```python
# This Python 3 environment comes with many helpful analytics
libraries installed
# It is defined by the kaggle/python Docker image:
https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import seaborn as sns
import matplotlib.pyplot as plt
# Input data files are available in the read-only "../input/"
directory
# For example, running this (by clicking run or pressing Shift+Enter)
will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/)
that gets preserved as output when you create a version using "Save &
Run All"
# You can also write temporary files to /kaggle/temp/, but they won't
be saved outside of the current session

data = pd.read_csv("student_data.csv")
data
```

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob
0	GP	F	18	U	GT3	A	4	4	at_home	teacher
1	GP	F	17	U	GT3	T	1	1	at_home	other
2	GP	F	15	U	LE3	T	1	1	at_home	other
3	GP	F	15	U	GT3	T	4	2	health	services
4	GP	F	16	U	GT3	T	3	3	other	other
..	...	..	...	...	...	...	...	...	...	...
390	MS	M	20	U	LE3	A	2	2	services	services
391	MS	M	17	U	LE3	T	3	1	services	services
392	MS	M	21	R	GT3	T	1	1	other	other
393	MS	M	18	R	LE3	T	3	2	services	

```
other
394      MS   M   19        U       LE3        T      1      1      other
at_home

     ... famrel freetime   goout   Dalc   Walc health absences   G1  G2
G3
0    ...      4        3       4      1      1       3        6    5   6
6
1    ...      5        3       3      1      1       3        4    5   5
6
2    ...      4        3       2      2      3       3       10    7   8
10
3    ...      3        2       2      1      1       5        2   15  14
15
4    ...      4        3       2      1      2       5        4    6  10
10
..   ...    ...      ...     ...    ...    ...     ...      ...   ..  ..  .
.
390  ...      5        5       4      4      5       4       11    9   9
9
391  ...      2        4       5      3      4       2        3   14  16
16
392  ...      5        5       3      3      3       3        3   10   8
7
393  ...      4        4       1      3      4       5        0   11  12
10
394  ...      3        2       3      3      3       5        5    8   9
9

[395 rows x 33 columns]

data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 395 entries, 0 to 394
Data columns (total 33 columns):
 #   Column        Non-Null Count   Dtype
---  ------        --------------   -----
 0   school        395 non-null     object
 1   sex           395 non-null     object
 2   age           395 non-null     int64
 3   address       395 non-null     object
 4   famsize       395 non-null     object
 5   Pstatus       395 non-null     object
 6   Medu          395 non-null     int64
 7   Fedu          395 non-null     int64
 8   Mjob          395 non-null     object
 9   Fjob          395 non-null     object
 10  reason        395 non-null     object
 11  guardian      395 non-null     object
```

```
 12   traveltime   395 non-null    int64
 13   studytime    395 non-null    int64
 14   failures     395 non-null    int64
 15   schoolsup    395 non-null    object
 16   famsup       395 non-null    object
 17   paid         395 non-null    object
 18   activities   395 non-null    object
 19   nursery      395 non-null    object
 20   higher       395 non-null    object
 21   internet     395 non-null    object
 22   romantic     395 non-null    object
 23   famrel       395 non-null    int64
 24   freetime     395 non-null    int64
 25   goout        395 non-null    int64
 26   Dalc         395 non-null    int64
 27   Walc         395 non-null    int64
 28   health       395 non-null    int64
 29   absences     395 non-null    int64
 30   G1           395 non-null    int64
 31   G2           395 non-null    int64
 32   G3           395 non-null    int64
dtypes: int64(16), object(17)
memory usage: 102.0+ KB
```

data.describe()

```
              age        Medu        Fedu    traveltime    studytime
failures  \
count  395.000000  395.000000  395.000000  395.000000  395.000000
395.000000
mean    16.696203    2.749367    2.521519    1.448101    2.035443
0.334177
std      1.276043    1.094735    1.088201    0.697505    0.839240
0.743651
min     15.000000    0.000000    0.000000    1.000000    1.000000
0.000000
25%     16.000000    2.000000    2.000000    1.000000    1.000000
0.000000
50%     17.000000    3.000000    2.000000    1.000000    2.000000
0.000000
75%     18.000000    4.000000    3.000000    2.000000    2.000000
0.000000
max     22.000000    4.000000    4.000000    4.000000    4.000000
3.000000

           famrel     freetime       goout         Dalc         Walc
health  \
count  395.000000  395.000000  395.000000  395.000000  395.000000
395.000000
mean     3.944304    3.235443    3.108861    1.481013    2.291139
```

```
                                                                                      3.554430
std       0.896659       0.998862       1.113278       0.890741       1.287897
1.390303
min       1.000000       1.000000       1.000000       1.000000       1.000000
1.000000
25%       4.000000       3.000000       2.000000       1.000000       1.000000
3.000000
50%       4.000000       3.000000       3.000000       1.000000       2.000000
4.000000
75%       5.000000       4.000000       4.000000       2.000000       3.000000
5.000000
max       5.000000       5.000000       5.000000       5.000000       5.000000
5.000000

          absences            G1            G2            G3
count   395.000000    395.000000    395.000000    395.000000
mean      5.708861     10.908861     10.713924     10.415190
std       8.003096      3.319195      3.761505      4.581443
min       0.000000      3.000000      0.000000      0.000000
25%       0.000000      8.000000      9.000000      8.000000
50%       4.000000     11.000000     11.000000     11.000000
75%       8.000000     13.000000     13.000000     14.000000
max      75.000000     19.000000     19.000000     20.000000
```

```
data.columns
```

```
Index(['school', 'sex', 'age', 'address', 'famsize', 'Pstatus',
'Medu', 'Fedu',
       'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime',
'studytime',
       'failures', 'schoolsup', 'famsup', 'paid', 'activities',
'nursery',
       'higher', 'internet', 'romantic', 'famrel', 'freetime',
'goout', 'Dalc',
       'Walc', 'health', 'absences', 'G1', 'G2', 'G3'],
      dtype='object')
```

Counting Number of Entities

```
data['school'].value_counts()
```

```
school
GP     349
MS      46
Name: count, dtype: int64
```

```
data['sex'].value_counts()
```

```
sex
F    208
M    187
Name: count, dtype: int64
```

```
data['age'].value_counts()
```

```
age
16    104
17     98
18     82
15     82
19     24
20      3
22      1
21      1
Name: count, dtype: int64
```

```
data['famsize'].value_counts()
```

```
famsize
GT3    281
LE3    114
Name: count, dtype: int64
```

```
data['traveltime'].value_counts()
```

```
traveltime
1    257
2    107
3     23
4      8
Name: count, dtype: int64
```

```
data['studytime'].value_counts()
```

```
studytime
2    198
1    105
3     65
4     27
Name: count, dtype: int64
```

```
data['failures'].value_counts()
```

```
failures
0    312
1     50
2     17
3     16
Name: count, dtype: int64
```

```
data['Mjob'].value_counts()

Mjob
other       141
services    103
at_home      59
teacher      58
health       34
Name: count, dtype: int64

data['Fjob'].value_counts()

Fjob
other       217
services    111
teacher      29
at_home      20
health       18
Name: count, dtype: int64
```

Failure

```
Fail_0 = data.where(data['failures']==0)
Fail_0 = Fail_0.dropna()
Fail_0

     school sex   age address famsize Pstatus  Medu  Fedu      Mjob
Fjob  \
0       GP   F  18.0       U     GT3       A   4.0   4.0   at_home
teacher
1       GP   F  17.0       U     GT3       T   1.0   1.0   at_home
other
3       GP   F  15.0       U     GT3       T   4.0   2.0    health
services
4       GP   F  16.0       U     GT3       T   3.0   3.0     other
other
5       GP   M  16.0       U     LE3       T   4.0   3.0  services
other
..      ...  ..   ...     ...     ...     ...   ...   ...       ...
...
386     MS   F  18.0       R     GT3       T   4.0   4.0   teacher
at_home
388     MS   F  18.0       U     LE3       T   3.0   1.0   teacher
services
391     MS   M  17.0       U     LE3       T   3.0   1.0  services
services
393     MS   M  18.0       R     LE3       T   3.0   2.0  services
other
```

```
394     MS   M   19.0        U       LE3       T   1.0   1.0      other
at_home

       ...  famrel freetime  goout   Dalc   Walc health absences      G1
G2     G3
0      ...      4.0      3.0    4.0    1.0    1.0      3.0       6.0    5.0
6.0    6.0
1      ...      5.0      3.0    3.0    1.0    1.0      3.0       4.0    5.0
5.0    6.0
3      ...      3.0      2.0    2.0    1.0    1.0      5.0       2.0   15.0
14.0   15.0
4      ...      4.0      3.0    2.0    1.0    2.0      5.0       4.0    6.0
10.0   10.0
5      ...      5.0      4.0    2.0    1.0    2.0      5.0      10.0   15.0
15.0   15.0
..     ...      ...      ...    ...    ...    ...      ...       ...    ...   ..
.    ...
386    ...      4.0      4.0    3.0    2.0    2.0      5.0       7.0    6.0
5.0    6.0
388    ...      4.0      3.0    4.0    1.0    1.0      1.0       0.0    7.0
9.0    8.0
391    ...      2.0      4.0    5.0    3.0    4.0      2.0       3.0   14.0
16.0   16.0
393    ...      4.0      4.0    1.0    3.0    4.0      5.0       0.0   11.0
12.0   10.0
394    ...      3.0      2.0    3.0    3.0    3.0      5.0       5.0    8.0
9.0    9.0

[312 rows x 33 columns]

Fail_1 = data.where(data['failures']==1)
Fail_1 = Fail_1.dropna()
Fail_1

    school sex   age address famsize Pstatus  Medu  Fedu      Mjob
Fjob  \
40      GP   F  16.0       U     LE3       T   2.0   2.0     other
other
44      GP   F  16.0       U     LE3       T   2.0   2.0     other
at_home
49      GP   F  15.0       U     GT3       T   4.0   4.0  services
teacher
52      GP   M  15.0       U     LE3       A   4.0   2.0    health
health
88      GP   M  16.0       U     GT3       T   2.0   2.0  services
other
95      GP   F  15.0       R     GT3       T   1.0   1.0   at_home
other
111     GP   F  16.0       R     GT3       T   3.0   3.0  services
other
```

112	GP	F	16.0	U	GT3	T	2.0	2.0	at_home	other
118	GP	M	17.0	R	GT3	T	1.0	3.0	other	other
138	GP	M	16.0	U	LE3	T	1.0	1.0	services	other
151	GP	M	16.0	U	LE3	T	2.0	1.0	at_home	other
159	GP	M	16.0	U	GT3	T	3.0	3.0	other	services
162	GP	M	16.0	U	LE3	T	1.0	2.0	other	other
165	GP	M	16.0	U	GT3	T	3.0	2.0	services	services
198	GP	F	17.0	U	GT3	T	4.0	4.0	services	teacher
205	GP	F	17.0	U	GT3	T	3.0	4.0	at_home	services
213	GP	M	18.0	U	GT3	T	2.0	2.0	services	other
217	GP	M	18.0	U	LE3	T	3.0	3.0	services	health
221	GP	F	17.0	U	GT3	T	1.0	1.0	at_home	other
225	GP	F	18.0	R	GT3	T	3.0	1.0	other	other
239	GP	M	18.0	U	GT3	T	2.0	2.0	other	services
248	GP	M	18.0	R	LE3	T	3.0	3.0	other	services
250	GP	M	18.0	U	GT3	T	3.0	2.0	services	other
252	GP	M	18.0	U	GT3	T	2.0	1.0	services	services
255	GP	M	17.0	U	LE3	T	1.0	1.0	health	other
278	GP	F	18.0	U	GT3	T	4.0	4.0	health	health
281	GP	M	17.0	U	LE3	A	3.0	2.0	teacher	services
292	GP	F	18.0	U	LE3	T	2.0	1.0	services	at_home
304	GP	M	19.0	U	GT3	T	3.0	3.0	other	other
305	GP	F	18.0	U	GT3	T	2.0	4.0	services	at_home
307	GP	M	19.0	U	GT3	T	4.0	4.0	teacher	services
308	GP	M	19.0	R	GT3	T	3.0	3.0	other	

services
```
309     GP  F  19.0     U     LE3      T  1.0  1.0   at_home
other
310     GP  F  19.0     U     LE3      T  1.0  2.0  services
services
312     GP  M  19.0     U     GT3      T  1.0  2.0     other
services
313     GP  F  19.0     U     LE3      T  3.0  2.0  services
other
315     GP  F  19.0     R     GT3      T  2.0  3.0     other
other
336     GP  F  19.0     R     GT3      A  3.0  1.0  services
at_home
340     GP  F  19.0     U     GT3      T  2.0  1.0  services
services
341     GP  M  18.0     U     GT3      T  4.0  4.0   teacher
services
343     GP  F  17.0     U     GT3      A  2.0  2.0   at_home
at_home
349     MS  M  18.0     R     GT3      T  3.0  2.0     other
other
352     MS  M  18.0     U     LE3      T  1.0  3.0   at_home
services
353     MS  M  19.0     R     GT3      T  1.0  1.0     other
other
361     MS  M  18.0     R     LE3      T  1.0  1.0   at_home
other
367     MS  F  17.0     R     GT3      T  1.0  1.0     other
services
383     MS  M  19.0     R     GT3      T  1.0  1.0     other
services
384     MS  M  18.0     R     GT3      T  4.0  2.0     other
other
387     MS  F  19.0     R     GT3      T  2.0  3.0  services
other
389     MS  F  18.0     U     GT3      T  1.0  1.0     other
other

     ... famrel  freetime  goout  Dalc  Walc  health  absences    G1
G2    G3
40   ...    3.0       3.0    3.0   1.0   2.0     3.0      25.0   7.0
10.0  11.0
44   ...    4.0       3.0    3.0   2.0   2.0     5.0      14.0  10.0
10.0   9.0
49   ...    4.0       4.0    4.0   1.0   1.0     3.0       2.0   7.0
7.0   7.0
52   ...    5.0       5.0    5.0   3.0   4.0     5.0       6.0  11.0
11.0  10.0
88   ...    4.0       4.0    2.0   1.0   1.0     3.0      12.0  11.0
```

									10.0	10.0
95	...	3.0	1.0	2.0	1.0	1.0	1.0	2.0	7.0	10.0 10.0
111	...	4.0	1.0	2.0	1.0	1.0	2.0	0.0	7.0	10.0 10.0
112	...	3.0	1.0	2.0	1.0	1.0	5.0	6.0	10.0	13.0 13.0
118	...	5.0	2.0	4.0	1.0	4.0	5.0	20.0	9.0	7.0 8.0
138	...	4.0	4.0	4.0	1.0	3.0	5.0	0.0	14.0	12.0 12.0
151	...	4.0	4.0	4.0	3.0	5.0	5.0	6.0	12.0	13.0 14.0
159	...	4.0	5.0	5.0	4.0	4.0	5.0	4.0	10.0	12.0 12.0
162	...	4.0	4.0	4.0	2.0	4.0	5.0	0.0	7.0	0.0 0.0
165	...	4.0	5.0	2.0	1.0	1.0	2.0	16.0	12.0	11.0 12.0
198	...	4.0	2.0	4.0	2.0	3.0	2.0	24.0	18.0	18.0 18.0
205	...	4.0	4.0	3.0	3.0	4.0	5.0	28.0	10.0	9.0 9.0
213	...	4.0	4.0	4.0	2.0	4.0	5.0	15.0	6.0	7.0 8.0
217	...	3.0	2.0	4.0	2.0	4.0	4.0	13.0	6.0	6.0 8.0
221	...	4.0	3.0	4.0	1.0	1.0	5.0	0.0	6.0	5.0 0.0
225	...	5.0	3.0	3.0	1.0	1.0	4.0	16.0	9.0	8.0 7.0
239	...	5.0	5.0	4.0	3.0	5.0	2.0	0.0	7.0	7.0 0.0
248	...	4.0	3.0	3.0	1.0	3.0	5.0	8.0	3.0	5.0 5.0
250	...	4.0	4.0	5.0	2.0	4.0	5.0	0.0	6.0	8.0 8.0
252	...	3.0	2.0	5.0	2.0	5.0	5.0	4.0	6.0	9.0 8.0
255	...	4.0	4.0	4.0	1.0	2.0	5.0	2.0	7.0	9.0 8.0
278	...	2.0	4.0	4.0	1.0	1.0	4.0	15.0	9.0	8.0 8.0
281	...	4.0	4.0	4.0	3.0	4.0	3.0	19.0	11.0	9.0 10.0
292	...	5.0	4.0	3.0	1.0	1.0	5.0	12.0	12.0	12.0 13.0
304	...	4.0	4.0	4.0	1.0	1.0	3.0	20.0	15.0	14.0 13.0

305	...	4.0	4.0	3.0	1.0	1.0	3.0	8.0	14.0	12.0 12.0
307	...	4.0	3.0	4.0	1.0	1.0	4.0	38.0	8.0	9.0 8.0
308	...	4.0	5.0	3.0	1.0	2.0	5.0	0.0	15.0	12.0 12.0
309	...	4.0	4.0	3.0	1.0	3.0	3.0	18.0	12.0	10.0 10.0
310	...	4.0	2.0	4.0	2.0	2.0	3.0	0.0	9.0	9.0 0.0
312	...	4.0	5.0	2.0	2.0	2.0	4.0	3.0	13.0	11.0 11.0
313	...	4.0	2.0	2.0	1.0	2.0	1.0	22.0	13.0	10.0 11.0
315	...	4.0	1.0	2.0	1.0	1.0	3.0	40.0	13.0	11.0 11.0
336	...	5.0	4.0	3.0	1.0	2.0	5.0	12.0	14.0	13.0 13.0
340	...	4.0	3.0	4.0	1.0	3.0	3.0	4.0	11.0	12.0 11.0
341	...	4.0	3.0	3.0	2.0	2.0	2.0	0.0	10.0	10.0 0.0
343	...	3.0	3.0	1.0	1.0	2.0	4.0	0.0	9.0	8.0 0.0
349	...	2.0	5.0	5.0	5.0	5.0	5.0	10.0	11.0	13.0 13.0
352	...	4.0	3.0	3.0	2.0	3.0	3.0	7.0	8.0	7.0 8.0
353	...	4.0	4.0	4.0	3.0	3.0	5.0	4.0	8.0	8.0 8.0
361	...	4.0	4.0	3.0	2.0	3.0	5.0	2.0	13.0	12.0 12.0
367	...	5.0	2.0	1.0	1.0	2.0	1.0	0.0	7.0	6.0 0.0
383	...	4.0	3.0	2.0	1.0	3.0	5.0	0.0	6.0	5.0 0.0
384	...	5.0	4.0	3.0	4.0	3.0	3.0	14.0	6.0	5.0 5.0
387	...	5.0	4.0	2.0	1.0	2.0	5.0	0.0	7.0	5.0 0.0
389	...	1.0	1.0	1.0	1.0	1.0	5.0	0.0	6.0	5.0 0.0

[50 rows x 33 columns]

```
Fail_2 = data.where(data['failures']==2)
Fail_2 = Fail_2.dropna()
Fail_2
```

```
     school sex   age address famsize Pstatus  Medu  Fedu      Mjob
Fjob  \
25       GP   F  16.0       U     GT3       T   2.0   2.0  services
services
72       GP   F  15.0       R     GT3       T   1.0   1.0     other
other
85       GP   F  15.0       U     GT3       T   4.0   4.0  services
services
128      GP   M  18.0       R     GT3       T   2.0   2.0  services
other
130      GP   F  15.0       R     GT3       T   3.0   4.0  services
teacher
137      GP   F  16.0       U     GT3       A   3.0   3.0     other
other
141      GP   M  16.0       U     LE3       T   2.0   2.0  services
services
152      GP   F  15.0       R     GT3       T   3.0   3.0  services
services
160      GP   M  17.0       R     LE3       T   2.0   1.0   at_home
other
161      GP   M  15.0       R     GT3       T   3.0   2.0     other
other
170      GP   M  16.0       U     GT3       T   3.0   4.0     other
other
216      GP   F  17.0       U     GT3       T   4.0   3.0     other
other
270      GP   F  19.0       U     GT3       T   3.0   3.0     other
services
314      GP   F  19.0       U     GT3       T   1.0   1.0   at_home
health
370      MS   F  19.0       U     LE3       T   3.0   2.0  services
services
376      MS   F  20.0       U     GT3       T   4.0   2.0    health
other
390      MS   M  20.0       U     LE3       A   2.0   2.0  services
services

     ... famrel freetime  goout  Dalc  Walc health absences    G1
G2    G3
25   ...    1.0      2.0    2.0   1.0   3.0    5.0     14.0   6.0
9.0   8.0
72   ...    3.0      3.0    4.0   2.0   4.0    5.0      2.0   8.0
6.0   5.0
85   ...    4.0      4.0    4.0   2.0   3.0    5.0      6.0   7.0
9.0   8.0
128  ...    3.0      3.0    3.0   1.0   2.0    4.0      0.0   7.0
4.0   0.0
130  ...    4.0      2.0    2.0   2.0   2.0    5.0      0.0  12.0
0.0   0.0
137  ...    4.0      3.0    2.0   1.0   1.0    5.0      0.0   4.0
```

```
0.0    0.0
141   ...    2.0     3.0    3.0    2.0    2.0    2.0     8.0    9.0
9.0    9.0
152   ...    4.0     2.0    1.0    2.0    3.0    3.0     8.0   10.0
10.0   10.0
160   ...    3.0     3.0    2.0    2.0    2.0    5.0     0.0    7.0
6.0    0.0
161   ...    4.0     4.0    4.0    1.0    4.0    3.0     6.0    5.0
9.0    7.0
170   ...    3.0     4.0    5.0    2.0    4.0    2.0     0.0    6.0
5.0    0.0
216   ...    3.0     4.0    5.0    2.0    4.0    1.0    22.0    6.0
6.0    4.0
270   ...    4.0     3.0    5.0    3.0    3.0    5.0    15.0    9.0
9.0    9.0
314   ...    4.0     1.0    2.0    1.0    1.0    3.0    14.0   15.0
13.0   13.0
370   ...    3.0     2.0    2.0    1.0    1.0    3.0     4.0    7.0
7.0    9.0
376   ...    5.0     4.0    3.0    1.0    1.0    3.0     4.0   15.0
14.0   15.0
390   ...    5.0     5.0    4.0    4.0    5.0    4.0    11.0    9.0
9.0    9.0

[17 rows x 33 columns]

Fail_3 = data.where(data['failures']==3)
Fail_3 = Fail_3.dropna()
Fail_3

     school sex   age address famsize Pstatus  Medu  Fedu      Mjob
Fjob  \
2        GP   F  15.0       U     LE3       T   1.0   1.0   at_home
other
18       GP   M  17.0       U     GT3       T   3.0   2.0  services
services
78       GP   M  17.0       U     GT3       T   2.0   1.0     other
other
127      GP   F  19.0       U     GT3       T   0.0   1.0   at_home
other
144      GP   M  17.0       U     GT3       T   2.0   1.0     other
other
146      GP   F  15.0       U     GT3       T   3.0   2.0    health
services
149      GP   M  15.0       U     LE3       A   2.0   1.0  services
other
150      GP   M  18.0       U     LE3       T   1.0   1.0     other
other
153      GP   M  19.0       U     GT3       T   3.0   2.0  services
at_home
```

```
157      GP    F   18.0       R       GT3       T    1.0    1.0    at_home
other
164      GP    M   17.0       R       LE3       T    1.0    1.0     other
services
173      GP    F   16.0       U       GT3       T    1.0    3.0    at_home
services
206      GP    F   16.0       U       GT3       A    3.0    1.0   services
other
247      GP    M   22.0       U       GT3       T    3.0    1.0   services
services
350      MS    M   19.0       R       GT3       T    1.0    1.0     other
services
392      MS    M   21.0       R       GT3       T    1.0    1.0     other
other

      ...  famrel  freetime   goout   Dalc   Walc  health  absences     G1    G2
G3
2     ...     4.0       3.0     2.0    2.0    3.0     3.0      10.0    7.0   8.0
10.0
18    ...     5.0       5.0     5.0    2.0    4.0     5.0      16.0    6.0   5.0
5.0
78    ...     4.0       5.0     1.0    1.0    1.0     3.0       2.0    8.0   8.0
10.0
127   ...     3.0       4.0     2.0    1.0    1.0     5.0       2.0    7.0   8.0
9.0
144   ...     5.0       4.0     5.0    1.0    2.0     5.0       0.0    5.0   0.0
0.0
146   ...     3.0       3.0     2.0    1.0    1.0     3.0       0.0    6.0   7.0
0.0
149   ...     4.0       5.0     5.0    2.0    5.0     5.0       0.0    8.0   9.0
10.0
150   ...     2.0       3.0     5.0    2.0    5.0     4.0       0.0    6.0   5.0
0.0
153   ...     4.0       5.0     4.0    1.0    1.0     4.0       0.0    5.0   0.0
0.0
157   ...     5.0       2.0     5.0    1.0    5.0     4.0       6.0    9.0   8.0
10.0
164   ...     5.0       3.0     5.0    1.0    5.0     5.0       0.0    5.0   8.0
7.0
173   ...     4.0       3.0     5.0    1.0    1.0     3.0       0.0    8.0   7.0
0.0
206   ...     2.0       3.0     3.0    2.0    2.0     4.0       5.0    7.0   7.0
7.0
247   ...     5.0       4.0     5.0    5.0    5.0     1.0      16.0    6.0   8.0
8.0
350   ...     5.0       4.0     4.0    3.0    3.0     2.0       8.0    8.0   7.0
8.0
392   ...     5.0       5.0     3.0    3.0    3.0     3.0       3.0   10.0   8.0
7.0
```

```
[16 rows x 33 columns]
```

