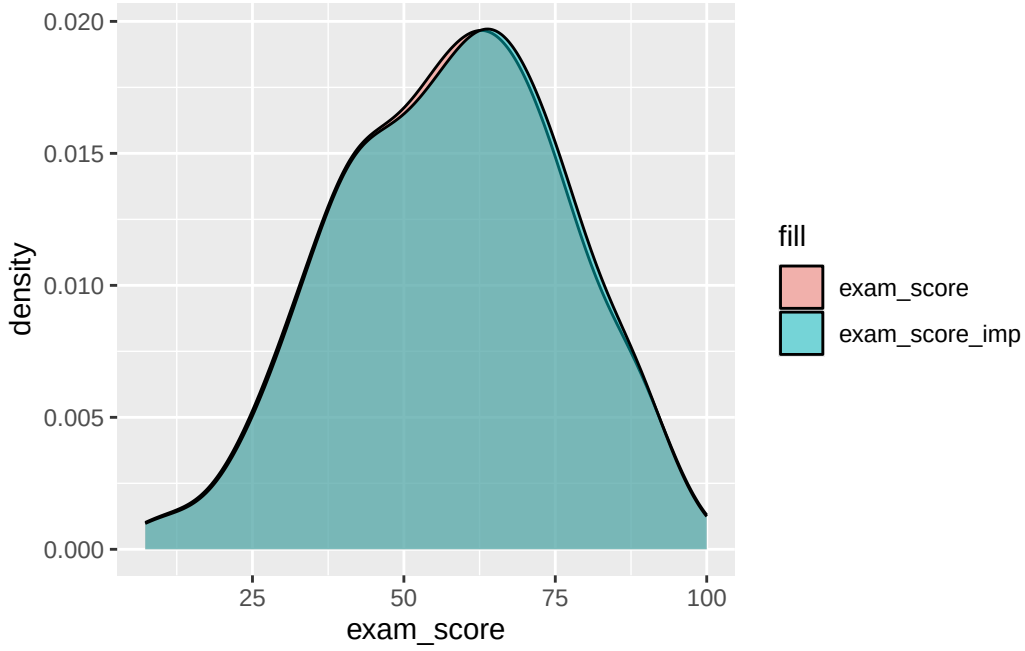
exam_score (NRMSE = 0.38)

Ninguno parece demasiado alto, pero es curioso que gpa sea uno de los más bajos a diferencia del resto de imputaciones.

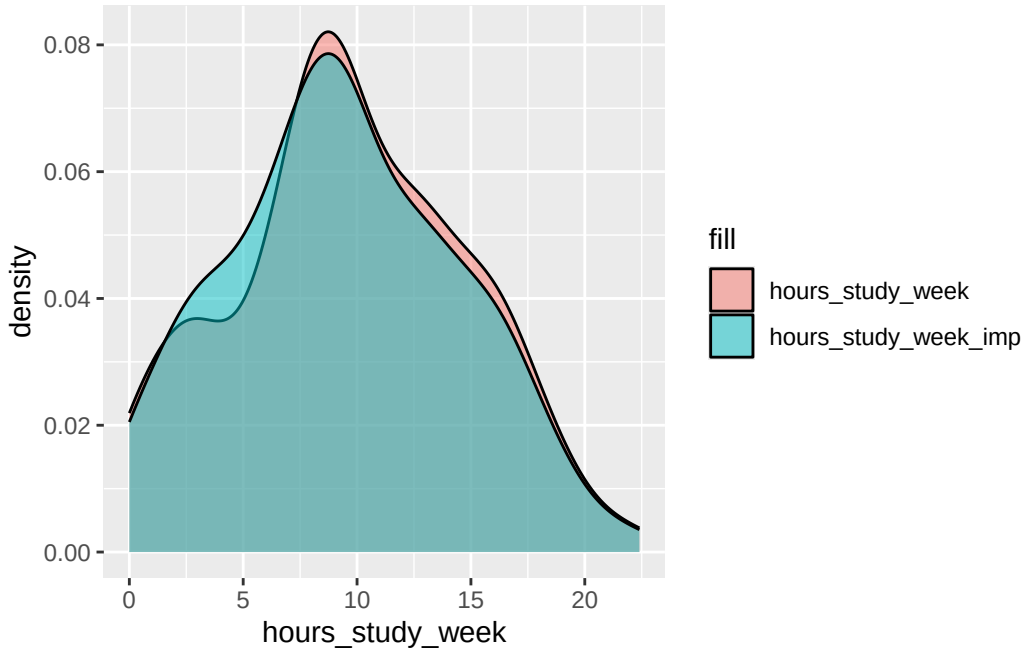
```
g1<-ggplot(data, aes(x =gpa, fill = "gpa")) +
  geom_density(alpha = 0.5, na.rm = TRUE) +
  geom_density(aes(x = imp$ximp$gpa,
                   fill = "gpa_imp"), alpha = 0.5)

g2<-ggplot(data, aes(x = exam_score, fill = "exam_score")) +
  geom_density(alpha = 0.5, na.rm = TRUE) +
  geom_density(aes(x = imp$ximp$exam_score,
                   fill = "exam_score_imp"),alpha = 0.5)

g3<-ggplot(data, aes(x = hours_study_week,
                     fill = "hours_study_week")) +
  geom_density(alpha = 0.5, na.rm = TRUE) +
  geom_density(aes(x = imp$ximp$hours_study_week,
                   fill = "hours_study_week_imp"), alpha = 0.5)

g4<-ggplot(data, aes(x = attendance_pct,
                     fill = "attendance_pct")) +
  geom_density(alpha = 0.5, na.rm = TRUE) +
  geom_density(aes(x = imp$ximp$attendance_pct,
                   fill = "attendance_pct_imp"), alpha = 0.5)

g1
```