Failure Outliers

```
#With Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_0,ax=ax)
```

```
<Axes: >
```
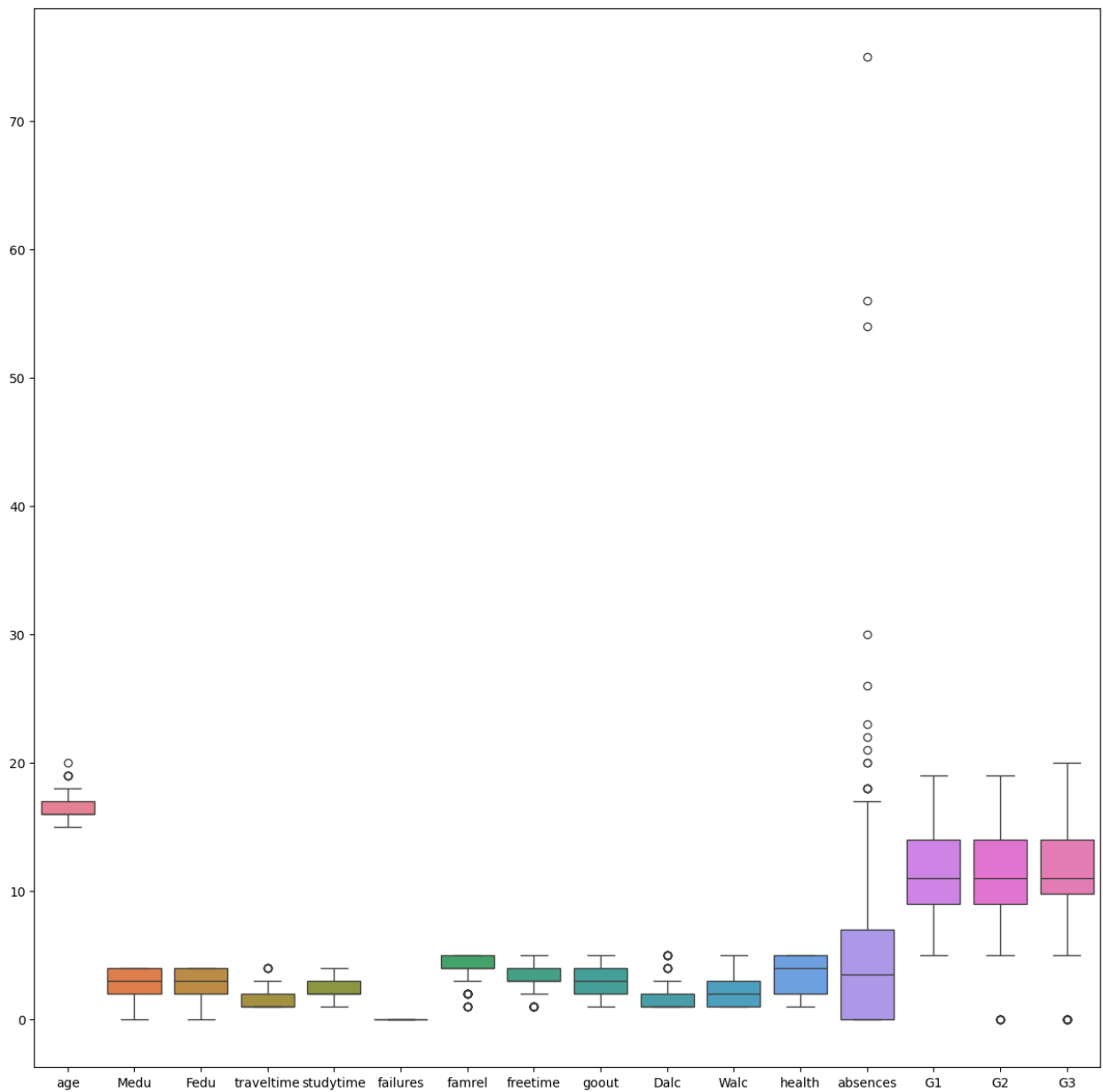

Handling Fail_0 Outliers

```python
Fail_0['age'].value_counts()

age
16.0    88
17.0    84
15.0    71
18.0    63
19.0     5
20.0     1
Name: count, dtype: int64

age_mean = round(Fail_0['age'].mean())
Fail_0['age']=Fail_0['age'].apply(lambda x : age_mean if x>18 else x)

Fail_0['traveltime'].value_counts()

traveltime
1.0    209
2.0     80
3.0     17
4.0      6
Name: count, dtype: int64

Travel_mean = round(Fail_0['traveltime'].mean())
# round(Medu_mean)
Fail_0['traveltime'] = Fail_0['traveltime'].replace(4,Travel_mean)

Fail_0['famrel'].value_counts()

famrel
4.0    151
5.0     88
3.0     54
2.0     13
1.0      6
Name: count, dtype: int64

Famrel_Mean = round(Fail_0['famrel'].mean())
Fail_0['famrel']= Fail_0['famrel'].apply(lambda x:x if x>=3 else
Famrel_Mean)

Fail_0['freetime'].value_counts()

freetime
3.0    134
4.0     86
2.0     52
5.0     27
1.0     13
Name: count, dtype: int64
```

```python
Freetime_Mean = round(Fail_0['freetime'].mean())
Fail_0['freetime']= Fail_0['freetime'].replace(1,Freetime_Mean)

Fail_0['Dalc'].value_counts()
```

```
Dalc
1.0     232
2.0      50
3.0      17
5.0       7
4.0       6
Name: count, dtype: int64
```

```python
Fail_0['Dalc'] = Fail_0['Dalc'].apply(lambda x : x if x<4 else 3)

Absence = Fail_0['absences'].value_counts()
Absence = pd.DataFrame(Absence)
Absence = Absence.reset_index()
Absence
```

```
    absences  count
0        0.0     89
1        2.0     58
2        4.0     47
3        6.0     25
4        8.0     17
5       10.0     15
6       12.0      9
7       14.0      8
8        3.0      6
9        7.0      6
10      18.0      4
11       5.0      4
12      16.0      3
13       9.0      3
14       1.0      3
15      11.0      2
16      13.0      2
17      20.0      2
18      30.0      1
19      23.0      1
20      56.0      1
21      22.0      1
22      75.0      1
23      21.0      1
24      26.0      1
25      54.0      1
26      17.0      1
```

```python
Fail_0['absences']=Absence['count'].apply(lambda x:x if x<=2 else
round(Fail_0['absences'].mean()))
```

```
Fail_0['G2'].value_counts()

G2
10.0    37
9.0     37
15.0    34
12.0    34
13.0    32
11.0    31
14.0    21
8.0     20
16.0    13
18.0    11
7.0     11
6.0      9
0.0      8
5.0      6
17.0     5
19.0     3
Name: count, dtype: int64

G2_Mean = round(Fail_0['G2'].mean())
Fail_0['G2'] = Fail_0['G2'].apply(lambda x:x if x>=8 else G2_Mean)

Fail_0['G3'].value_counts()

G3
10.0    45
11.0    42
15.0    32
14.0    26
13.0    25
12.0    25
9.0     21
0.0     18
8.0     18
16.0    16
6.0     15
18.0    11
17.0     6
19.0     5
5.0      3
7.0      3
20.0     1
Name: count, dtype: int64

G3_Mean = round(Fail_0['G3'].mean())
G3_Mean
Fail_0['G3'] = Fail_0['G3'].apply(lambda x:x if x>6 else G2_Mean)
```

```
#After Removing Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_0,ax=ax)

<Axes: >
```
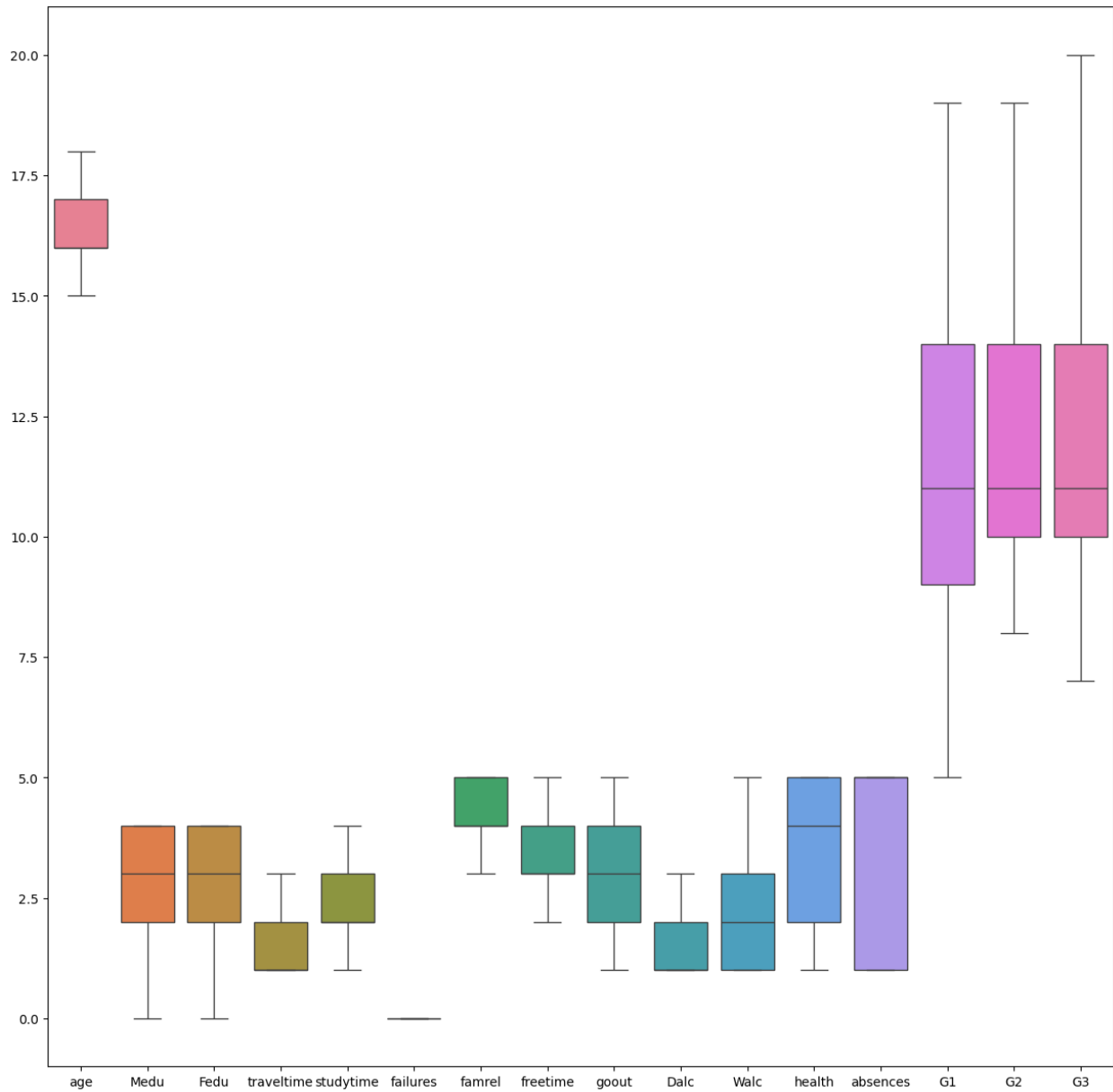


```
# fig,ax = plt.subplots(figsize = (15,15))
sns.displot(data = Fail_0,kind='kde')

C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\2969027086.py:2:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(data = Fail_0,kind='kde')
```

```
<seaborn.axisgrid.FacetGrid at 0x20294760860>
```
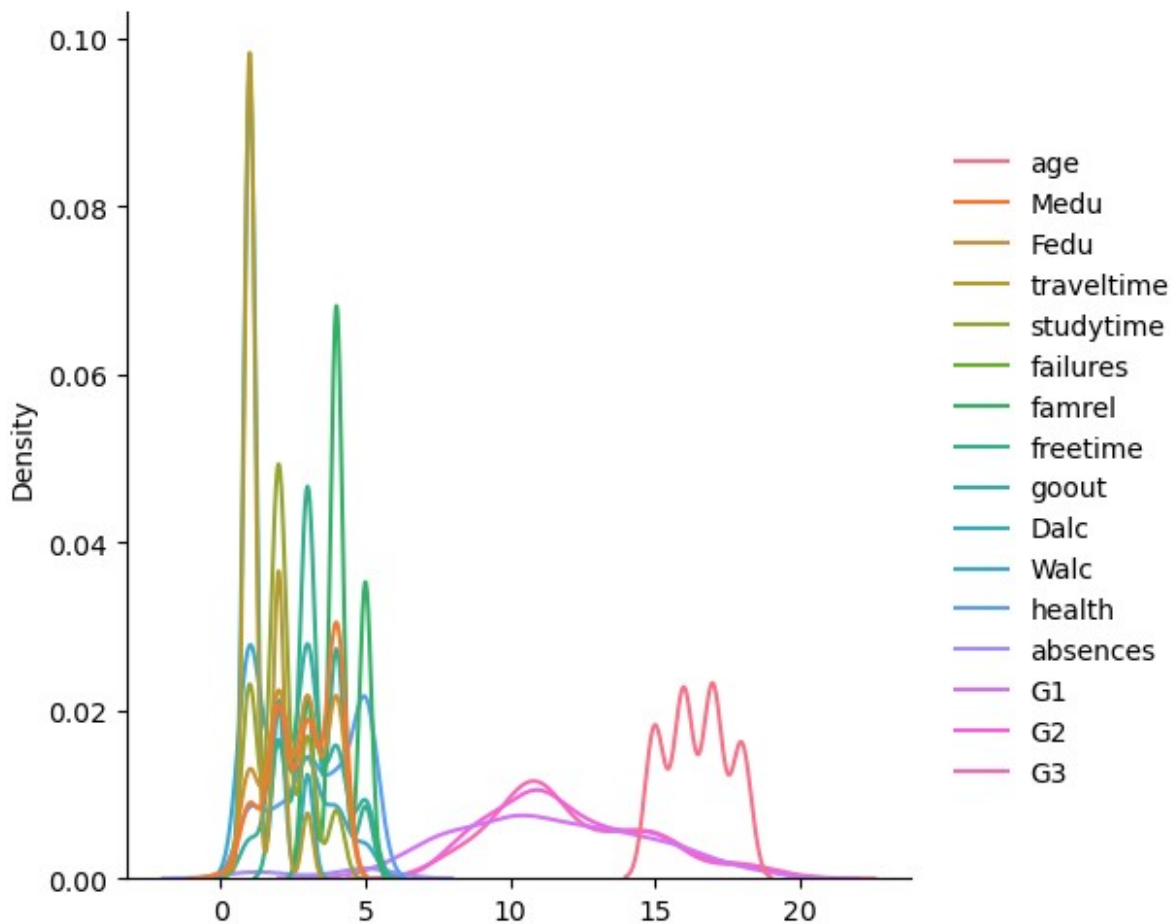

Fail_0 Linear Regression

```
fail_0_X =  Fail_0.iloc[:,:-1].select_dtypes(include="float64")
fail_0_Y = Fail_0.iloc[:,:-1].select_dtypes(include="float64")

from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test=train_test_split(fail_0_X,fail_0_Y,test_
size=0.3,random_state=45)

X_train.dtypes
```

```
age            float64
Medu           float64
Fedu           float64
traveltime     float64
studytime      float64
failures       float64
famrel         float64
freetime       float64
```

```
goout          float64
Dalc           float64
Walc           float64
health         float64
absences       float64
G1             float64
G2             float64
dtype: object
```

```python
from sklearn.preprocessing import StandardScaler
scaler =  StandardScaler()
X_train_norm =
scaler.fit_transform(X_train.select_dtypes(include="float64").dropna()
)

X_train_norm
```

```
array([[ 0.71818485, -1.8973666 , -1.15470054,  1.17932379,
0.29559878,
         0.         ,  1.41421356,  0.83205029,  1.336198  ,
3.87298335,
         2.29227953,  1.0059702 , -1.04560405,  0.38646346,
0.10259784],
       [-0.92338052,  0.63245553,  0.         ,  1.17932379,
0.29559878,
         0.         ,  1.41421356,  0.83205029,  0.31814238, -
0.25819889,
         0.13483997, -0.23214697,  1.11306237, -0.70472748, -
1.12857619],
       [ 0.71818485, -0.63245553,  0.         ,  2.89470384,
0.29559878,
         0.         ,  1.41421356, -0.2773501 , -0.69991324, -
0.25819889,
        -0.94387981,  0.38691162, -0.50593744, -1.43218811, -
1.12857619],
       [-0.92338052,  0.63245553,  1.15470054, -0.53605627, -
1.28092806,
         0.         ,  1.41421356,  0.83205029, -0.69991324, -
0.25819889,
        -0.94387981,  1.0059702 , -1.04560405,  0.02273314,
0.92338052],
       [ 0.71818485,  0.63245553,  1.15470054, -0.53605627, -
1.28092806,
         0.         ,  0.         ,  0.83205029,  1.336198  , -
0.25819889,
         0.13483997, -0.85120556, -0.50593744,  0.75019377,
0.51298918],
       [-0.92338052,  0.63245553, -1.15470054, -0.53605627,
1.87212563,
         0.         , -1.41421356, -1.38675049, -0.69991324, -
```

```
0.25819889,
        -0.94387981,  1.0059702 ,  1.11306237,  1.11392409,
0.51298918],
       [-0.92338052, -1.8973666 , -1.15470054, -0.53605627, -
1.28092806,
         0.        ,  0.        , -1.38675049, -0.69991324, -
0.25819889,
         0.13483997,  1.0059702 , -1.04560405,  0.02273314, -
0.30779351],
       [-0.92338052, -0.63245553, -1.15470054, -0.53605627,
0.29559878,
         0.        ,  0.        , -1.38675049, -0.69991324, -
0.25819889,
        -0.94387981, -1.47026414,  1.11306237,  1.4776544 ,
2.15455454],
       [-0.92338052,  0.63245553,  1.15470054, -0.53605627,
0.29559878,
         0.        , -1.41421356, -0.2773501 ,  0.31814238, -
0.25819889,
         0.13483997, -0.85120556,  1.11306237, -0.70472748, -
1.94935887],
       [ 0.71818485,  0.63245553, -1.15470054, -0.53605627,
0.29559878,
         0.        ,  0.        ,  1.94145069, -1.71796886, -
0.25819889,
         1.21355975,  1.0059702 , -1.04560405,  1.11392409,
0.92338052],
       [-0.92338052,  0.63245553,  0.        , -0.53605627,
0.29559878,
         0.        ,  0.        ,  0.83205029, -1.71796886, -
0.25819889,
        -0.94387981, -1.47026414, -1.04560405,  0.38646346,
0.51298918],
       [ 0.71818485, -1.8973666 , -1.15470054, -0.53605627,
0.29559878,
         0.        ,  0.        ,  0.83205029,  1.336198  , -
0.25819889,
        -0.94387981, -0.23214697,  1.11306237,  0.02273314, -
0.30779351],
       [-0.92338052,  0.63245553,  1.15470054, -0.53605627, -
1.28092806,
         0.        ,  0.        , -0.2773501 ,  0.31814238, -
0.25819889,
         1.21355975,  1.0059702 ,  1.11306237,  0.75019377,
0.51298918],
       [ 2.35975021,  0.63245553,  1.15470054,  1.17932379,
0.29559878,
         0.        ,  0.        , -0.2773501 ,  1.336198  , -
0.25819889,
```

```
        -0.94387981, -1.47026414,  1.11306237, -2.15964874, -
0.71818485],
       [ 0.71818485,  0.63245553,  0.        , -0.53605627, -
1.28092806,
         0.        , -1.41421356, -0.2773501 ,  0.31814238, -
0.25819889,
         1.21355975,  1.0059702 , -1.04560405, -1.43218811, -
1.12857619],
       [ 0.71818485,  0.63245553,  1.15470054, -0.53605627,
1.87212563,
         0.        , -1.41421356, -1.38675049,  0.31814238, -
0.25819889,
         0.13483997, -0.85120556, -0.50593744,  0.38646346,
0.51298918]])

fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=X_train_norm,ax=ax)

<Axes: >
```
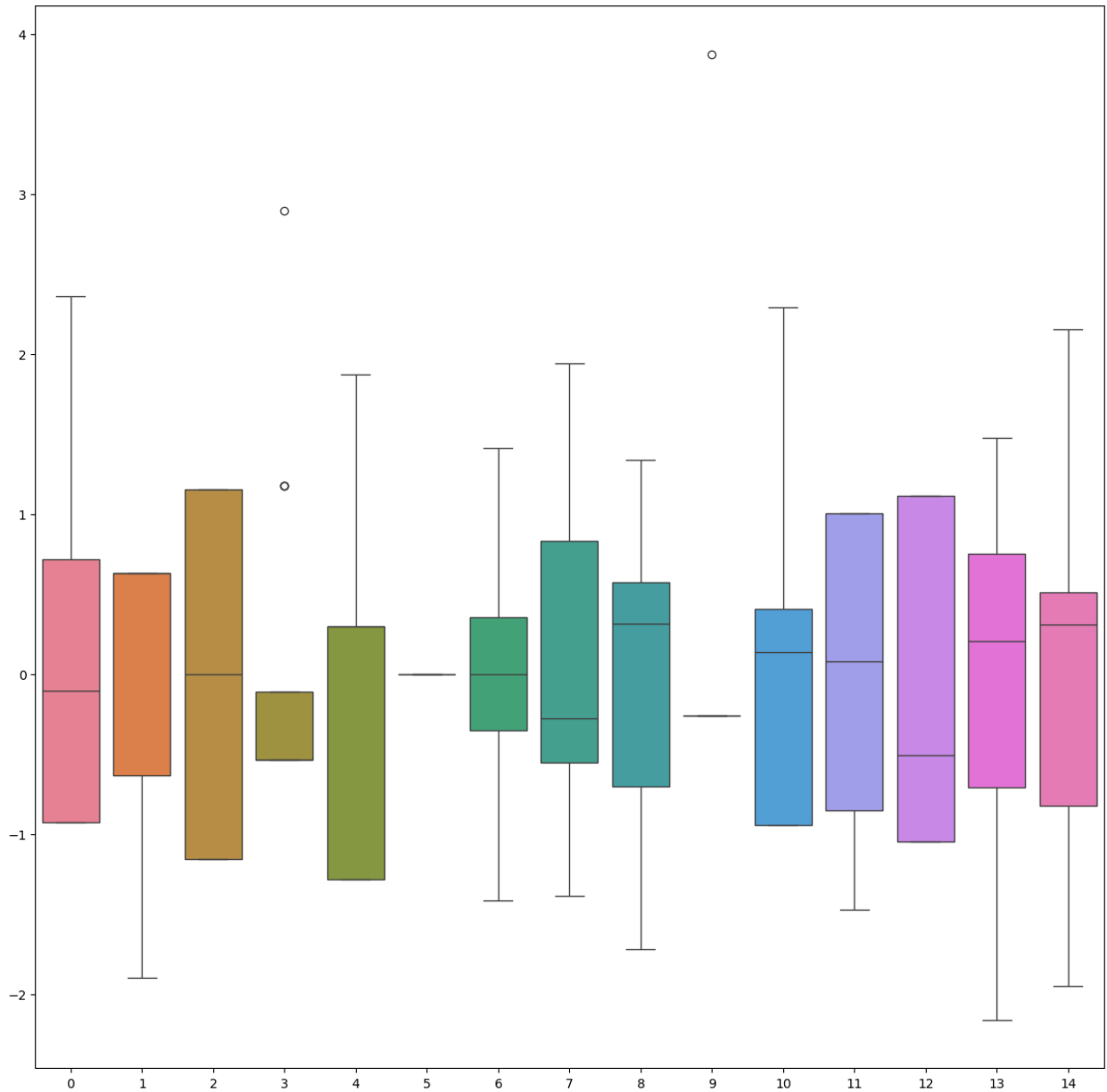
```
X_test_norm =
scaler.transform(X_test.select_dtypes(include="float64").dropna())

X_test_norm

array([[ 4.00131557,  0.63245553,  1.15470054,  1.17932379,
0.29559878,
         0.        ,  0.        , -0.2773501 ,  1.336198  , -
0.25819889,
        -0.94387981, -0.23214697,  1.11306237, -2.52337906, -
0.71818485],
       [ 0.71818485,  0.63245553,  0.        , -0.53605627,
0.29559878,
```

```
        0.        ,  1.41421356,  0.83205029, -0.69991324, -
0.25819889,
        0.13483997,  1.0059702 ,  1.11306237,  1.11392409,
0.92338052],
      [-0.92338052, -1.8973666 , -1.15470054, -0.53605627,
1.87212563,
        0.        ,  0.        ,  1.94145069, -0.69991324, -
0.25819889,
       -0.94387981, -0.23214697,  1.11306237,  0.75019377,
1.33377186],
      [ 0.71818485, -0.63245553,  0.        , -0.53605627,
0.29559878,
        0.        ,  0.        , -0.2773501 , -0.69991324, -
0.25819889,
        0.13483997,  1.0059702 ,  1.11306237, -2.15964874, -
1.12857619],
      [ 2.35975021, -3.16227766, -2.30940108, -0.53605627,
0.29559878,
        0.        ,  1.41421356, -0.2773501 ,  0.31814238, -
0.25819889,
       -0.94387981, -0.23214697,  1.11306237, -2.52337906, -
0.71818485],
      [-0.92338052, -0.63245553,  1.15470054, -0.53605627,
0.29559878,
        0.        ,  1.41421356,  1.94145069, -1.71796886, -
0.25819889,
       -0.94387981,  1.0059702 ,  1.11306237,  0.75019377,
0.92338052],
      [-0.92338052, -1.8973666 ,  1.15470054, -0.53605627,
1.87212563,
        0.        ,  0.        , -0.2773501 , -0.69991324, -
0.25819889,
       -0.94387981,  1.0059702 , -1.04560405, -0.70472748, -
1.53896753],
      [-0.92338052, -1.8973666 , -2.30940108,  2.89470384,
1.87212563,
        0.        ,  1.41421356, -1.38675049, -0.69991324, -
0.25819889,
       -0.94387981,  0.38691162,  1.11306237, -0.70472748, -
0.30779351]])
```

```python
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=X_test_norm,ax=ax)
```

```
<Axes: >
```
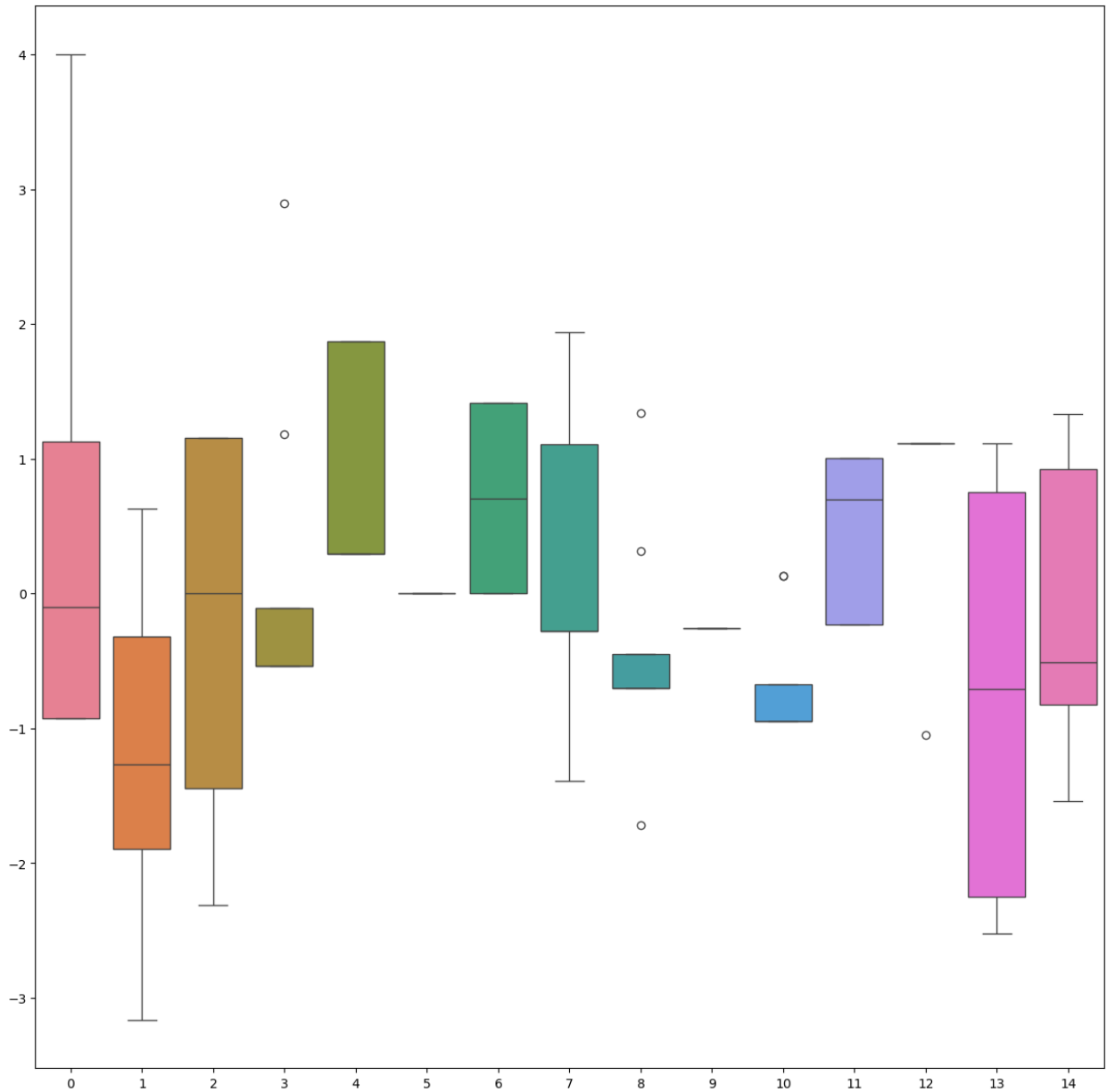
```
from sklearn.linear_model import LinearRegression
regression = LinearRegression()
regression.fit(X_train_norm,Y_train.dropna())
```

```
LinearRegression()
```

```
print(regression.coef_)
```

```
[[ 6.09174647e-01 -5.07205469e-18 -4.77846044e-17  2.94638273e-16
   -7.18875381e-17 -2.47383308e-16  1.00105438e-16  1.88341462e-16
   -1.19465458e-17 -5.47060832e-17  2.76527754e-17  3.46767336e-16
    6.98792983e-17 -1.01038781e-16  8.66674560e-17]
 [ 7.70442602e-16  7.90569415e-01 -4.13827465e-16 -4.77888205e-16
```

```
   -5.64735913e-16 -2.22044605e-16  7.06150233e-16 -6.91351067e-16
   -1.76258950e-16  3.97760489e-16 -6.36402380e-16  1.86837556e-16
   -1.84946370e-17  6.92205041e-16 -9.95206477e-16]
 [ 2.02974765e-16 -2.80795484e-16  8.66025404e-01 -6.51694482e-16
    3.62803494e-16  1.11022302e-16  6.90431464e-16 -1.58657993e-16
   -3.00476391e-16 -2.41666345e-16  2.62889634e-16 -1.63807638e-16
    4.21757109e-17 -4.30627833e-17 -4.53135085e-16]
 [ 7.04187905e-16  2.37861003e-16 -1.19158685e-16  5.82961191e-01
   -4.79122459e-16  2.49800181e-16  4.78254141e-16 -5.98107879e-16
   -4.70247479e-17  1.76052047e-16 -7.25351336e-18  5.57385846e-18
   -4.88443053e-17  5.62564601e-16 -8.43180087e-16]
 [-2.97572189e-16  1.10819733e-17 -1.88671217e-16 -8.24517915e-17
    6.34305723e-01  4.85722573e-17 -5.20353869e-17 -1.97750644e-16
    3.58708725e-16  9.96687712e-17 -3.23299378e-16  2.92161720e-16
   -2.52837548e-16  7.69235685e-17 -1.09477873e-16]
 [ 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
    0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
    0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
    0.00000000e+00  0.00000000e+00  0.00000000e+00]
 [ 7.74084399e-16  2.73829164e-16 -5.61337225e-17 -3.64182905e-16
   -5.49637804e-17 -4.64905892e-16  7.07106781e-01 -4.60204623e-16
    5.49940692e-17 -1.81785731e-16 -8.61387635e-17 -8.05329018e-17
    1.68594802e-16  6.33111720e-16 -7.30176531e-16]
 [ 4.83676364e-16  3.08455802e-16 -4.24893952e-16  4.03202605e-17
   -3.60285892e-16  3.88578059e-16  5.13395513e-16  9.01387819e-01
    2.25850169e-16  1.22063747e-16 -6.45607560e-16 -1.61039997e-16
    1.79819713e-17  6.44107323e-16 -8.31847206e-16]
 [-1.22139106e-15 -1.90484471e-16 -2.17675877e-16  4.90522142e-16
    2.54794694e-16  2.22044605e-16  1.87123731e-16  7.44521351e-16
    9.82264603e-01 -2.19939054e-16  5.18941174e-16  1.76759342e-16
   -2.34965082e-16 -7.86029468e-16  8.29926307e-16]
 [ 3.04768157e-17  5.32762715e-17 -7.80002939e-17 -9.50915759e-17
    2.97831111e-18 -2.08166817e-17  9.02497218e-17 -6.62999986e-17
    9.73564684e-17  2.42061459e-01 -5.68101379e-17 -4.12593606e-17
   -1.21617358e-16 -7.53018091e-17 -9.40546865e-17]
 [-2.98025944e-16 -8.43107298e-16  5.44746267e-16  7.58827217e-16
    3.05139501e-17 -1.66533454e-16 -7.72302328e-16  2.79881865e-16
   -3.49951982e-16 -3.34736334e-16  9.27024811e-01 -1.21451821e-16
    5.71345985e-17 -8.25428974e-16  9.73924310e-16]
 [ 1.52889922e-15 -1.74237552e-16  7.32637742e-16 -6.54481377e-16
   -2.05921233e-16  3.33066907e-16  9.51694443e-18 -4.06954854e-16
   -1.18532167e-15  5.31945102e-16  4.01691244e-17  1.61535600e+00
    1.08337968e-15  1.87108234e-16 -6.45373431e-16]
 [ 1.05061918e-15 -1.06990642e-16 -2.81733099e-17 -2.08596637e-15
   -4.54287813e-16  1.11022302e-16  7.19165124e-16 -4.96795878e-17
   -6.08263616e-17  6.54468955e-16  1.75056873e-16  5.62875502e-16
    1.85299589e+00  6.33967639e-17  4.13637190e-17]
 [ 1.44929588e-15  4.63118827e-16 -2.80825858e-16  8.15482154e-16
   -9.21034383e-16 -8.88178420e-16 -1.12021806e-16 -7.93996658e-17
```

```
      7.24726983e-16 -1.04389433e-15 -7.93443125e-16 -2.08535830e-16
     -4.56853352e-16  2.74928968e+00 -5.61343962e-16]
    [ 3.47644661e-15  1.18299035e-15  7.93487173e-16 -1.86151318e-15
     -1.54687080e-15  1.33226763e-15  2.01719902e-15 -2.11009577e-15
     -2.59333698e-15  6.00515749e-17 -7.90541945e-16  5.23872871e-16
      1.40463877e-15  2.08755891e-15  2.43669859e+00]]
```

```python
reg_predict = regression.predict(X_test_norm)
reg_predict
```

```
array([[18.,  4.,  4.,  2.,  2.,  0.,  4.,  3.,  4.,  1.,  1.,  3.,
5.,
         5., 11.],
       [16.,  4.,  3.,  1.,  2.,  0.,  5.,  4.,  2.,  1.,  2.,  5.,
5.,
        15., 15.],
       [15.,  2.,  2.,  1.,  3.,  0.,  4.,  5.,  2.,  1.,  1.,  3.,
5.,
        14., 16.],
       [16.,  3.,  3.,  1.,  2.,  0.,  4.,  3.,  2.,  1.,  2.,  5.,
5.,
         6., 10.],
       [17.,  1.,  1.,  1.,  2.,  0.,  5.,  3.,  3.,  1.,  1.,  3.,
5.,
         5., 11.],
       [15.,  3.,  4.,  1.,  2.,  0.,  5.,  5.,  1.,  1.,  1.,  5.,
5.,
        14., 15.],
       [15.,  2.,  4.,  1.,  3.,  0.,  4.,  3.,  2.,  1.,  1.,  5.,
1.,
        10.,  9.],
       [15.,  2.,  1.,  3.,  3.,  0.,  5.,  2.,  2.,  1.,  1.,  4.,
5.,
        10., 12.]])
```

```python
residual = Y_test.select_dtypes(include='float64').dropna()-
reg_predict
residual
```

```
            age          Medu          Fedu     traveltime
studytime  \
0    0.000000e+00 -1.776357e-15  4.440892e-16 -1.332268e-15  1.332268e-
15
5    0.000000e+00 -8.881784e-16 -8.881784e-16 -6.661338e-16  8.881784e-
16
14   0.000000e+00  2.886580e-15 -8.881784e-16  3.552714e-15  1.776357e-
15
4    0.000000e+00 -4.440892e-16 -1.776357e-15 -2.220446e-16  6.661338e-
16
1    0.000000e+00 -3.108624e-15 -3.108624e-15 -1.110223e-15  4.440892e-
```

```
16
9  -1.776357e-15   8.881784e-16   0.000000e+00   1.776357e-15   8.881784e-
16
24  0.000000e+00   2.220446e-16  -1.776357e-15   7.771561e-16   8.881784e-
16
11 -1.776357e-15   2.220446e-16  -8.881784e-16   1.776357e-15   8.881784e-
16
```

```
      failures          famrel         freetime           goout
Dalc  \
0          0.0 -2.664535e-15 -1.776357e-15   3.552714e-15 -2.220446e-16

5          0.0 -1.776357e-15   0.000000e+00   4.440892e-16   2.220446e-16

14         0.0   2.220446e-15   2.664535e-15 -3.108624e-15   4.440892e-16

4          0.0   0.000000e+00   8.881784e-16   4.440892e-16   0.000000e+00

1          0.0 -1.776357e-15 -1.776357e-15   8.881784e-16 -4.440892e-16

9          0.0   0.000000e+00   1.776357e-15 -1.110223e-15   4.440892e-16

24         0.0   4.440892e-16   8.881784e-16   0.000000e+00 -2.220446e-16

11         0.0   0.000000e+00 -1.332268e-15 -2.664535e-15   0.000000e+00
```

```
          Walc          health        absences              G1
G2
0  -2.220446e-16 -5.329071e-15   0.000000e+00 -2.664535e-15 -8.881784e-
15
5    1.554312e-15 -2.664535e-15 -1.776357e-15 -1.776357e-15 -7.105427e-
15
14 -1.776357e-15   1.776357e-15   1.776357e-15   3.552714e-15   1.065814e-
14
4  -8.881784e-16 -3.552714e-15 -1.776357e-15   4.440892e-15 -7.105427e-
15
1    0.000000e+00 -3.552714e-15 -3.552714e-15   3.552714e-15 -5.329071e-
15
9  -8.881784e-16 -1.776357e-15   0.000000e+00   3.552714e-15 -1.776357e-
15
24 -4.440892e-16   0.000000e+00 -2.220446e-16   5.329071e-15
0.000000e+00
11 -1.332268e-15   2.664535e-15   7.105427e-15   1.776357e-15   5.329071e-
15
```

```python
sns.displot(residual,kind = 'kde')
```

```
C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\4192276842.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
```

```
`warn_singular=False` to disable this warning.
  sns.displot(residual,kind = 'kde')

<seaborn.axisgrid.FacetGrid at 0x2029507b050>
```
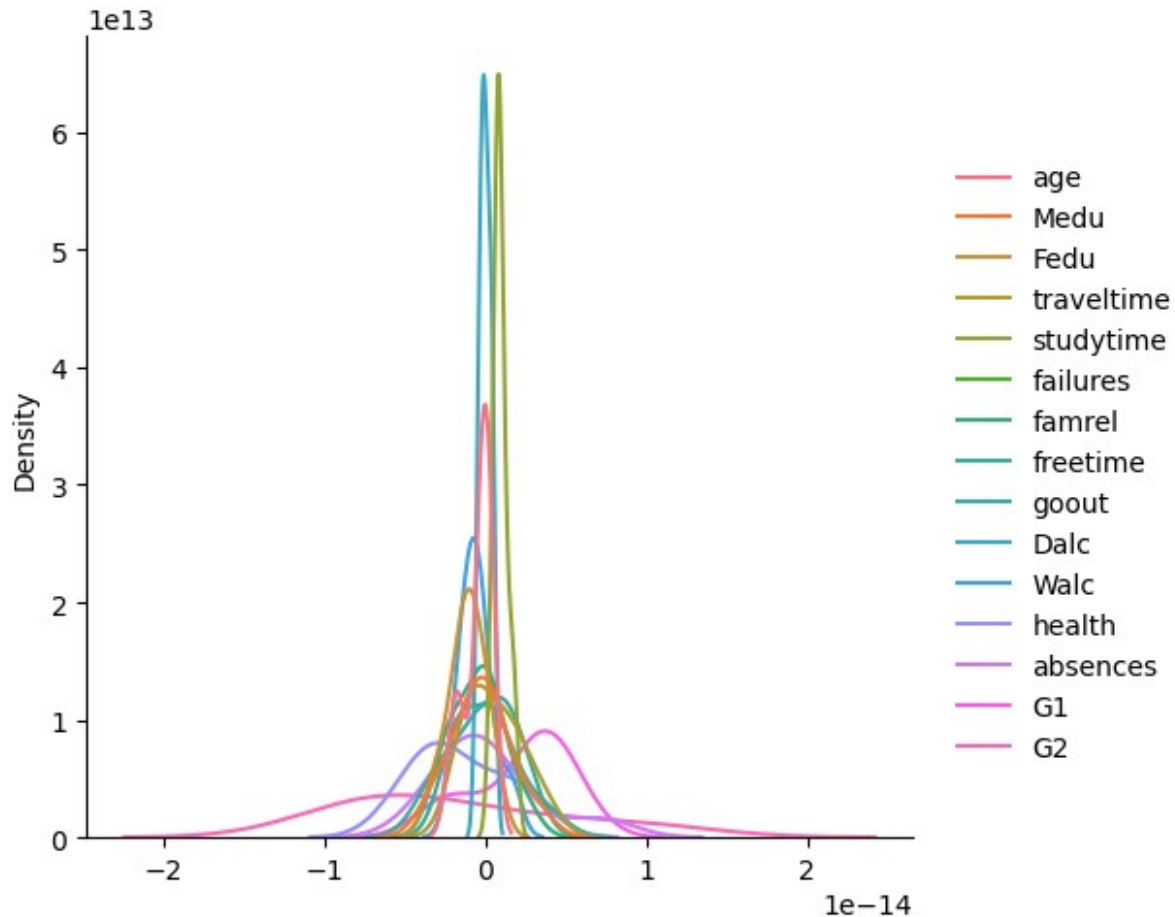


```
from sklearn.metrics import
mean_squared_error,mean_absolute_error,r2_score
print(mean_squared_error(Y_test.dropna(),reg_predict))
print(mean_absolute_error(Y_test.dropna(),reg_predict))
print(r2_score(Y_test.dropna(),reg_predict))

6.328246290333504e-30
1.6200004300988743e-15
0.9333333333333333
```