density
0.2
0.1
0.0
0.0
2.5
5.0
7.5
10.0
gpa
fill
gpa
gpa_imp

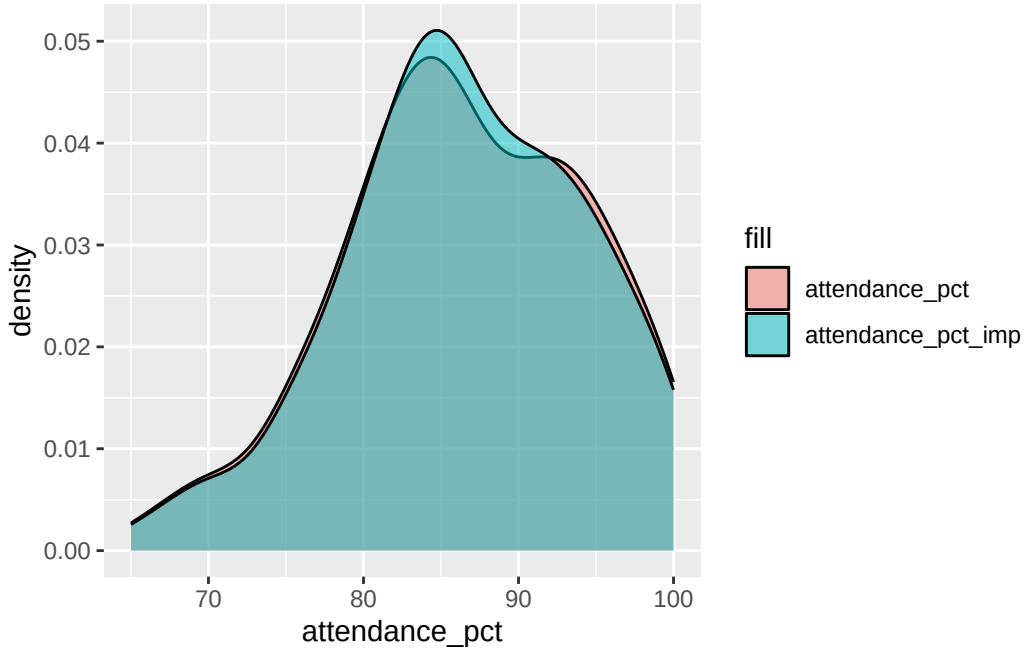
```
g2
```


density
0.020
0.015
0.010
0.005
0.000
25
50
75
100
exam_score
fill
exam_score
exam_score_imp

```
g3
```


0.08
0.06
0.04
0.02
0.00
density
0
5
10
15
20
hours_study_week
fill
hours_study_week
hours_study_week_imp

```
g4
```


0.05
0.04
0.03
0.02
0.01
0.00
density
70
80
90
100
attendance_pct
fill
attendance_pct
attendance_pct_imp

El gráfico de gpa es el que peor se ajusta a su distribución, pero mejora mucho a lo que veíamos con el resto de imputaciones.