Handling Fail_1 Outliers

```
# Before Removing Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_1,ax=ax)
```

```
<Axes: >
```
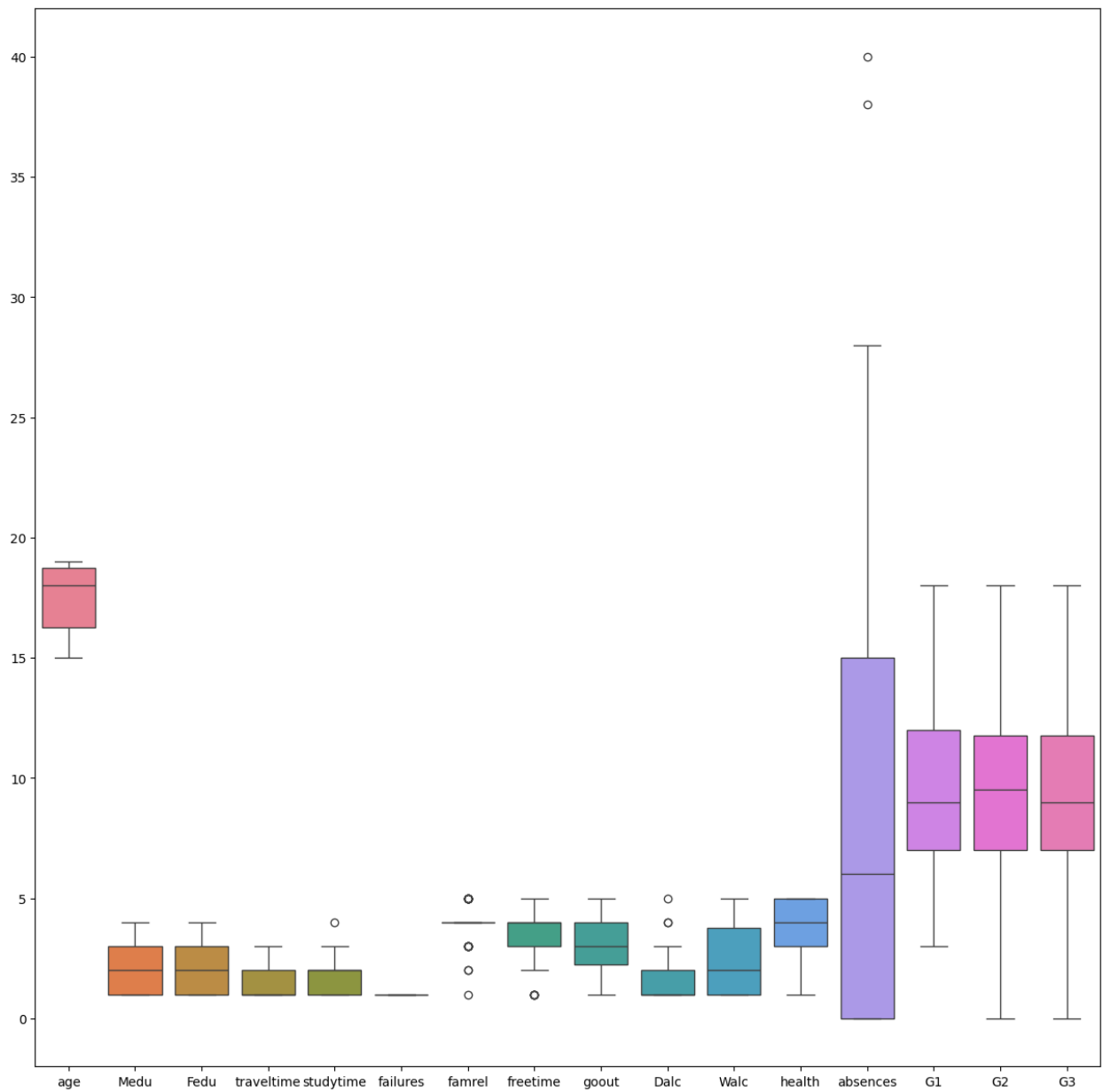


```
Fail_1['studytime'].value_counts()

studytime
2.0    26
1.0    16
3.0     7
4.0     1
Name: count, dtype: int64
```

```python
Studytime_mean = round(Fail_1['studytime'].mean())
Fail_1['studytime']=Fail_1['studytime'].replace(4,Studytime_mean)

Fail_1['famrel'].value_counts()

famrel
4.0    32
5.0     9
3.0     6
2.0     2
1.0     1
Name: count, dtype: int64

famrel_mean = round(Fail_1['famrel'].mean())
Fail_1['famrel']=Fail_1['famrel'].apply(lambda x:x if x==famrel_mean
else famrel_mean)

Fail_1['freetime'].value_counts()

freetime
4.0    20
3.0    11
5.0     7
2.0     7
1.0     5
Name: count, dtype: int64

Freetime_mean = round(Fail_1['freetime'].mean())
Fail_1['freetime']=Fail_1['freetime'].replace(1,Freetime_mean)

Fail_1['Dalc'].value_counts()

Dalc
1.0    29
2.0    12
3.0     6
4.0     2
5.0     1
Name: count, dtype: int64

Dalc_Mean = round(Fail_1['Dalc'].mean())
Fail_1['Dalc'] = Fail_1['Dalc'].apply(lambda x:x if x<4 else
Dalc_Mean)

Fail_1['absences'].value_counts()

absences
0.0    14
2.0     4
4.0     4
6.0     3
12.0    3
```

```
15.0     2
20.0     2
16.0     2
14.0     2
8.0      2
18.0     1
10.0     1
40.0     1
22.0     1
3.0      1
25.0     1
38.0     1
19.0     1
13.0     1
28.0     1
24.0     1
7.0      1
Name: count, dtype: int64

Absences_Mean = round(Fail_1['absences'].mean())
Fail_1['absences'] = Fail_1['absences'].apply(lambda x:x if x<30 else
Absences_Mean)

#After Removing Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_1,ax=ax)

<Axes: >
```
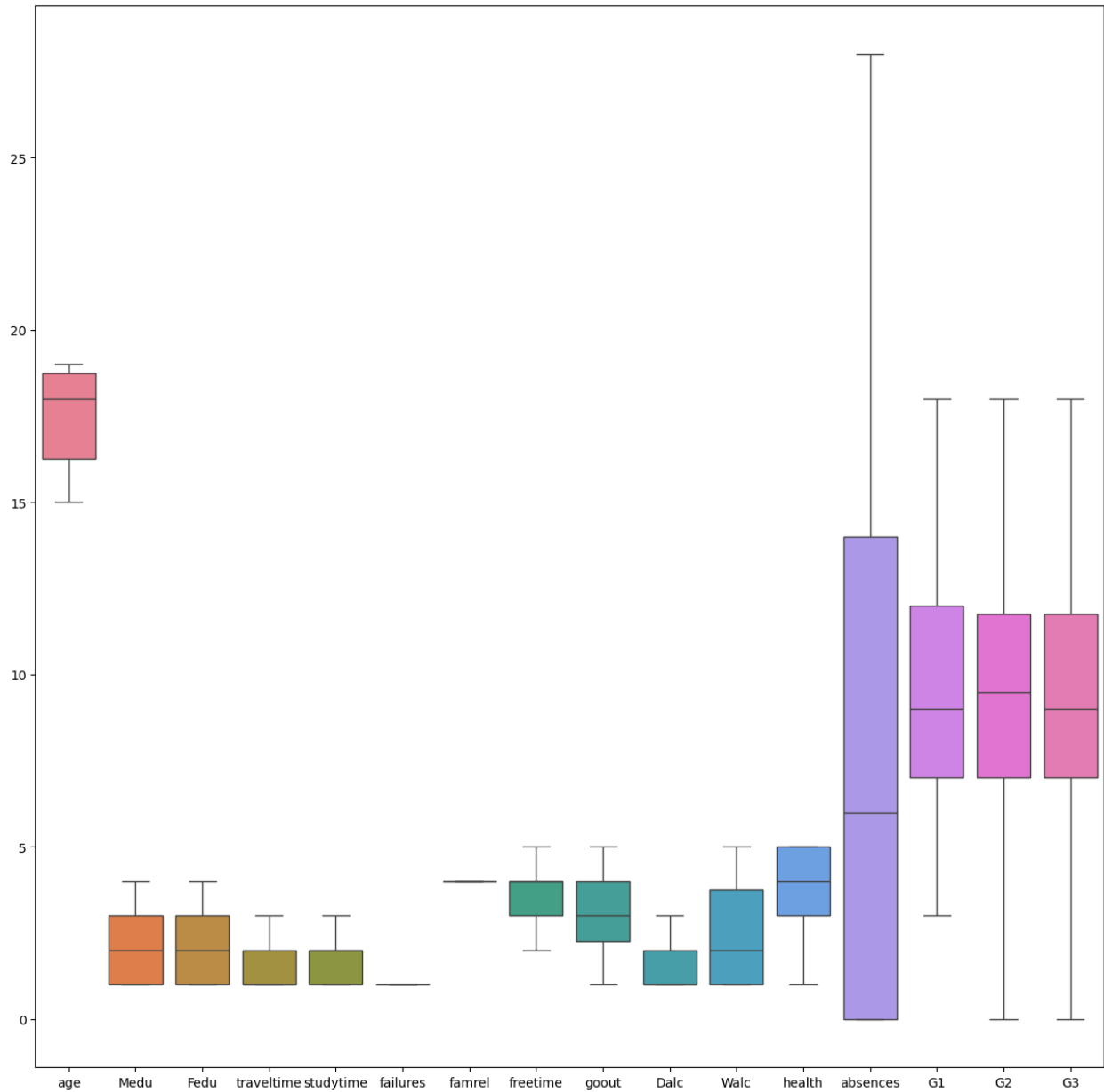
```
sns.displot(Fail_1,kind='kde')
```

C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\3221834503.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(Fail_1,kind='kde')

<seaborn.axisgrid.FacetGrid at 0x2029ac582f0>
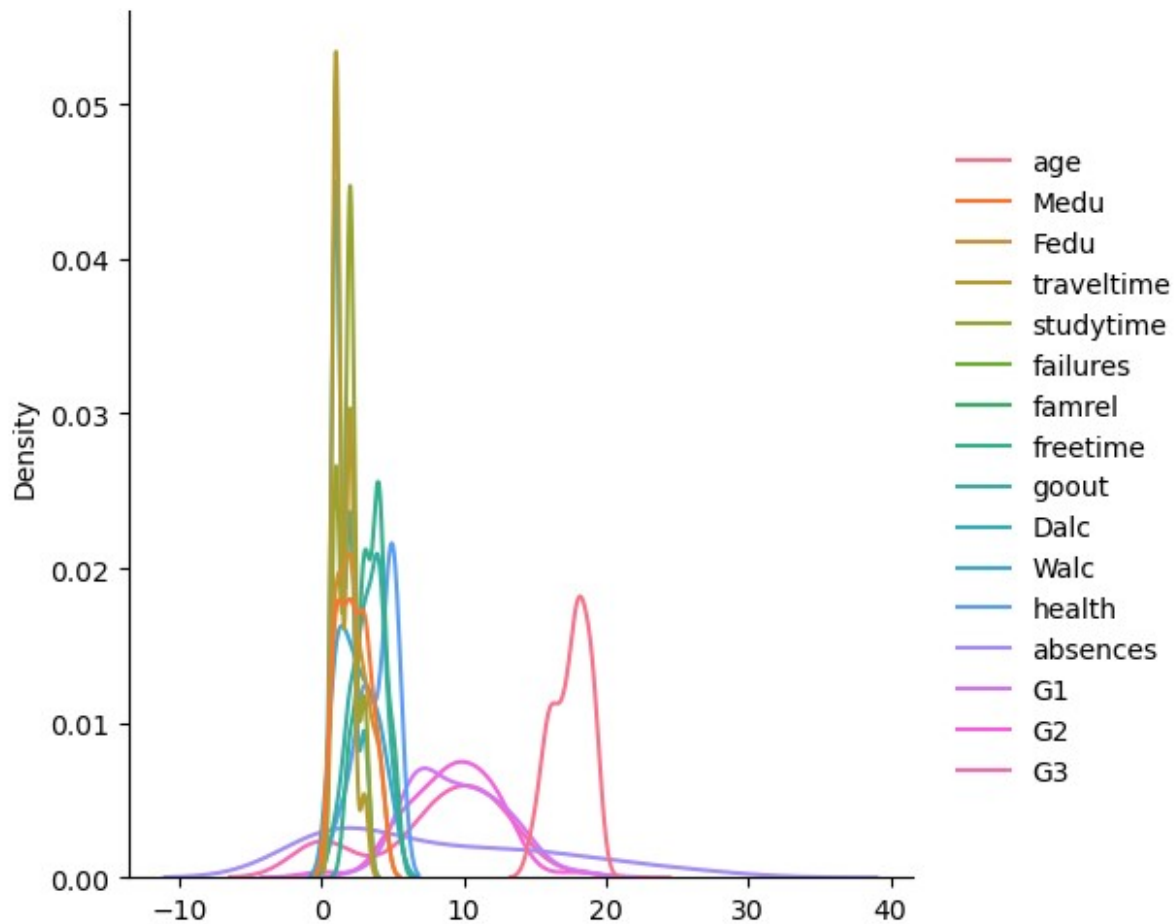
Fail_1 Linear Regression

```
X = Fail_1.iloc[::-1]
Y = Fail_1.iloc[::-1]

from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test =
train_test_split(X,Y,test_size=0.4,random_state=45)

X_train
```

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob
Fjob \									
88	GP	M	16.0	U	GT3	T	2.0	2.0	services
other									
118	GP	M	17.0	R	GT3	T	1.0	3.0	other
other									
250	GP	M	18.0	U	GT3	T	3.0	2.0	services
other									
292	GP	F	18.0	U	LE3	T	2.0	1.0	services
at_home									
239	GP	M	18.0	U	GT3	T	2.0	2.0	other

services

341	GP	M	18.0	U	GT3	T	4.0	4.0	teacher	services
248	GP	M	18.0	R	LE3	T	3.0	3.0	other	services
213	GP	M	18.0	U	GT3	T	2.0	2.0	services	other
151	GP	M	16.0	U	LE3	T	2.0	1.0	at_home	other
112	GP	F	16.0	U	GT3	T	2.0	2.0	at_home	other
353	MS	M	19.0	R	GT3	T	1.0	1.0	other	other
309	GP	F	19.0	U	LE3	T	1.0	1.0	at_home	other
111	GP	F	16.0	R	GT3	T	3.0	3.0	services	other
278	GP	F	18.0	U	GT3	T	4.0	4.0	health	health
205	GP	F	17.0	U	GT3	T	3.0	4.0	at_home	services
336	GP	F	19.0	R	GT3	A	3.0	1.0	services	at_home
159	GP	M	16.0	U	GT3	T	3.0	3.0	other	services
281	GP	M	17.0	U	LE3	A	3.0	2.0	teacher	services
349	MS	M	18.0	R	GT3	T	3.0	2.0	other	other
313	GP	F	19.0	U	LE3	T	3.0	2.0	services	other
387	MS	F	19.0	R	GT3	T	2.0	3.0	services	other
312	GP	M	19.0	U	GT3	T	1.0	2.0	other	services
367	MS	F	17.0	R	GT3	T	1.0	1.0	other	services
198	GP	F	17.0	U	GT3	T	4.0	4.0	services	teacher
304	GP	M	19.0	U	GT3	T	3.0	3.0	other	other
221	GP	F	17.0	U	GT3	T	1.0	1.0	at_home	other
217	GP	M	18.0	U	LE3	T	3.0	3.0	services	health
383	MS	M	19.0	R	GT3	T	1.0	1.0	other	services
225	GP	F	18.0	R	GT3	T	3.0	1.0	other	other

```
340     GP    F    19.0         U      GT3         T    2.0    1.0    services
services
```

	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
88	...	4.0	4.0	2.0	1.0	1.0	3.0	12.0	11.0	10.0	10.0
118	...	4.0	2.0	4.0	1.0	4.0	5.0	20.0	9.0	7.0	8.0
250	...	4.0	4.0	5.0	2.0	4.0	5.0	0.0	6.0	8.0	8.0
292	...	4.0	4.0	3.0	1.0	1.0	5.0	12.0	12.0	12.0	13.0
239	...	4.0	5.0	4.0	3.0	5.0	2.0	0.0	7.0	7.0	0.0
341	...	4.0	3.0	3.0	2.0	2.0	2.0	0.0	10.0	10.0	0.0
248	...	4.0	3.0	3.0	1.0	3.0	5.0	8.0	3.0	5.0	5.0
213	...	4.0	4.0	4.0	2.0	4.0	5.0	15.0	6.0	7.0	8.0
151	...	4.0	4.0	4.0	3.0	5.0	5.0	6.0	12.0	13.0	14.0
112	...	4.0	3.0	2.0	1.0	1.0	5.0	6.0	10.0	13.0	13.0
353	...	4.0	4.0	4.0	3.0	3.0	5.0	4.0	8.0	8.0	8.0
309	...	4.0	4.0	3.0	1.0	3.0	3.0	18.0	12.0	10.0	10.0
111	...	4.0	3.0	2.0	1.0	1.0	2.0	0.0	7.0	10.0	10.0
278	...	4.0	4.0	4.0	1.0	1.0	4.0	15.0	9.0	8.0	8.0
205	...	4.0	4.0	3.0	3.0	4.0	5.0	28.0	10.0	9.0	9.0
336	...	4.0	4.0	3.0	1.0	2.0	5.0	12.0	14.0	13.0	13.0
159	...	4.0	5.0	5.0	2.0	4.0	5.0	4.0	10.0	12.0	12.0
281	...	4.0	4.0	4.0	3.0	4.0	3.0	19.0	11.0	9.0	10.0
349	...	4.0	5.0	5.0	2.0	5.0	5.0	10.0	11.0	13.0	13.0
313	...	4.0	2.0	2.0	1.0	2.0	1.0	22.0	13.0	10.0	11.0
387	...	4.0	4.0	2.0	1.0	2.0	5.0	0.0	7.0	5.0	0.0
312	...	4.0	5.0	2.0	2.0	2.0	4.0	3.0	13.0	11.0	11.0

367	...	4.0	2.0	1.0	1.0	2.0	1.0	0.0	7.0	6.0 0.0
198	...	4.0	2.0	4.0	2.0	3.0	2.0	24.0	18.0	18.0 18.0
304	...	4.0	4.0	4.0	1.0	1.0	3.0	20.0	15.0	14.0 13.0
221	...	4.0	3.0	4.0	1.0	1.0	5.0	0.0	6.0	5.0 0.0
217	...	4.0	2.0	4.0	2.0	4.0	4.0	13.0	6.0	6.0 8.0
383	...	4.0	3.0	2.0	1.0	3.0	5.0	0.0	6.0	5.0 0.0
225	...	4.0	3.0	3.0	1.0	1.0	4.0	16.0	9.0	8.0 7.0
340	...	4.0	3.0	4.0	1.0	3.0	3.0	4.0	11.0	12.0 11.0

[30 rows x 33 columns]

X_test

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob \
49	GP	F	15.0	U	GT3	T	4.0	4.0	services	teacher
352	MS	M	18.0	U	LE3	T	1.0	3.0	at_home	services
384	MS	M	18.0	R	GT3	T	4.0	2.0	other	other
343	GP	F	17.0	U	GT3	A	2.0	2.0	at_home	at_home
361	MS	M	18.0	R	LE3	T	1.0	1.0	at_home	other
252	GP	M	18.0	U	GT3	T	2.0	1.0	services	services
315	GP	F	19.0	R	GT3	T	2.0	3.0	other	other
162	GP	M	16.0	U	LE3	T	1.0	2.0	other	other
389	MS	F	18.0	U	GT3	T	1.0	1.0	other	other
307	GP	M	19.0	U	GT3	T	4.0	4.0	teacher	services
305	GP	F	18.0	U	GT3	T	2.0	4.0	services	at_home
308	GP	M	19.0	R	GT3	T	3.0	3.0	other	services
310	GP	F	19.0	U	LE3	T	1.0	2.0	services	services
138	GP	M	16.0	U	LE3	T	1.0	1.0	services	

```
other
95        GP   F  15.0       R    GT3      T   1.0   1.0    at_home
other
165       GP   M  16.0       U    GT3      T   3.0   2.0   services
services
40        GP   F  16.0       U    LE3      T   2.0   2.0     other
other
255       GP   M  17.0       U    LE3      T   1.0   1.0    health
other
52        GP   M  15.0       U    LE3      A   4.0   2.0    health
health
44        GP   F  16.0       U    LE3      T   2.0   2.0     other
at_home

     ... famrel freetime  goout  Dalc  Walc health absences      G1
G2    G3
49   ...    4.0      4.0    4.0   1.0   1.0    3.0      2.0     7.0
7.0   7.0
352  ...    4.0      3.0    3.0   2.0   3.0    3.0      7.0     8.0
7.0   8.0
384  ...    4.0      4.0    3.0   2.0   3.0    3.0     14.0     6.0
5.0   5.0
343  ...    4.0      3.0    1.0   1.0   2.0    4.0      0.0     9.0
8.0   0.0
361  ...    4.0      4.0    3.0   2.0   3.0    5.0      2.0    13.0
12.0  12.0
252  ...    4.0      2.0    5.0   2.0   5.0    5.0      4.0     6.0
9.0   8.0
315  ...    4.0      3.0    2.0   1.0   1.0    3.0      9.0    13.0
11.0  11.0
162  ...    4.0      4.0    4.0   2.0   4.0    5.0      0.0     7.0
0.0   0.0
389  ...    4.0      3.0    1.0   1.0   1.0    5.0      0.0     6.0
5.0   0.0
307  ...    4.0      3.0    4.0   1.0   1.0    4.0      9.0     8.0
9.0   8.0
305  ...    4.0      4.0    3.0   1.0   1.0    3.0      8.0    14.0
12.0  12.0
308  ...    4.0      5.0    3.0   1.0   2.0    5.0      0.0    15.0
12.0  12.0
310  ...    4.0      2.0    4.0   2.0   2.0    3.0      0.0     9.0
9.0   0.0
138  ...    4.0      4.0    4.0   1.0   3.0    5.0      0.0    14.0
12.0  12.0
95   ...    4.0      3.0    2.0   1.0   1.0    1.0      2.0     7.0
10.0  10.0
165  ...    4.0      5.0    2.0   1.0   1.0    2.0     16.0    12.0
11.0  12.0
40   ...    4.0      3.0    3.0   1.0   2.0    3.0     25.0     7.0
```

```
      10.0  11.0
255   ...     4.0      4.0     4.0    1.0    2.0    5.0     2.0    7.0
9.0    8.0
52    ...     4.0      5.0     5.0    3.0    4.0    5.0     6.0   11.0
11.0  10.0
44    ...     4.0      3.0     3.0    2.0    2.0    5.0    14.0   10.0
10.0   9.0

[20 rows x 33 columns]

Y_train

     school sex   age address famsize Pstatus  Medu  Fedu        Mjob
Fjob  \
88      GP   M  16.0       U     GT3       T   2.0   2.0   services
other
118     GP   M  17.0       R     GT3       T   1.0   3.0      other
other
250     GP   M  18.0       U     GT3       T   3.0   2.0   services
other
292     GP   F  18.0       U     LE3       T   2.0   1.0   services
at_home
239     GP   M  18.0       U     GT3       T   2.0   2.0      other
services
341     GP   M  18.0       U     GT3       T   4.0   4.0    teacher
services
248     GP   M  18.0       R     LE3       T   3.0   3.0      other
services
213     GP   M  18.0       U     GT3       T   2.0   2.0   services
other
151     GP   M  16.0       U     LE3       T   2.0   1.0    at_home
other
112     GP   F  16.0       U     GT3       T   2.0   2.0    at_home
other
353     MS   M  19.0       R     GT3       T   1.0   1.0      other
other
309     GP   F  19.0       U     LE3       T   1.0   1.0    at_home
other
111     GP   F  16.0       R     GT3       T   3.0   3.0   services
other
278     GP   F  18.0       U     GT3       T   4.0   4.0     health
health
205     GP   F  17.0       U     GT3       T   3.0   4.0    at_home
services
336     GP   F  19.0       R     GT3       A   3.0   1.0   services
at_home
159     GP   M  16.0       U     GT3       T   3.0   3.0      other
services
281     GP   M  17.0       U     LE3       A   3.0   2.0    teacher
services
```

```
349      MS   M  18.0      R      GT3      T   3.0  2.0    other
other
313      GP   F  19.0      U      LE3      T   3.0  2.0  services
other
387      MS   F  19.0      R      GT3      T   2.0  3.0  services
other
312      GP   M  19.0      U      GT3      T   1.0  2.0    other
services
367      MS   F  17.0      R      GT3      T   1.0  1.0    other
services
198      GP   F  17.0      U      GT3      T   4.0  4.0  services
teacher
304      GP   M  19.0      U      GT3      T   3.0  3.0    other
other
221      GP   F  17.0      U      GT3      T   1.0  1.0  at_home
other
217      GP   M  18.0      U      LE3      T   3.0  3.0  services
health
383      MS   M  19.0      R      GT3      T   1.0  1.0    other
services
225      GP   F  18.0      R      GT3      T   3.0  1.0    other
other
340      GP   F  19.0      U      GT3      T   2.0  1.0  services
services

     ... famrel freetime  goout  Dalc  Walc health absences    G1
G2    G3
88   ...    4.0      4.0    2.0   1.0   1.0    3.0     12.0  11.0
10.0  10.0
118  ...    4.0      2.0    4.0   1.0   4.0    5.0     20.0   9.0
7.0   8.0
250  ...    4.0      4.0    5.0   2.0   4.0    5.0      0.0   6.0
8.0   8.0
292  ...    4.0      4.0    3.0   1.0   1.0    5.0     12.0  12.0
12.0  13.0
239  ...    4.0      5.0    4.0   3.0   5.0    2.0      0.0   7.0
7.0   0.0
341  ...    4.0      3.0    3.0   2.0   2.0    2.0      0.0  10.0
10.0   0.0
248  ...    4.0      3.0    3.0   1.0   3.0    5.0      8.0   3.0
5.0   5.0
213  ...    4.0      4.0    4.0   2.0   4.0    5.0     15.0   6.0
7.0   8.0
151  ...    4.0      4.0    4.0   3.0   5.0    5.0      6.0  12.0
13.0  14.0
112  ...    4.0      3.0    2.0   1.0   1.0    5.0      6.0  10.0
13.0  13.0
353  ...    4.0      4.0    4.0   3.0   3.0    5.0      4.0   8.0
8.0   8.0
```

```
309    ...     4.0      4.0      3.0      1.0      3.0      3.0      18.0   12.0
10.0   10.0
111    ...     4.0      3.0      2.0      1.0      1.0      2.0       0.0    7.0
10.0   10.0
278    ...     4.0      4.0      4.0      1.0      1.0      4.0      15.0    9.0
8.0    8.0
205    ...     4.0      4.0      3.0      3.0      4.0      5.0      28.0   10.0
9.0    9.0
336    ...     4.0      4.0      3.0      1.0      2.0      5.0      12.0   14.0
13.0   13.0
159    ...     4.0      5.0      5.0      2.0      4.0      5.0       4.0   10.0
12.0   12.0
281    ...     4.0      4.0      4.0      3.0      4.0      3.0      19.0   11.0
9.0    10.0
349    ...     4.0      5.0      5.0      2.0      5.0      5.0      10.0   11.0
13.0   13.0
313    ...     4.0      2.0      2.0      1.0      2.0      1.0      22.0   13.0
10.0   11.0
387    ...     4.0      4.0      2.0      1.0      2.0      5.0       0.0    7.0
5.0    0.0
312    ...     4.0      5.0      2.0      2.0      2.0      4.0       3.0   13.0
11.0   11.0
367    ...     4.0      2.0      1.0      1.0      2.0      1.0       0.0    7.0
6.0    0.0
198    ...     4.0      2.0      4.0      2.0      3.0      2.0      24.0   18.0
18.0   18.0
304    ...     4.0      4.0      4.0      1.0      1.0      3.0      20.0   15.0
14.0   13.0
221    ...     4.0      3.0      4.0      1.0      1.0      5.0       0.0    6.0
5.0    0.0
217    ...     4.0      2.0      4.0      2.0      4.0      4.0      13.0    6.0
6.0    8.0
383    ...     4.0      3.0      2.0      1.0      3.0      5.0       0.0    6.0
5.0    0.0
225    ...     4.0      3.0      3.0      1.0      1.0      4.0      16.0    9.0
8.0    7.0
340    ...     4.0      3.0      4.0      1.0      3.0      3.0       4.0   11.0
12.0   11.0

[30 rows x 33 columns]

Y_train

     school sex   age address famsize Pstatus  Medu  Fedu      Mjob
Fjob  \
88       GP   M  16.0       U     GT3       T   2.0   2.0  services
other
118      GP   M  17.0       R     GT3       T   1.0   3.0     other
other
250      GP   M  18.0       U     GT3       T   3.0   2.0  services
```

										other
292	GP	F	18.0	U	LE3	T	2.0	1.0	services	at_home
239	GP	M	18.0	U	GT3	T	2.0	2.0	other	services
341	GP	M	18.0	U	GT3	T	4.0	4.0	teacher	services
248	GP	M	18.0	R	LE3	T	3.0	3.0	other	services
213	GP	M	18.0	U	GT3	T	2.0	2.0	services	other
151	GP	M	16.0	U	LE3	T	2.0	1.0	at_home	other
112	GP	F	16.0	U	GT3	T	2.0	2.0	at_home	other
353	MS	M	19.0	R	GT3	T	1.0	1.0	other	other
309	GP	F	19.0	U	LE3	T	1.0	1.0	at_home	other
111	GP	F	16.0	R	GT3	T	3.0	3.0	services	other
278	GP	F	18.0	U	GT3	T	4.0	4.0	health	health
205	GP	F	17.0	U	GT3	T	3.0	4.0	at_home	services
336	GP	F	19.0	R	GT3	A	3.0	1.0	services	at_home
159	GP	M	16.0	U	GT3	T	3.0	3.0	other	services
281	GP	M	17.0	U	LE3	A	3.0	2.0	teacher	services
349	MS	M	18.0	R	GT3	T	3.0	2.0	other	other
313	GP	F	19.0	U	LE3	T	3.0	2.0	services	other
387	MS	F	19.0	R	GT3	T	2.0	3.0	services	other
312	GP	M	19.0	U	GT3	T	1.0	2.0	other	services
367	MS	F	17.0	R	GT3	T	1.0	1.0	other	services
198	GP	F	17.0	U	GT3	T	4.0	4.0	services	teacher
304	GP	M	19.0	U	GT3	T	3.0	3.0	other	other
221	GP	F	17.0	U	GT3	T	1.0	1.0	at_home	other
217	GP	M	18.0	U	LE3	T	3.0	3.0	services	health

```
383      MS   M   19.0          R       GT3        T   1.0   1.0       other
services
225      GP   F   18.0          R       GT3        T   3.0   1.0       other
other
340      GP   F   19.0          U       GT3        T   2.0   1.0    services
services

         ... famrel freetime  goout  Dalc  Walc health absences    G1
G2    G3
88       ...    4.0      4.0    2.0   1.0   1.0    3.0     12.0  11.0
10.0  10.0
118      ...    4.0      2.0    4.0   1.0   4.0    5.0     20.0   9.0
7.0   8.0
250      ...    4.0      4.0    5.0   2.0   4.0    5.0      0.0   6.0
8.0   8.0
292      ...    4.0      4.0    3.0   1.0   1.0    5.0     12.0  12.0
12.0  13.0
239      ...    4.0      5.0    4.0   3.0   5.0    2.0      0.0   7.0
7.0   0.0
341      ...    4.0      3.0    3.0   2.0   2.0    2.0      0.0  10.0
10.0   0.0
248      ...    4.0      3.0    3.0   1.0   3.0    5.0      8.0   3.0
5.0   5.0
213      ...    4.0      4.0    4.0   2.0   4.0    5.0     15.0   6.0
7.0   8.0
151      ...    4.0      4.0    4.0   3.0   5.0    5.0      6.0  12.0
13.0  14.0
112      ...    4.0      3.0    2.0   1.0   1.0    5.0      6.0  10.0
13.0  13.0
353      ...    4.0      4.0    4.0   3.0   3.0    5.0      4.0   8.0
8.0   8.0
309      ...    4.0      4.0    3.0   1.0   3.0    3.0     18.0  12.0
10.0  10.0
111      ...    4.0      3.0    2.0   1.0   1.0    2.0      0.0   7.0
10.0  10.0
278      ...    4.0      4.0    4.0   1.0   1.0    4.0     15.0   9.0
8.0   8.0
205      ...    4.0      4.0    3.0   3.0   4.0    5.0     28.0  10.0
9.0   9.0
336      ...    4.0      4.0    3.0   1.0   2.0    5.0     12.0  14.0
13.0  13.0
159      ...    4.0      5.0    5.0   2.0   4.0    5.0      4.0  10.0
12.0  12.0
281      ...    4.0      4.0    4.0   3.0   4.0    3.0     19.0  11.0
9.0  10.0
349      ...    4.0      5.0    5.0   2.0   5.0    5.0     10.0  11.0
13.0  13.0
313      ...    4.0      2.0    2.0   1.0   2.0    1.0     22.0  13.0
10.0  11.0
```

```
387  ...      4.0     4.0    2.0   1.0   2.0   5.0        0.0    7.0
5.0    0.0
312  ...      4.0     5.0    2.0   2.0   2.0   4.0        3.0   13.0
11.0   11.0
367  ...      4.0     2.0    1.0   1.0   2.0   1.0        0.0    7.0
6.0    0.0
198  ...      4.0     2.0    4.0   2.0   3.0   2.0       24.0   18.0
18.0   18.0
304  ...      4.0     4.0    4.0   1.0   1.0   3.0       20.0   15.0
14.0   13.0
221  ...      4.0     3.0    4.0   1.0   1.0   5.0        0.0    6.0
5.0    0.0
217  ...      4.0     2.0    4.0   2.0   4.0   4.0       13.0    6.0
6.0    8.0
383  ...      4.0     3.0    2.0   1.0   3.0   5.0        0.0    6.0
5.0    0.0
225  ...      4.0     3.0    3.0   1.0   1.0   4.0       16.0    9.0
8.0    7.0
340  ...      4.0     3.0    4.0   1.0   3.0   3.0        4.0   11.0
12.0   11.0

[30 rows x 33 columns]
```

```python
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_train_norm =
scaler.fit_transform(X_train.select_dtypes(include='float64'))

X_train_norm
```

```
array([[-1.67516978, -0.3867394 , -0.16095569,  0.90453403,
0.09805807,
         0.        ,  0.        ,  0.50650222, -1.2528605 , -
0.79240582,
        -1.26360005, -0.64918902,  0.27449115,  0.42724562,
0.16937145,
         0.33551414],
       [-0.72696047, -1.44148324,  0.80477847,  2.41209076,
0.09805807,
         0.        ,  0.        , -1.66422159,  0.67461719, -
0.79240582,
         0.96628239,  0.84893949,  1.22924298, -0.19799187, -
0.78334295,
        -0.0753195 ],
       [ 0.22124884,  0.66800443, -0.16095569,  0.90453403, -
1.37281295,
         0.        ,  0.        ,  0.50650222,  1.63835604,
0.52827054,
         0.96628239,  0.84893949, -1.1576366 , -1.13584811, -
0.46577148,
```

```
        -0.0753195 ],
       [ 0.22124884, -0.3867394 , -1.12668986, -0.60302269,
0.09805807,
         0.        ,  0.        ,  0.50650222, -0.28912165, -
0.79240582,
        -1.26360005,  0.84893949,  0.27449115,  0.73986437,
0.80451438,
         0.95176461],
       [ 0.22124884, -0.3867394 , -0.16095569, -0.60302269,
0.09805807,
         0.        ,  0.        ,  1.59186413,  0.67461719,  1.8489469
,
         1.70957654, -1.39825327, -1.1576366 , -0.82322937, -
0.78334295,
        -1.71865408],
       [ 0.22124884,  1.72274826,  1.77051264, -0.60302269,
0.09805807,
         0.        ,  0.        , -0.57885968, -0.28912165,
0.52827054,
        -0.5203059 , -1.39825327, -1.1576366 ,  0.11462687,
0.16937145,
        -1.71865408],
       [ 0.22124884,  0.66800443,  0.80477847, -0.60302269,
0.09805807,
         0.        ,  0.        , -0.57885968, -0.28912165, -
0.79240582,
         0.22298824,  0.84893949, -0.20288476, -2.07370436, -
1.41848588,
        -0.69156997],
       [ 0.22124884, -0.3867394 , -0.16095569, -0.60302269,
0.09805807,
         0.        ,  0.        ,  0.50650222,  0.67461719,
0.52827054,
         0.96628239,  0.84893949,  0.63252309, -1.13584811, -
0.78334295,
        -0.0753195 ],
       [-1.67516978, -0.3867394 , -1.12668986, -0.60302269, -
1.37281295,
         0.        ,  0.        ,  0.50650222,  0.67461719,  1.8489469
,
         1.70957654,  0.84893949, -0.44157272,  0.73986437,
1.12208585,
         1.15718143],
       [-1.67516978, -0.3867394 , -0.16095569, -0.60302269,
0.09805807,
         0.        ,  0.        , -0.57885968, -1.2528605 , -
0.79240582,
        -1.26360005,  0.84893949, -0.44157272,  0.11462687,
1.12208585,
```

```
         0.95176461],
       [ 1.16945815, -1.44148324, -1.12668986,  2.41209076, -
1.37281295,
         0.        ,  0.        ,  0.50650222,  0.67461719,  1.8489469
,
         0.22298824,  0.84893949, -0.68026068, -0.51061062, -
0.46577148,
        -0.0753195 ],
       [ 1.16945815, -1.44148324, -1.12668986, -0.60302269,
0.09805807,
         0.        ,  0.        ,  0.50650222, -0.28912165, -
0.79240582,
         0.22298824, -0.64918902,  0.99055503,  0.73986437,
0.16937145,
         0.33551414],
       [-1.67516978,  0.66800443,  0.80477847, -0.60302269,
1.56892908,
         0.        ,  0.        , -0.57885968, -1.2528605 , -
0.79240582,
        -1.26360005, -1.39825327, -1.1576366 , -0.82322937,
0.16937145,
         0.33551414],
       [ 0.22124884,  1.72274826,  1.77051264, -0.60302269,
0.09805807,
         0.        ,  0.        ,  0.50650222,  0.67461719, -
0.79240582,
        -1.26360005,  0.09987523,  0.63252309, -0.19799187, -
0.46577148,
        -0.0753195 ],
       [-0.72696047,  0.66800443,  1.77051264, -0.60302269,
1.56892908,
         0.        ,  0.        ,  0.50650222, -0.28912165,  1.8489469
,
         0.96628239,  0.84893949,  2.18399482,  0.11462687, -
0.14820002,
         0.13009732],
       [ 1.16945815,  0.66800443, -1.12668986, -0.60302269,
1.56892908,
         0.        ,  0.        ,  0.50650222, -0.28912165, -
0.79240582,
        -0.5203059 ,  0.84893949,  0.27449115,  1.36510186,
1.12208585,
         0.95176461],
       [-1.67516978,  0.66800443,  0.80477847, -0.60302269,
0.09805807,
         0.        ,  0.        ,  1.59186413,  1.63835604,
0.52827054,
         0.96628239,  0.84893949, -0.68026068,  0.11462687,
0.80451438,
```

```
         0.74634779],
        [-0.72696047,  0.66800443, -0.16095569, -0.60302269, -
1.37281295,
         0.        ,  0.        ,  0.50650222,  0.67461719,  1.8489469
,
         0.96628239, -0.64918902,  1.109899  ,  0.42724562, -
0.14820002,
         0.33551414],
        [ 0.22124884,  0.66800443, -0.16095569,  0.90453403, -
1.37281295,
         0.        ,  0.        ,  1.59186413,  1.63835604,
0.52827054,
         1.70957654,  0.84893949,  0.03580319,  0.42724562,
1.12208585,
         0.95176461],
        [ 1.16945815,  0.66800443, -0.16095569,  0.90453403,
0.09805807,
         0.        ,  0.        , -1.66422159, -1.2528605 , -
0.79240582,
        -0.5203059 , -2.14731753,  1.46793094,  1.05248312,
0.16937145,
         0.54093097],
        [ 1.16945815, -0.3867394 ,  0.80477847, -0.60302269,
1.56892908,
         0.        ,  0.        ,  0.50650222, -1.2528605 , -
0.79240582,
        -0.5203059 ,  0.84893949, -1.1576366 , -0.82322937, -
1.41848588,
        -1.71865408],
        [ 1.16945815, -1.44148324, -0.16095569, -0.60302269,
0.09805807,
         0.        ,  0.        ,  1.59186413, -1.2528605 ,
0.52827054,
        -0.5203059 ,  0.09987523, -0.79960466,  1.05248312,
0.48694291,
         0.54093097],
        [-0.72696047, -1.44148324, -1.12668986,  2.41209076, -
1.37281295,
         0.        ,  0.        , -1.66422159, -2.21659935, -
0.79240582,
        -0.5203059 , -2.14731753, -1.1576366 , -0.82322937, -
1.10091442,
        -1.71865408],
        [-0.72696047,  1.72274826,  1.77051264,  0.90453403, -
1.37281295,
         0.        ,  0.        , -1.66422159,  0.67461719,
0.52827054,
         0.22298824, -1.39825327,  1.7066189 ,  2.61557685,
2.70994318,
```

```
           1.97884873],
       [ 1.16945815,  0.66800443,  0.80477847, -0.60302269,
0.09805807,
          0.        ,  0.        ,  0.50650222,  0.67461719, -
0.79240582,
         -1.26360005, -0.64918902,  1.22924298,  1.67772061,
1.43965731,
          0.95176461],
       [-0.72696047, -1.44148324, -1.12668986, -0.60302269,
1.56892908,
          0.        ,  0.        , -0.57885968,  0.67461719, -
0.79240582,
         -1.26360005,  0.84893949, -1.1576366 , -1.13584811, -
1.41848588,
         -1.71865408],
       [ 0.22124884,  0.66800443,  0.80477847, -0.60302269,
0.09805807,
          0.        ,  0.        , -1.66422159,  0.67461719,
0.52827054,
          0.96628239,  0.09987523,  0.39383513, -1.13584811, -
1.10091442,
         -0.0753195 ],
       [ 1.16945815, -1.44148324, -1.12668986,  0.90453403, -
1.37281295,
          0.        ,  0.        , -0.57885968, -1.2528605 , -
0.79240582,
          0.22298824,  0.84893949, -1.1576366 , -1.13584811, -
1.41848588,
         -1.71865408],
       [ 0.22124884,  0.66800443, -1.12668986, -0.60302269,
0.09805807,
          0.        ,  0.        , -0.57885968, -0.28912165, -
0.79240582,
         -1.26360005,  0.09987523,  0.75186707, -0.19799187, -
0.46577148,
         -0.28073632],
       [ 1.16945815, -0.3867394 , -1.12668986, -0.60302269,
1.56892908,
          0.        ,  0.        , -0.57885968,  0.67461719, -
0.79240582,
          0.22298824, -0.64918902, -0.68026068,  0.42724562,
0.80451438,
          0.54093097]])

fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=X_train_norm,ax=ax)

<Axes: >
```
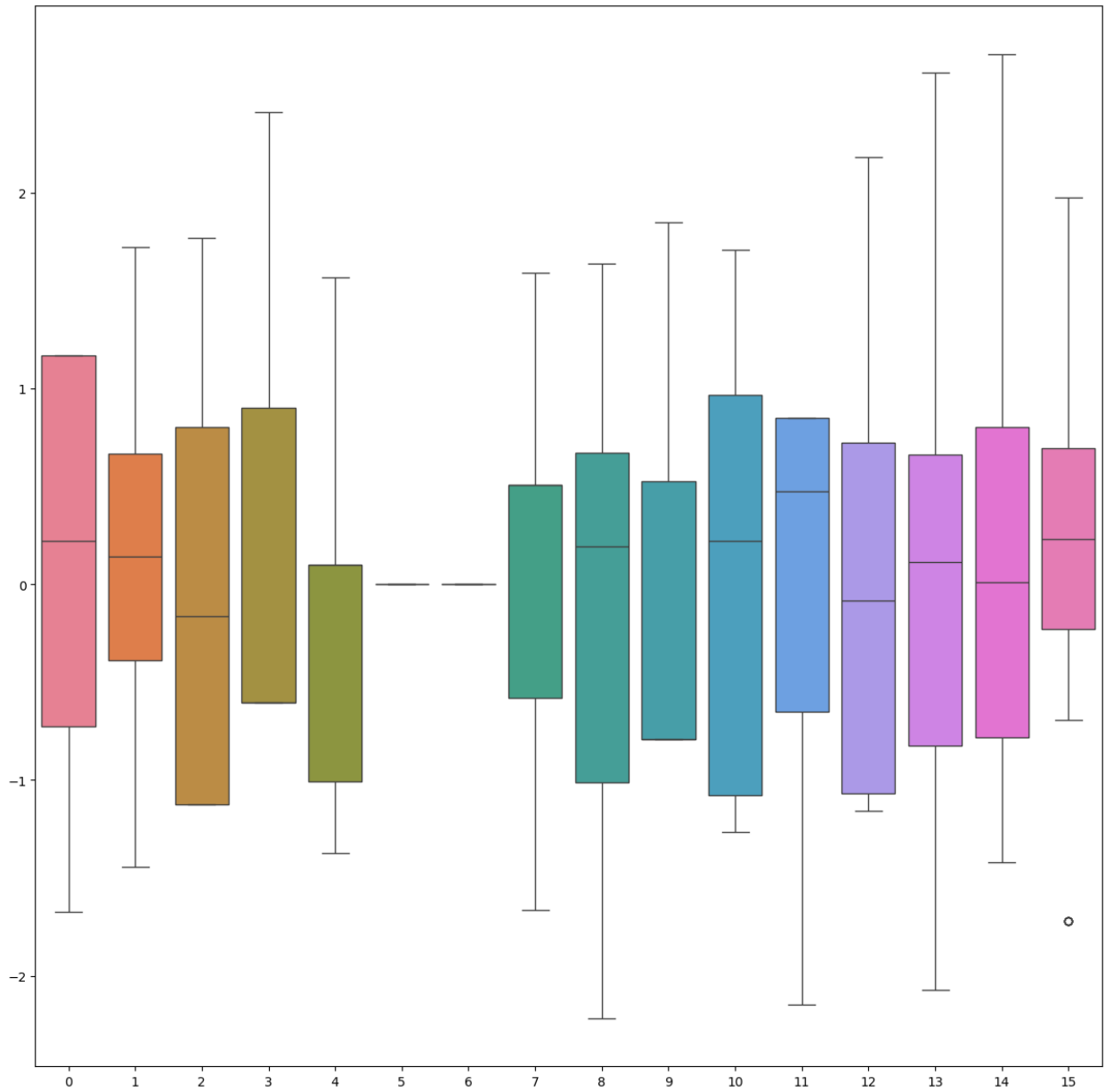
```
X_test_norm =
scaler.transform(X_test.select_dtypes(include='float64'))

X_test_norm

array([[-2.6233791 ,  1.72274826,  1.77051264, -0.60302269,
0.09805807,
          0.        ,  0.        ,  0.50650222,  0.67461719, -
0.79240582,
         -1.26360005, -0.64918902, -0.91894864, -0.82322937, -
0.78334295,
         -0.28073632],
        [ 0.22124884, -1.44148324,  0.80477847, -0.60302269, -
```

```
 1.37281295,
          0.        ,  0.        , -0.57885968, -0.28912165,
 0.52827054,
          0.22298824, -0.64918902, -0.32222874, -0.51061062, -
 0.78334295,
         -0.0753195 ],
        [ 0.22124884,  1.72274826, -0.16095569,  0.90453403, -
 1.37281295,
          0.        ,  0.        ,  0.50650222, -0.28912165,
 0.52827054,
          0.22298824, -0.64918902,  0.51317911, -1.13584811, -
 1.41848588,
         -0.69156997],
        [-0.72696047, -0.3867394 , -0.16095569, -0.60302269,
 0.09805807,
          0.        ,  0.        , -0.57885968, -2.21659935, -
 0.79240582,
         -0.5203059 ,  0.09987523, -1.1576366 , -0.19799187, -
 0.46577148,
         -1.71865408],
        [ 0.22124884, -1.44148324, -1.12668986,  0.90453403,
 0.09805807,
          0.        ,  0.        ,  0.50650222, -0.28912165,
 0.52827054,
          0.22298824,  0.84893949, -0.91894864,  1.05248312,
 0.80451438,
          0.74634779],
        [ 0.22124884, -0.3867394 , -1.12668986, -0.60302269, -
 1.37281295,
          0.        ,  0.        , -1.66422159,  1.63835604,
 0.52827054,
          1.70957654,  0.84893949, -0.68026068, -1.13584811, -
 0.14820002,
         -0.0753195 ],
        [ 1.16945815, -0.3867394 ,  0.80477847, -0.60302269,
 1.56892908,
          0.        ,  0.        , -0.57885968, -1.2528605 , -
 0.79240582,
         -1.26360005, -0.64918902, -0.08354079,  1.05248312,
 0.48694291,
          0.54093097],
        [-1.67516978, -1.44148324, -0.16095569,  0.90453403, -
 1.37281295,
          0.        ,  0.        ,  0.50650222,  0.67461719,
 0.52827054,
          0.96628239,  0.84893949, -1.1576366 , -0.82322937, -
 3.00634321,
         -1.71865408],
        [ 0.22124884, -1.44148324, -1.12668986,  0.90453403,
```

```
0.09805807,
        0.        ,  0.        , -0.57885968, -2.21659935, -
0.79240582,
       -1.26360005,  0.84893949, -1.1576366 , -1.13584811, -
1.41848588,
       -1.71865408],
      [ 1.16945815,  1.72274826,  1.77051264,  0.90453403, -
1.37281295,
        0.        ,  0.        , -0.57885968,  0.67461719, -
0.79240582,
       -1.26360005,  0.09987523, -0.08354079, -0.51061062, -
0.14820002,
       -0.0753195 ],
      [ 0.22124884, -0.3867394 ,  1.77051264, -0.60302269,
0.09805807,
        0.        ,  0.        ,  0.50650222, -0.28912165, -
0.79240582,
       -1.26360005, -0.64918902, -0.20288476,  1.36510186,
0.80451438,
        0.74634779],
      [ 1.16945815,  0.66800443,  0.80477847, -0.60302269,
0.09805807,
        0.        ,  0.        ,  1.59186413, -0.28912165, -
0.79240582,
       -0.5203059 ,  0.84893949, -1.1576366 ,  1.67772061,
0.80451438,
        0.74634779],
      [ 1.16945815, -1.44148324, -0.16095569, -0.60302269,
0.09805807,
        0.        ,  0.        , -1.66422159,  0.67461719,
0.52827054,
       -0.5203059 , -0.64918902, -1.1576366 , -0.19799187, -
0.14820002,
       -1.71865408],
      [-1.67516978, -1.44148324, -1.12668986, -0.60302269,
0.09805807,
        0.        ,  0.        ,  0.50650222,  0.67461719, -
0.79240582,
        0.22298824,  0.84893949, -1.1576366 ,  1.36510186,
0.80451438,
        0.74634779],
      [-2.6233791 , -1.44148324, -1.12668986,  0.90453403,
0.09805807,
        0.        ,  0.        , -0.57885968, -1.2528605 , -
0.79240582,
       -1.26360005, -2.14731753, -0.91894864, -0.82322937,
0.16937145,
        0.33551414],
      [-1.67516978,  0.66800443, -0.16095569,  0.90453403, -
```

```
1.37281295,
        0.        ,  0.        ,  1.59186413, -1.2528605 , -
0.79240582,
       -1.26360005, -1.39825327,  0.75186707,  0.73986437,
0.48694291,
        0.74634779],
      [-1.67516978, -0.3867394 , -0.16095569,  0.90453403,
0.09805807,
        0.        ,  0.        , -0.57885968, -0.28912165, -
0.79240582,
       -0.5203059 , -0.64918902,  1.82596288, -0.82322937,
0.16937145,
        0.54093097],
      [-0.72696047, -1.44148324, -1.12668986,  0.90453403, -
1.37281295,
        0.        ,  0.        ,  0.50650222,  0.67461719, -
0.79240582,
       -0.5203059 ,  0.84893949, -0.91894864, -0.82322937, -
0.14820002,
       -0.0753195 ],
      [-2.6233791 ,  1.72274826, -0.16095569,  0.90453403, -
1.37281295,
        0.        ,  0.        ,  1.59186413,  1.63835604,  1.8489469
,
        0.96628239,  0.84893949, -0.44157272,  0.42724562,
0.48694291,
        0.33551414],
      [-1.67516978, -0.3867394 , -0.16095569,  0.90453403,
0.09805807,
        0.        ,  0.        , -0.57885968, -0.28912165,
0.52827054,
       -0.5203059 ,  0.84893949,  0.51317911,  0.11462687,
0.16937145,
        0.13009732]])

fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=X_test_norm,ax=ax)

<Axes: >
```
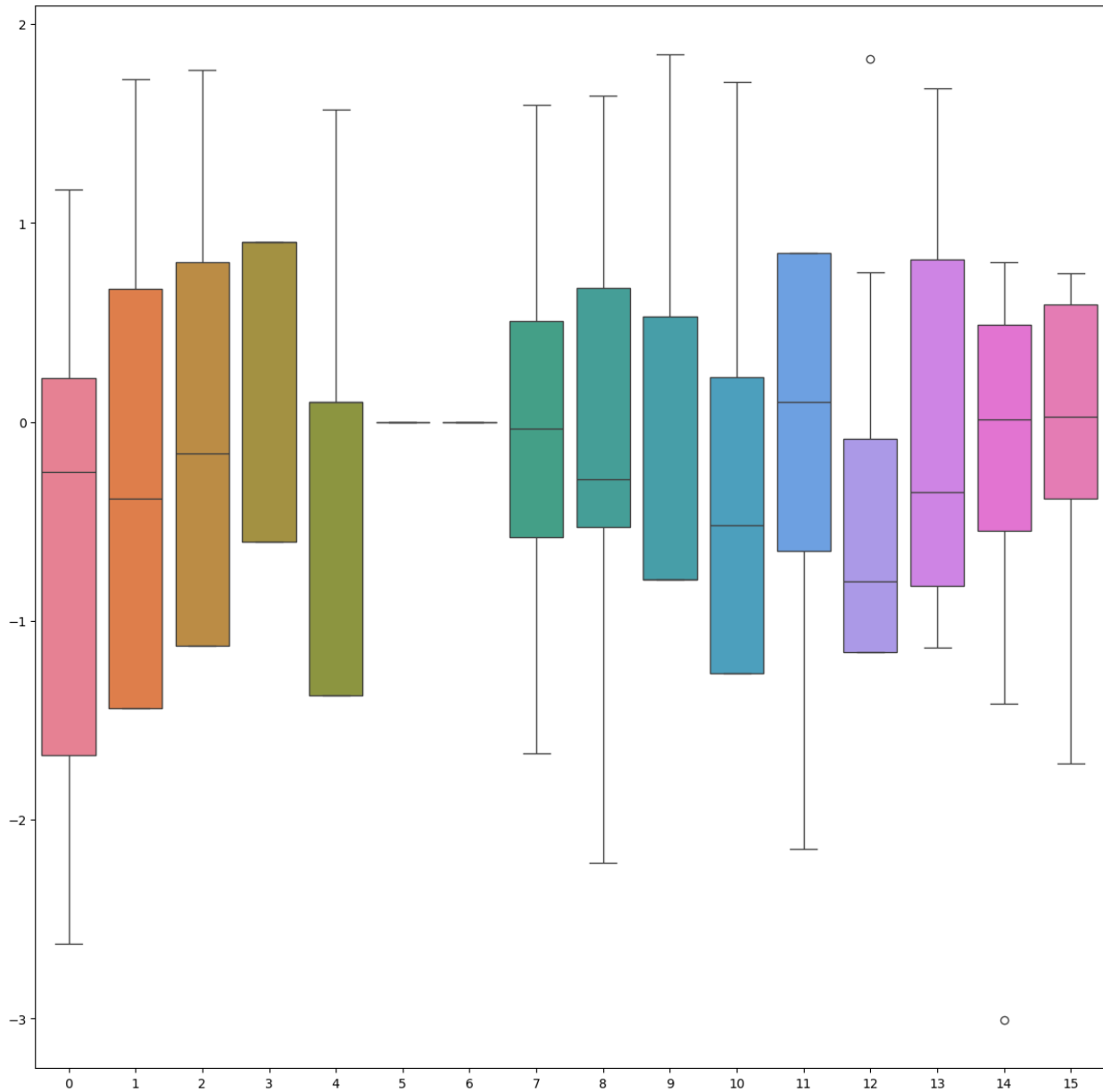
```
from sklearn.linear_model import LinearRegression
fail_1_regression = LinearRegression()
fail_1_regression.fit(X_train_norm,Y_train.select_dtypes(include='floa
t64'))
```

```
LinearRegression()
```

```
print(fail_1_regression.coef_)
```

```
[[ 1.05461947e+00 -2.71211837e-16 -3.98822574e-16 -9.21232772e-17
  -1.29101966e-16 -3.76201908e-17 -8.26860808e-18  3.01570692e-16
  -6.19971286e-17  2.94249842e-16  4.15219306e-17 -3.75335317e-16
  -4.34149225e-16 -2.21052588e-16  1.77870971e-16 -3.37665232e-16]]
```

```
[ 2.94424503e-16   9.48097510e-01   3.91378098e-16  -4.28222927e-16
  -6.49841216e-16  -1.38777878e-16   5.55111512e-17   4.26398872e-16
  -5.17097897e-16   5.84169146e-16  -1.21564649e-16  -9.49760581e-16
  -3.16377671e-16  -3.35691091e-16   6.79969581e-16  -6.25190596e-16]
[ 1.23050527e-16   4.45894245e-16   1.03548164e+00  -7.00776753e-16
   1.27148446e-16   2.77555756e-17  -5.55111512e-17   7.51756292e-18
   4.75507315e-16  -1.02438000e-16  -1.29248521e-16  -3.54646238e-16
  -4.78826526e-16  -2.73323534e-16   2.47026326e-16  -1.18997617e-16]
[ 1.78475339e-16  -2.83384755e-16  -1.44124081e-16   6.63324958e-01
  -3.89924602e-16  -6.93889390e-18  -5.55111512e-17  -3.65803782e-16
  -2.91023635e-16   3.51567152e-16   1.32670167e-16  -6.64262678e-16
  -4.33388200e-16   2.13965593e-16   3.03702095e-17   1.82103964e-16]
[-2.88495143e-16   1.42575353e-16   2.46322981e-16  -1.55979710e-16
   6.79869268e-01   8.32667268e-17  -8.32667268e-17  -1.75885275e-16
   1.46480573e-16  -4.54319663e-16   1.22874518e-16   8.34162402e-16
   2.65055649e-16   5.63917094e-17  -8.13468790e-16   4.84652962e-16]
[ 0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00]
[ 0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00
   0.00000000e+00   0.00000000e+00   0.00000000e+00   0.00000000e+00]
[ 4.71452820e-17  -1.71242058e-17  -7.12810182e-17  -2.23761941e-16
  -1.67994947e-16  -5.55111512e-17  -5.55111512e-17   9.21351664e-01
   7.24024906e-16  -4.44653633e-16  -4.35294092e-16   1.69098204e-16
   9.93800090e-17  -1.54678827e-16   4.53549715e-17  -2.00732903e-17]
[ 3.93210351e-16   1.62485627e-17   3.27995402e-16  -5.01657994e-16
   4.11506491e-16   1.80411242e-16   0.00000000e+00   2.18035706e-16
   1.03762549e+00   9.04215278e-16   3.10308548e-16  -8.29287921e-16
  -1.66066704e-15   1.31751003e-15  -7.16297973e-16   4.03413565e-16]
[-4.19197433e-17   2.07955091e-16   1.77982409e-17   7.48151174e-16
   2.59513872e-16  -5.55111512e-17  -5.55111512e-17  -7.16859737e-16
  -5.91988253e-17   7.57187779e-01   2.32768220e-16   1.48117260e-15
   3.83719269e-16  -3.16426778e-16   1.67759670e-16   5.99222568e-16]
[-5.71570188e-16  -7.40111230e-16  -1.26317334e-16   1.01568054e-15
   6.50452633e-16   2.22044605e-16   0.00000000e+00   3.75719254e-16
   3.49287558e-16  -4.41008308e-16   1.34536240e+00   3.51944837e-16
   1.60625883e-15  -9.47427570e-16   2.45355506e-16   2.22933541e-16]
[ 1.59799920e-16   6.41145291e-16  -6.24520101e-16  -1.22322595e-15
   4.68397770e-16   0.00000000e+00   2.77555756e-17   3.38205273e-16
  -4.72429205e-16   1.05609230e-15   1.70673712e-16   1.33499896e+00
  -6.10409696e-16  -2.14308610e-16   8.10750783e-16  -4.80040403e-16]
[-5.37308877e-15  -1.27885533e-15  -7.47008508e-16   4.04559434e-15
   4.07917025e-15  -2.66453526e-15   4.44089210e-16  -6.21352538e-15
   8.07748817e-17  -4.03794397e-15  -1.23913268e-16   9.37920612e-15
   8.37914077e+00   3.46668377e-15  -8.93400524e-15   3.13096627e-15]
[ 4.28957252e-16  -1.89211744e-15  -8.28918807e-16   5.21621081e-16
```

```
    -2.01870150e-15  1.27675648e-15  0.00000000e+00 -9.81514868e-16
     3.86912875e-16 -2.61425178e-16  1.53233283e-16  7.67510088e-17
    -5.85281095e-17  3.19878449e+00 -3.58446699e-15  1.60980831e-15]
   [ 1.35588299e-15  1.69222598e-17 -1.23890525e-16 -2.94039678e-16
    -1.20868334e-15 -4.44089210e-16 -6.66133815e-16  2.63445679e-16
    -8.82801001e-16  1.59294715e-16  1.77184554e-16 -9.21298342e-16
    -2.98005839e-16 -8.84428771e-16  3.14889751e+00 -2.05185043e-15]
   [-1.22348987e-15 -1.18005160e-16 -1.74762463e-15 -2.88468910e-16
     1.80488131e-15 -2.22044605e-16  8.88178420e-16  2.12638568e-15
     1.22748717e-15  3.00162103e-16  1.62894644e-15 -2.33491792e-15
     8.24763298e-16  1.25573661e-15 -2.02451606e-15  4.86815046e+00]]
```

```
fail_1_predict = fail_1_regression.predict(X_test_norm)
fail_1_predict
```

```
array([[ 1.50000000e+01,  4.00000000e+00,  4.00000000e+00,
         1.00000000e+00,  2.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  4.00000000e+00,  4.00000000e+00,
         1.00000000e+00,  1.00000000e+00,  3.00000000e+00,
         2.00000000e+00,  7.00000000e+00,  7.00000000e+00,
         7.00000000e+00],
       [ 1.80000000e+01,  1.00000000e+00,  3.00000000e+00,
         1.00000000e+00,  1.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  3.00000000e+00,  3.00000000e+00,
         2.00000000e+00,  3.00000000e+00,  3.00000000e+00,
         7.00000000e+00,  8.00000000e+00,  7.00000000e+00,
         8.00000000e+00],
       [ 1.80000000e+01,  4.00000000e+00,  2.00000000e+00,
         2.00000000e+00,  1.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  4.00000000e+00,  3.00000000e+00,
         2.00000000e+00,  3.00000000e+00,  3.00000000e+00,
         1.40000000e+01,  6.00000000e+00,  5.00000000e+00,
         5.00000000e+00],
       [ 1.70000000e+01,  2.00000000e+00,  2.00000000e+00,
         1.00000000e+00,  2.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  3.00000000e+00,  1.00000000e+00,
         1.00000000e+00,  2.00000000e+00,  4.00000000e+00,
        -3.55271368e-15,  9.00000000e+00,  8.00000000e+00,
        -5.32907052e-15],
       [ 1.80000000e+01,  1.00000000e+00,  1.00000000e+00,
         2.00000000e+00,  2.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  4.00000000e+00,  3.00000000e+00,
         2.00000000e+00,  3.00000000e+00,  5.00000000e+00,
         2.00000000e+00,  1.30000000e+01,  1.20000000e+01,
         1.20000000e+01],
       [ 1.80000000e+01,  2.00000000e+00,  1.00000000e+00,
         1.00000000e+00,  1.00000000e+00,  1.00000000e+00,
         4.00000000e+00,  2.00000000e+00,  5.00000000e+00,
         2.00000000e+00,  5.00000000e+00,  5.00000000e+00,
         4.00000000e+00,  6.00000000e+00,  9.00000000e+00,
```

```
         8.00000000e+00],
       [ 1.90000000e+01,   2.00000000e+00,   3.00000000e+00,
         1.00000000e+00,   3.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   2.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   3.00000000e+00,
         9.00000000e+00,   1.30000000e+01,   1.10000000e+01,
         1.10000000e+01],
       [ 1.60000000e+01,   1.00000000e+00,   2.00000000e+00,
         2.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   4.00000000e+00,   4.00000000e+00,
         2.00000000e+00,   4.00000000e+00,   5.00000000e+00,
         1.77635684e-14,   7.00000000e+00,  -1.59872116e-14,
         3.55271368e-15],
       [ 1.80000000e+01,   1.00000000e+00,   1.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   1.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   5.00000000e+00,
         1.24344979e-14,   6.00000000e+00,   5.00000000e+00,
        -7.10542736e-15],
       [ 1.90000000e+01,   4.00000000e+00,   4.00000000e+00,
         2.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   4.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   4.00000000e+00,
         9.00000000e+00,   8.00000000e+00,   9.00000000e+00,
         8.00000000e+00],
       [ 1.80000000e+01,   2.00000000e+00,   4.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   4.00000000e+00,   3.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   3.00000000e+00,
         8.00000000e+00,   1.40000000e+01,   1.20000000e+01,
         1.20000000e+01],
       [ 1.90000000e+01,   3.00000000e+00,   3.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   5.00000000e+00,   3.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   5.00000000e+00,
        -1.95399252e-14,   1.50000000e+01,   1.20000000e+01,
         1.20000000e+01],
       [ 1.90000000e+01,   1.00000000e+00,   2.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   2.00000000e+00,   4.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   3.00000000e+00,
        -1.95399252e-14,   9.00000000e+00,   9.00000000e+00,
        -5.32907052e-15],
       [ 1.60000000e+01,   1.00000000e+00,   1.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   4.00000000e+00,   4.00000000e+00,
         1.00000000e+00,   3.00000000e+00,   5.00000000e+00,
         5.32907052e-15,   1.40000000e+01,   1.20000000e+01,
         1.20000000e+01],
```

```
       [ 1.50000000e+01,   1.00000000e+00,   1.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   2.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         2.00000000e+00,   7.00000000e+00,   1.00000000e+01,
         1.00000000e+01],
       [ 1.60000000e+01,   3.00000000e+00,   2.00000000e+00,
         2.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   5.00000000e+00,   2.00000000e+00,
         1.00000000e+00,   1.00000000e+00,   2.00000000e+00,
         1.60000000e+01,   1.20000000e+01,   1.10000000e+01,
         1.20000000e+01],
       [ 1.60000000e+01,   2.00000000e+00,   2.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   3.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   3.00000000e+00,
         2.50000000e+01,   7.00000000e+00,   1.00000000e+01,
         1.10000000e+01],
       [ 1.70000000e+01,   1.00000000e+00,   1.00000000e+00,
         2.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   4.00000000e+00,   4.00000000e+00,
         1.00000000e+00,   2.00000000e+00,   5.00000000e+00,
         2.00000000e+00,   7.00000000e+00,   9.00000000e+00,
         8.00000000e+00],
       [ 1.50000000e+01,   4.00000000e+00,   2.00000000e+00,
         2.00000000e+00,   1.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   5.00000000e+00,   5.00000000e+00,
         3.00000000e+00,   4.00000000e+00,   5.00000000e+00,
         6.00000000e+00,   1.10000000e+01,   1.10000000e+01,
         1.00000000e+01],
       [ 1.60000000e+01,   2.00000000e+00,   2.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   1.00000000e+00,
         4.00000000e+00,   3.00000000e+00,   3.00000000e+00,
         2.00000000e+00,   2.00000000e+00,   5.00000000e+00,
         1.40000000e+01,   1.00000000e+01,   1.00000000e+01,
         9.00000000e+00]])

fail_1_residual = Y_test.select_dtypes(include='float64')-
fail_1_predict
fail_1_residual
```

	age	Medu	Fedu	traveltime	studytime
49	3.552714e-15	-1.776357e-15	-3.552714e-15	1.554312e-15	-1.332268e-15
352	0.000000e+00	-1.110223e-15	-4.440892e-16	-1.998401e-15	3.330669e-16
384	0.000000e+00	-2.664535e-15	4.440892e-16	0.000000e+00	3.330669e-16
343	0.000000e+00	-1.332268e-15	2.220446e-16	-4.440892e-16	

```
4.440892e-16
361  0.000000e+00  2.220446e-15  2.220446e-15 -8.881784e-16
8.881784e-16
252  0.000000e+00  1.110223e-15  0.000000e+00 -8.881784e-16 -
4.440892e-16
315  0.000000e+00  0.000000e+00 -8.881784e-16 -6.661338e-16
4.440892e-16
162  0.000000e+00  2.220446e-15  1.776357e-15  0.000000e+00 -
1.776357e-15
389  0.000000e+00  8.881784e-16  2.886580e-15 -4.440892e-16
0.000000e+00
307  0.000000e+00 -1.776357e-15 -1.776357e-15  8.881784e-16 -
6.661338e-16
305  0.000000e+00 -1.332268e-15 -1.776357e-15 -2.220446e-16
2.220446e-16
308  0.000000e+00 -8.881784e-16 -1.776357e-15  5.551115e-16 -
4.440892e-16
310 -3.552714e-15 -1.110223e-15 -8.881784e-16 -1.776357e-15
1.998401e-15
138  0.000000e+00  3.108624e-15  1.110223e-15  0.000000e+00 -
8.881784e-16
95   0.000000e+00 -2.220446e-16  1.776357e-15 -1.776357e-15
1.776357e-15
165  1.776357e-15 -1.776357e-15  1.332268e-15  0.000000e+00
1.332268e-15
40   1.776357e-15  1.554312e-15  1.554312e-15  8.881784e-16 -
4.440892e-16
255  0.000000e+00  2.220446e-15  1.998401e-15  2.220446e-16
0.000000e+00
52   1.776357e-15  0.000000e+00  2.220446e-16  6.661338e-16
0.000000e+00
44   1.776357e-15  1.776357e-15  1.776357e-15  4.440892e-16 -
8.881784e-16

     failures  famrel      freetime           goout          Dalc
Walc  \
49        0.0     0.0 -8.881784e-16  0.000000e+00  2.220446e-15
6.661338e-16
352       0.0     0.0  4.440892e-16 -1.776357e-15  1.554312e-15
1.776357e-15
384       0.0     0.0  4.440892e-16  1.332268e-15  4.440892e-16
4.440892e-16
343       0.0     0.0  8.881784e-16 -8.881784e-16  1.443290e-15
2.664535e-15
361       0.0     0.0  4.440892e-16 -1.776357e-15 -1.332268e-15
0.000000e+00
252       0.0     0.0 -4.440892e-16  1.776357e-15 -1.776357e-15
1.776357e-15
315       0.0     0.0  4.440892e-16 -2.664535e-15  9.992007e-16
```

```
1.110223e-15
162        0.0      0.0  0.000000e+00   0.000000e+00   4.440892e-16
0.000000e+00
389        0.0      0.0  8.881784e-16   1.776357e-15  -4.440892e-16 -
2.220446e-16
307        0.0      0.0 -1.332268e-15   1.776357e-15  -8.881784e-16
1.332268e-15
305        0.0      0.0 -8.881784e-16  -2.664535e-15   2.220446e-15
1.332268e-15
308        0.0      0.0 -8.881784e-16  -3.996803e-15   8.881784e-16
3.552714e-15
310        0.0      0.0  2.220446e-16  -1.776357e-15   1.998401e-15
2.886580e-15
138        0.0      0.0 -8.881784e-16  -1.776357e-15   2.220446e-16
0.000000e+00
95         0.0      0.0  8.881784e-16   2.220446e-16   2.553513e-15 -
2.220446e-15
165        0.0      0.0  0.000000e+00   1.554312e-15   2.442491e-15 -
1.776357e-15
40         0.0      0.0  0.000000e+00   5.329071e-15  -1.110223e-15 -
5.773160e-15
255        0.0      0.0 -1.776357e-15   2.664535e-15  -4.440892e-16 -
2.220446e-15
52         0.0      0.0 -8.881784e-16   0.000000e+00  -1.332268e-15
0.000000e+00
44         0.0      0.0  4.440892e-16   2.220446e-15  -2.664535e-15 -
2.664535e-15

           health      absences             G1             G2
G3
49    8.881784e-16   3.552714e-15   5.329071e-15   7.105427e-15 -
8.881784e-16
352   1.332268e-15   1.065814e-14  -7.105427e-15   1.776357e-15
2.664535e-15
384   8.881784e-16   3.552714e-15  -1.776357e-15   2.664535e-15 -
1.776357e-15
343  -1.776357e-15   3.552714e-15   1.776357e-15  -1.776357e-15
5.329071e-15
361  -8.881784e-16   2.664535e-15  -3.552714e-15  -3.552714e-15
0.000000e+00
252  -8.881784e-16   1.776357e-15  -7.105427e-15   0.000000e+00
2.664535e-15
315   0.000000e+00   1.776357e-15   3.552714e-15  -1.776357e-15
3.552714e-15
162   2.664535e-15  -1.776357e-14  -1.243450e-14   1.598721e-14 -
3.552714e-15
389   0.000000e+00  -1.243450e-14  -5.329071e-15   2.664535e-15
7.105427e-15
307   3.108624e-15   5.329071e-15   0.000000e+00  -1.776357e-15
```

```
1.065814e-14
305  1.332268e-15   1.243450e-14   3.552714e-15  -3.552714e-15
3.552714e-15
308 -1.776357e-15   1.953993e-14   3.552714e-15  -3.552714e-15
3.552714e-15
310 -8.881784e-16   1.953993e-14  -1.776357e-15  -3.552714e-15
5.329071e-15
138  0.000000e+00  -5.329071e-15  -1.776357e-15   0.000000e+00 -
3.552714e-15
95   3.108624e-15   5.329071e-15  -1.776357e-15   0.000000e+00 -
3.552714e-15
165  2.442491e-15   3.552714e-15   0.000000e+00  -1.776357e-15 -
3.552714e-15
40   3.552714e-15  -2.842171e-14   0.000000e+00   1.776357e-15 -
1.776357e-15
255  1.776357e-15  -1.776357e-15  -6.217249e-15   0.000000e+00
2.664535e-15
52  -1.776357e-15   6.217249e-15   3.552714e-15   1.776357e-15 -
5.329071e-15
44   8.881784e-16  -2.842171e-14   0.000000e+00   3.552714e-15
1.776357e-15
```

```
sns.displot(fail_1_residual,kind = 'kde')
```

```
C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\3240972038.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(fail_1_residual,kind = 'kde')
```

```
<seaborn.axisgrid.FacetGrid at 0x2029b83ecf0>
```
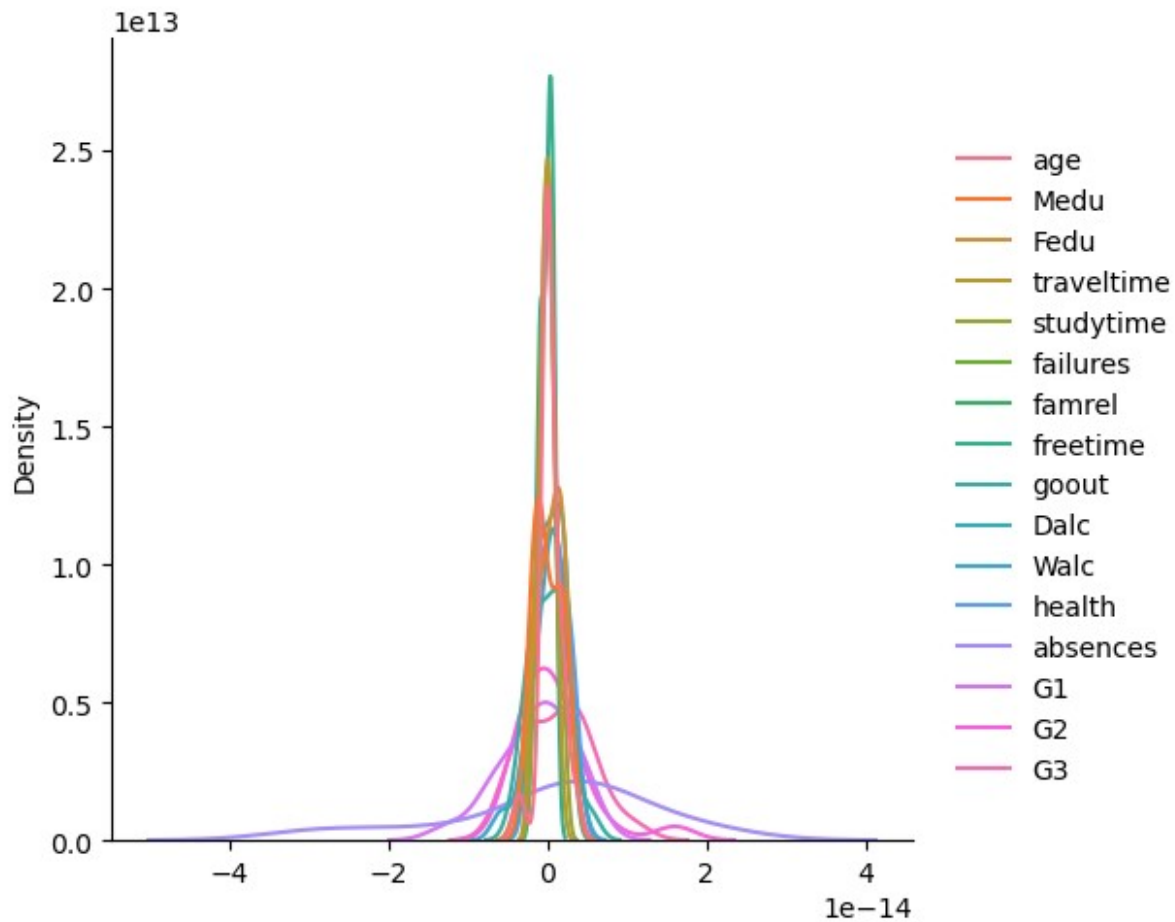
```python
from sklearn.metrics import
mean_squared_error,mean_absolute_error,r2_score
print("Mean Squared
Error",mean_squared_error(Y_test.select_dtypes("float64").dropna(),fai
l_1_predict))
print("Mean Absolute
Error",mean_absolute_error(Y_test.select_dtypes("float64").dropna(),fa
il_1_predict))
print("R2_Score",r2_score(Y_test.select_dtypes("float64").dropna(),fai
l_1_predict))

Mean Squared Error 1.5678379379674284e-29
Mean Absolute Error 1.9838297671270764e-15
R2_Score 1.0
```

Handling Fail_2 Outliers

```python
#Before Handling Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_2,ax=ax)
```

<Axes: >
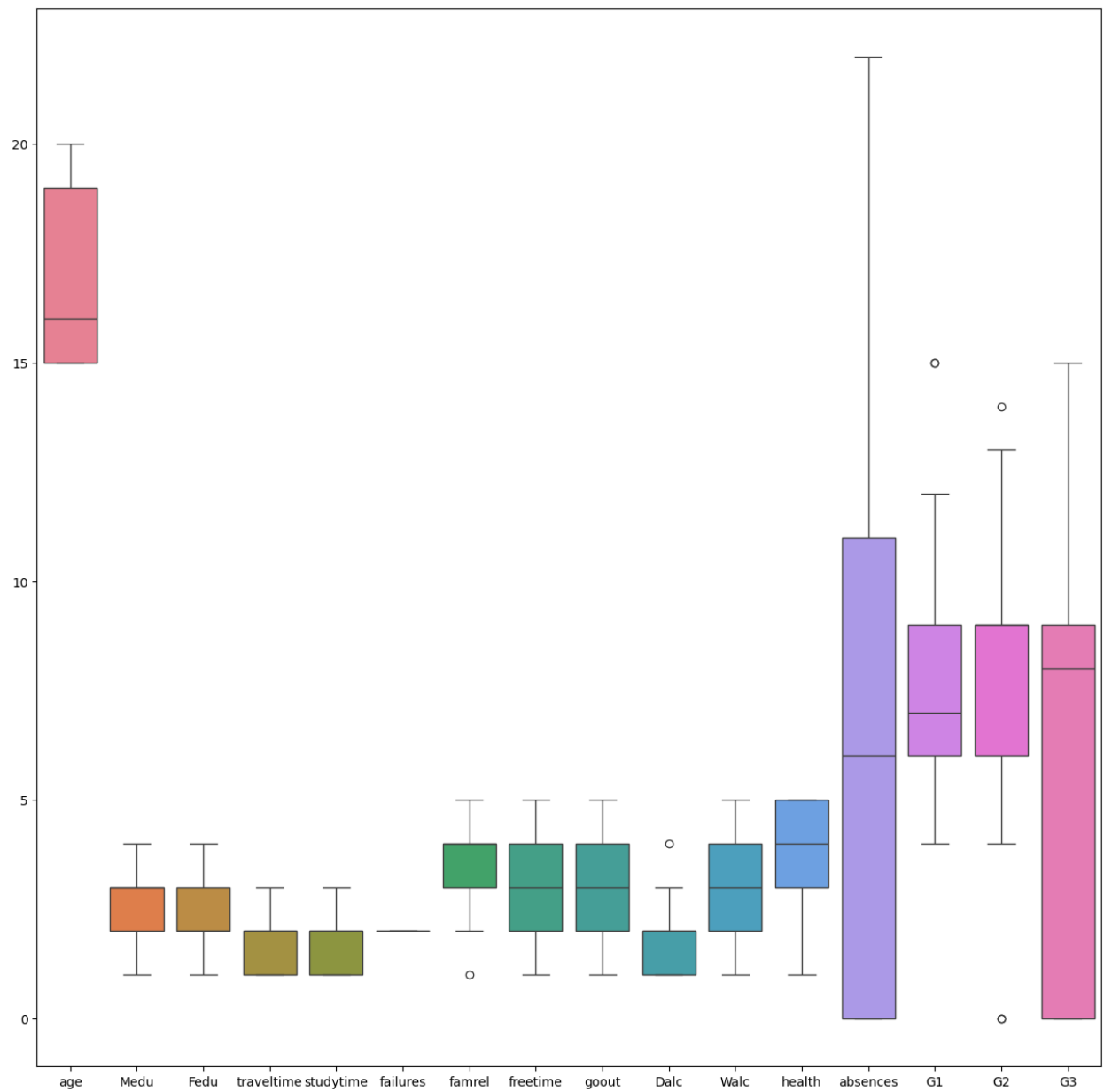

```
Fail_2['famrel'].value_counts()

famrel
4.0    7
3.0    6
5.0    2
1.0    1
2.0    1
Name: count, dtype: int64
```

```
Fail_2['famrel']=Fail_2['famrel'].apply(lambda x : x if
x>=round(Fail_2['famrel'].mean()) else round(Fail_2['famrel'].mean()))

Fail_2['Dalc'].value_counts()

Dalc
2.0    8
1.0    7
3.0    1
4.0    1
Name: count, dtype: int64

Fail_2['Dalc'] = Fail_2['Dalc'].apply(lambda x : x if
x<round(Fail_2['Dalc'].mean()) else round(Fail_2['Dalc'].mean()))

Fail_2['G1'].value_counts()

G1
7.0     4
6.0     3
9.0     3
15.0    2
8.0     1
12.0    1
4.0     1
10.0    1
5.0     1
Name: count, dtype: int64

Fail_2['G1']=Fail_2['G1'].apply(lambda x:x if x <
round(Fail_2['G1'].mean()) else round(Fail_2['G1'].mean()))

Fail_2['G2'].value_counts()

G2
9.0     6
6.0     3
0.0     2
4.0     1
10.0    1
5.0     1
13.0    1
7.0     1
14.0    1
Name: count, dtype: int64

Fail_2['G2']=Fail_2['G2'].apply(lambda x : x if x < 9 else
round(Fail_2['G2'].mean()))
Fail_2['G2'] = Fail_2['G2'].apply(lambda x : x if x>5 else 5)

Fail_2['G3'].value_counts()
```

```
G3
0.0     5
9.0     4
8.0     2
5.0     1
10.0    1
7.0     1
4.0     1
13.0    1
15.0    1
Name: count, dtype: int64
```

```python
#After Handling Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_2,ax=ax)
```

```
<Axes: >
```
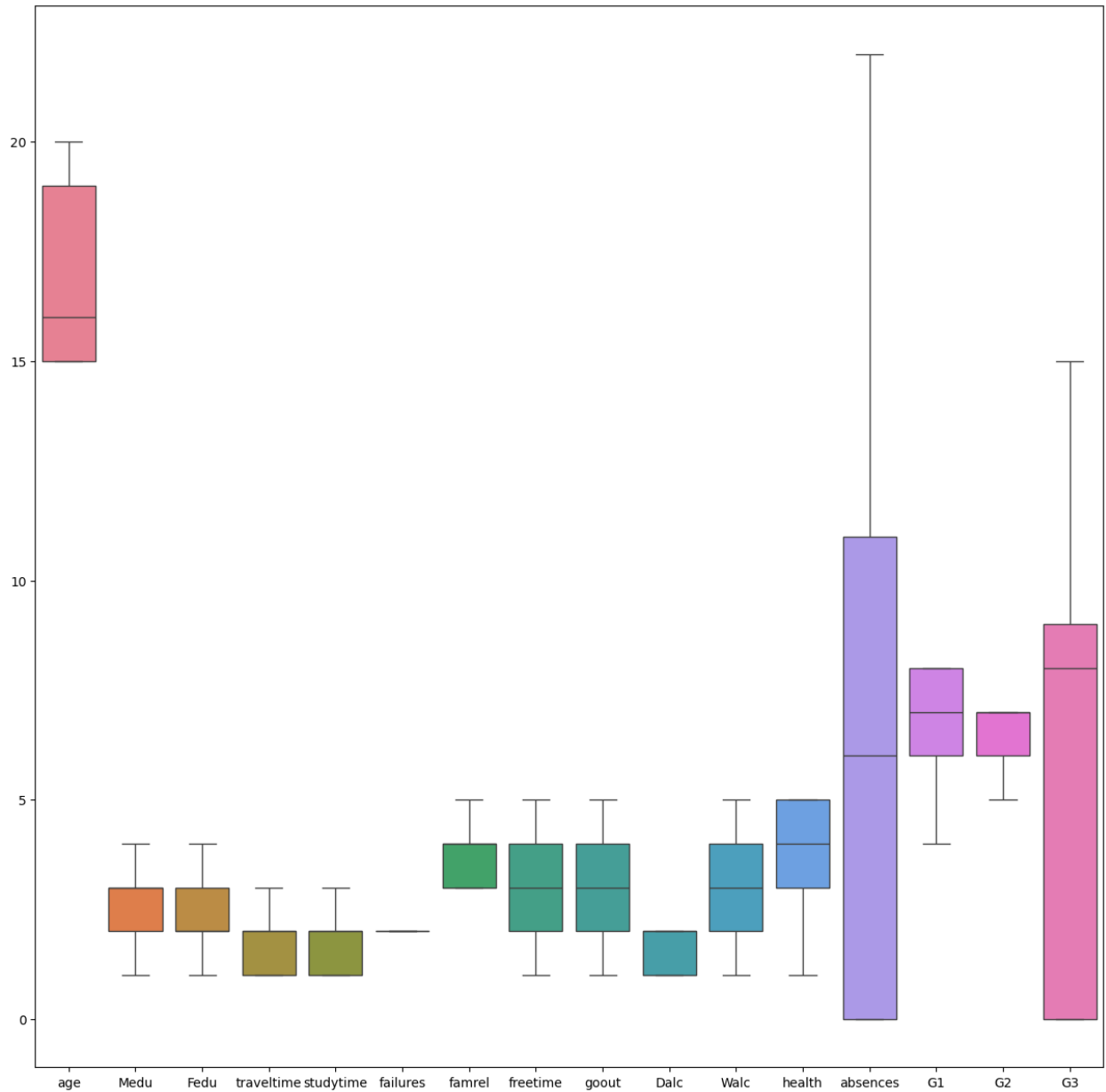
```
sns.displot(data = Fail_2,kind='kde')
```

```
C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\4063247573.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(data = Fail_2,kind='kde')
```

```
<seaborn.axisgrid.FacetGrid at 0x20294603aa0>
```
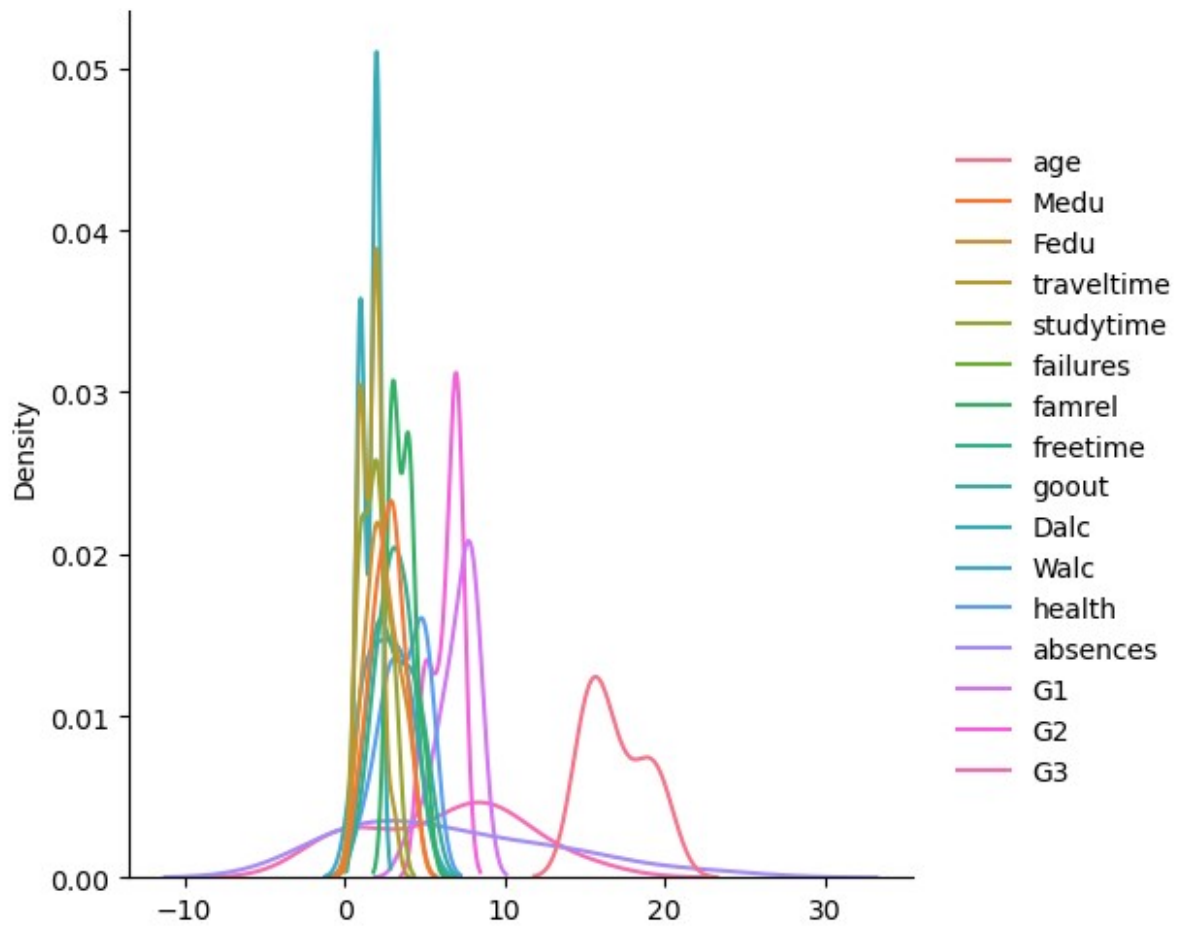
Handling Fail_3 Outliers

```
#Before Removing Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_3,ax=ax)

<Axes: >
```
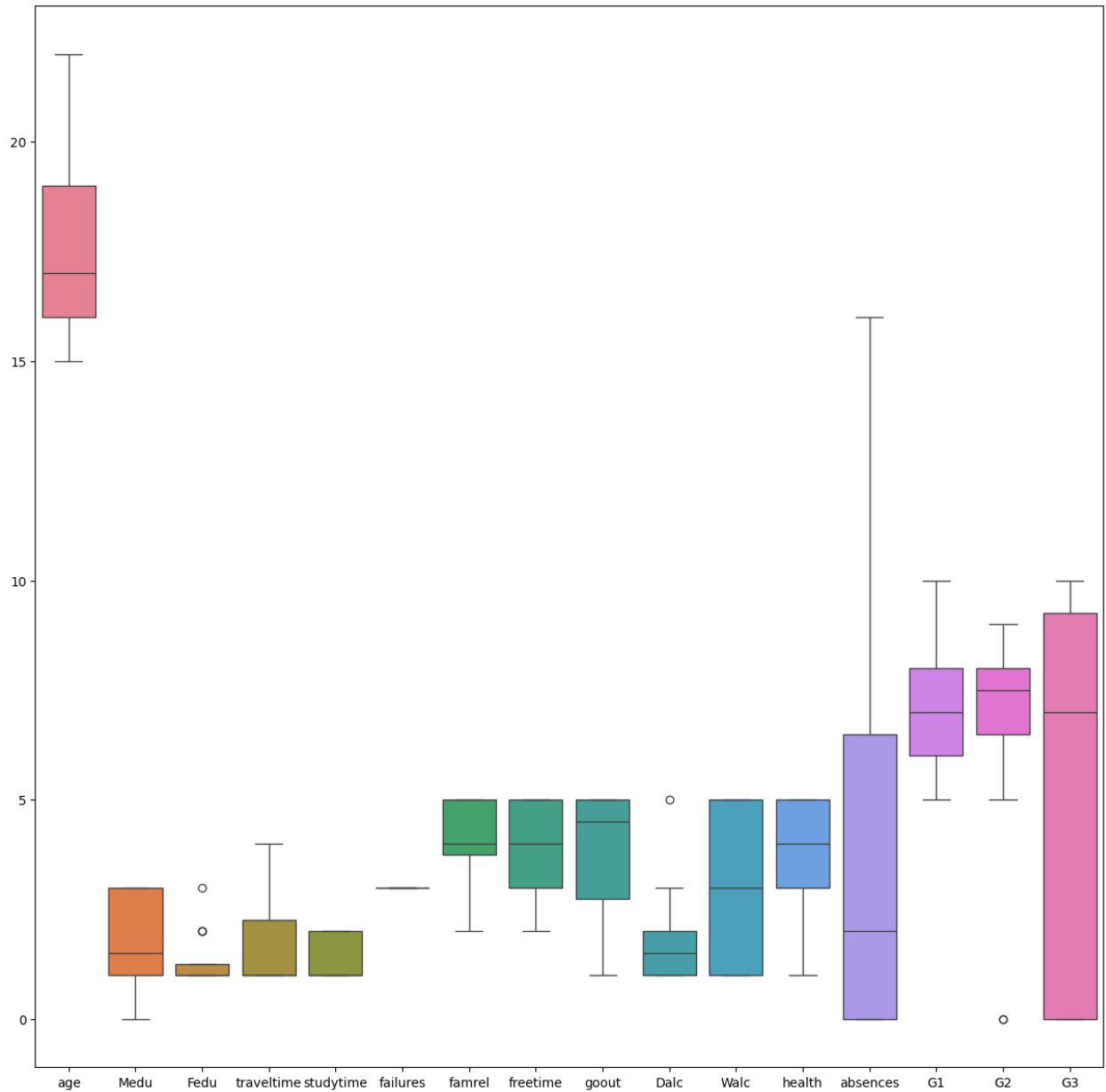
```
sns.displot(data = Fail_3,kind='kde')
```

C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\3551733757.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(data = Fail_3,kind='kde')

<seaborn.axisgrid.FacetGrid at 0x202950fedb0>

Legend:
- age
- Medu
- Fedu
- traveltime
- studytime
- failures
- famrel
- freetime
- goout
- Dalc
- Walc
- health
- absences
- G1
- G2
- G3
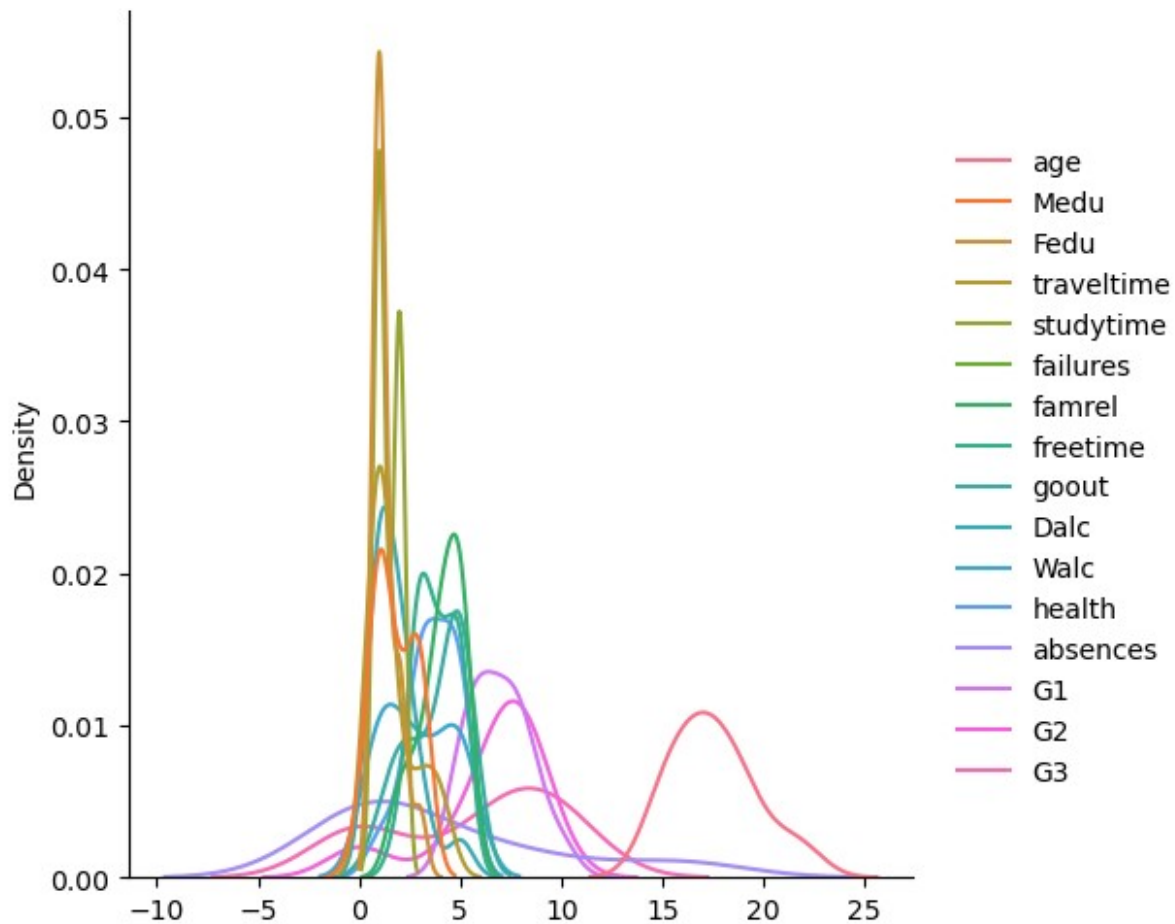
```
Fail_3['Fedu'].value_counts()

Fedu
1.0    12
2.0     3
3.0     1
Name: count, dtype: int64

Fail_3['Fedu']=Fail_3['Fedu'].apply(lambda x:x if
x==round(Fail_3['Fedu'].mean()) else round(Fail_3['Fedu'].mean()))

Fail_3['Dalc'].value_counts()

Dalc
1.0    8
2.0    5
3.0    2
5.0    1
Name: count, dtype: int64

Fail_3['Dalc'] =
Fail_3['Dalc'].replace(5,round(Fail_3['Dalc'].mean()))
```

```python
Fail_3['G2'].value_counts()

G2
8.0    7
7.0    4
5.0    2
0.0    2
9.0    1
Name: count, dtype: int64

Fail_3["G2"] = Fail_3['G2'].replace(0,5)

#After Removing Outliers
fig,ax = plt.subplots(figsize = (15,15))
sns.boxplot(data=Fail_3,ax=ax)

<Axes: >
```
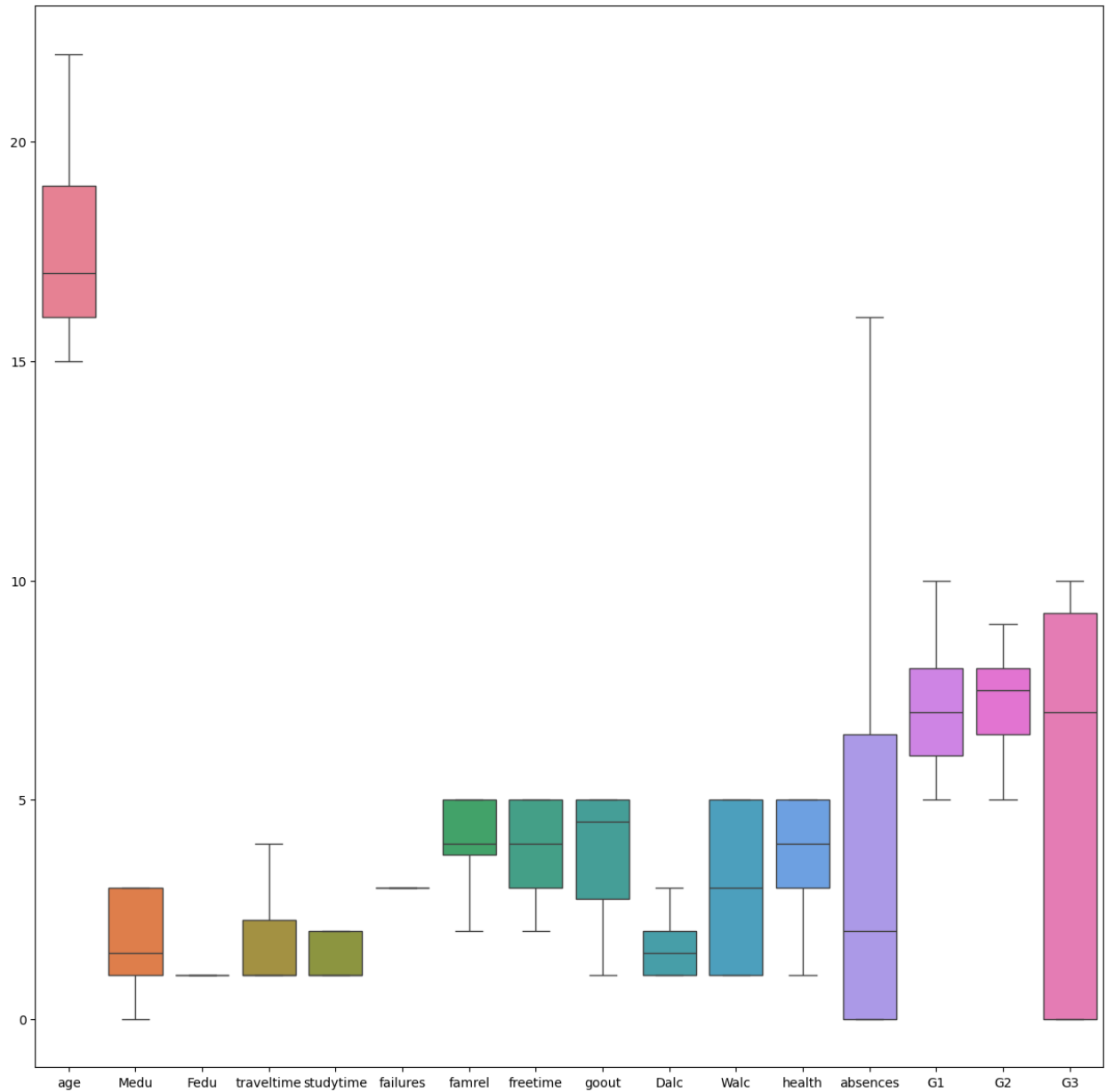
```
sns.displot(Fail_3,kind='kde')
```

C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\4245046575.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
`warn_singular=False` to disable this warning.
  sns.displot(Fail_3,kind='kde')

<seaborn.axisgrid.FacetGrid at 0x2029b0703e0>
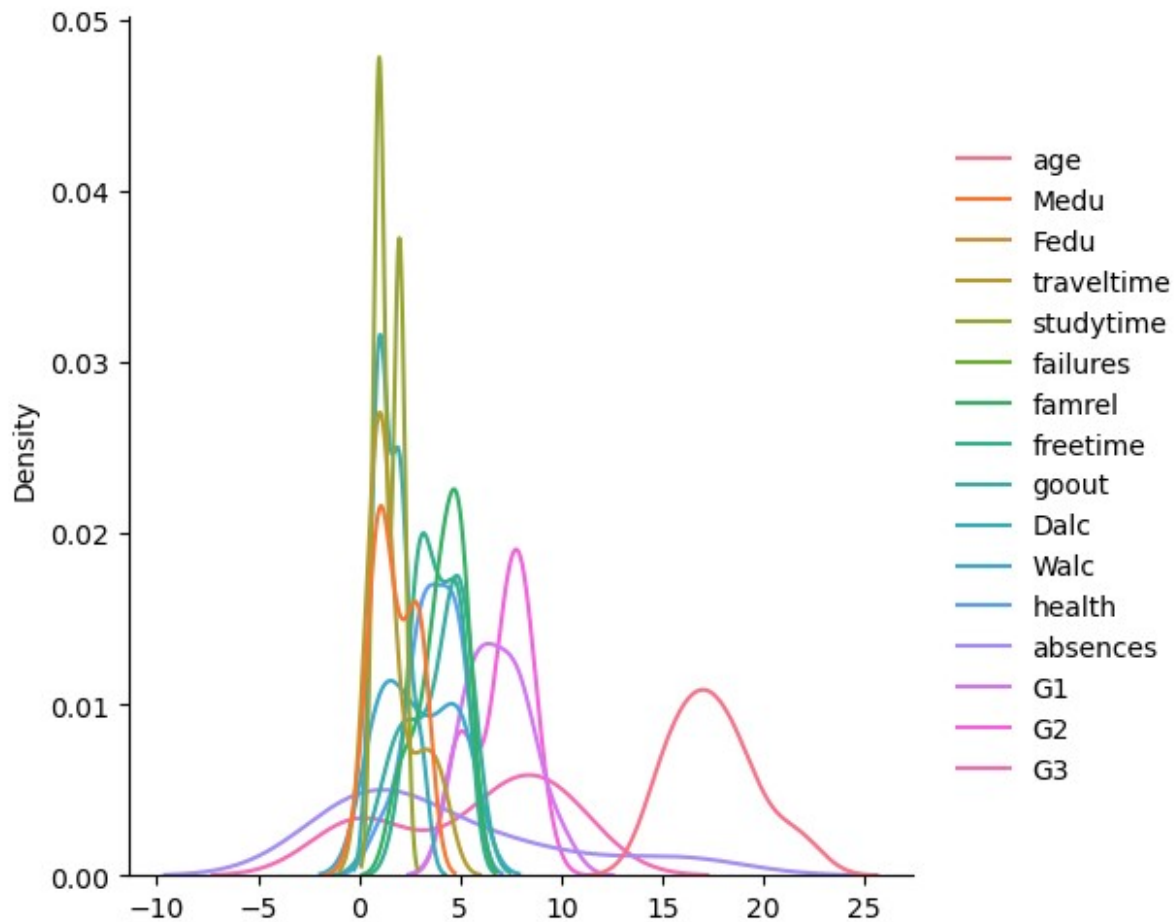
Fail_3 Linear Regression

```
X = Fail_3[::-1]
Y = Fail_3[::-1]

from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test =
train_test_split(X,Y,test_size=0.3,random_state=45)

from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_train_norm =
scaler.fit_transform(X_train.select_dtypes(include='float64').dropna()
)
X_test_norm =
scaler.transform(X_test.select_dtypes(include='float64').dropna())

from sklearn.linear_model import LinearRegression
Fail_3_regression = LinearRegression()
Fail_3_regression.fit(X_train_norm,Y_train.select_dtypes(include='floa
t64').dropna())
```

```
LinearRegression()

print(Fail_3_regression.coef_)

[[ 1.32012118e+00 -1.75141110e-01  0.00000000e+00  4.70056055e-03
   1.42161780e-01  0.00000000e+00 -1.45912494e-01  4.85974384e-01
   3.70317751e-02  2.77487660e-01  8.50913399e-02  3.50483867e-02
  -2.02122774e-01 -1.03275743e-01  4.89812544e-02 -1.20537934e-01]
 [-1.02602689e-01  7.41504639e-01  5.55111512e-17  7.99775602e-02
  -1.81844438e-01  4.16333634e-17 -2.75658020e-01  9.80561012e-02
   3.63527666e-02  4.18973995e-02  1.32542485e-01 -8.69353985e-02
   1.70420635e-04 -3.83130685e-02 -6.75697791e-02 -7.97765059e-02]
 [ 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00]
 [ 2.77736082e-03  8.06640754e-02  0.00000000e+00  8.53049691e-01
   1.04768095e-01 -2.77555756e-17  8.62925510e-02  4.79152393e-02
   1.33679664e-01 -4.36904877e-02 -9.54660072e-03 -4.14855553e-02
  -1.12512548e-01 -1.09210568e-02 -2.74191151e-02  2.28823209e-01]
 [ 4.23530929e-02 -9.24765575e-02  1.11022302e-16  5.28261146e-02
   2.97316299e-01  0.00000000e+00 -1.29024630e-01 -9.56776922e-02
   6.72167695e-02 -1.89542722e-02 -1.06578875e-02 -6.05448363e-02
   9.15159398e-02 -4.87312728e-02  3.40668101e-02 -1.18597256e-02]
 [ 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
   0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00]
 [-8.17122302e-02 -2.63508465e-01 -1.38777878e-17  8.17871806e-02
  -2.42529698e-01  0.00000000e+00  6.20608386e-01  6.30790696e-02
   7.19731389e-02  1.77996930e-02  1.39387507e-01 -1.19092094e-01
   1.77803687e-02 -4.81323504e-02 -7.57420473e-02 -5.13123690e-02]
 [ 2.47968414e-01  8.54057209e-02  3.60822483e-16  4.13784286e-02
  -1.63866961e-01  0.00000000e+00  5.74742989e-02  4.88673827e-01
  -4.19084213e-02 -1.76518405e-01 -6.83857541e-02 -2.18370480e-02
   1.88181247e-01  5.77503102e-02 -6.90977780e-03  2.06507434e-02]
 [ 3.04147046e-02  5.09654540e-02  3.33066907e-16  1.85819705e-01
   1.85303846e-01  2.77555756e-17  1.05556488e-01 -6.74570568e-02
   8.34477142e-01 -3.34055850e-02  1.39780525e-01  1.86168275e-01
   8.44647761e-02  2.56007578e-01 -1.91513284e-01 -4.43154605e-01]
 [ 1.28071228e-01  3.30083885e-02 -4.44089210e-16 -3.41281021e-02
  -2.93638323e-02  2.77555756e-17  1.46698552e-02 -1.59666904e-01
  -1.87723182e-02  5.96545173e-01  7.79830300e-02 -6.40924273e-02
   1.29516112e-02  9.59778008e-02 -1.60034086e-01  1.16373993e-01]
 [ 6.61061852e-02  1.75768561e-01  1.11022302e-16 -1.25522864e-02
  -2.77923702e-02  0.00000000e+00  1.93368622e-01 -1.04121253e-01
   1.32219140e-01  1.31264999e-01  9.53959425e-01  1.01735895e-01
   1.28940803e-01 -1.70517858e-01  3.91701372e-01 -2.72357349e-04]
 [ 2.15682380e-02 -9.13214165e-02 -1.66533454e-16 -4.32076742e-02
  -1.25060876e-01 -5.55111512e-17 -1.30868458e-01 -2.63364708e-02
```

```
   1.39490052e-01  -8.54565698e-02   8.05868370e-02   9.02832295e-01
  -2.89743331e-02  -1.51065903e-02  -9.43071997e-02   1.61634740e-01]
 [-5.98745791e-01   8.61744969e-04   1.99840144e-15  -5.64086271e-01
   9.09958717e-01  -1.94289029e-16   9.40531624e-02   1.09249782e+00
   3.04644811e-01   8.31271946e-02   4.91655454e-01  -1.39474251e-01
   3.97269342e+00   8.78513038e-02  -5.39945386e-01   7.20584370e-01]
 [-9.29855059e-02  -5.88834467e-02   3.33066907e-16  -1.66417481e-02
  -1.47272754e-01   5.55111512e-17  -7.73854690e-02   1.01903304e-01
   2.80647177e-01   1.87231893e-01  -1.97619623e-01  -2.21022633e-02
   2.67016390e-02   1.26318079e+00   3.12036101e-01   5.77296300e-02]
 [ 3.63352394e-02  -8.55618126e-02   3.33066907e-16  -3.44246000e-02
   8.48256732e-02   5.55111512e-17  -1.00332256e-01  -1.00456772e-02
  -1.72976837e-01  -2.57218806e-01   3.74021426e-01  -1.13683212e-01
  -1.35213711e-01   2.57090503e-01   7.43576452e-01   9.24136033e-02]
 [-2.71754643e-01  -3.07013587e-01  -6.10622664e-16   8.73113389e-01
  -8.97481681e-02  -1.11022302e-16  -2.06576404e-01   9.12443401e-02
  -1.21646445e+00   5.68461777e-01  -7.90379368e-04   5.92163051e-01
   5.48416918e-01   1.44555651e-01   2.80860725e-01   1.96127023e+00]]
```

```
Fail_3_reg_predict = Fail_3_regression.predict(X_test_norm)
Fail_3_reg_predict
```

```
array([[16.34888365,  2.62969451,  1.        ,  3.97048285,
1.06754475,
         3.        ,  4.62577012,  4.10139703,  4.33231047,
1.47755928,
         4.8513789 ,  5.12186491,  1.77263548,  8.72720481,
8.63481613,
        10.43971344],
       [18.32017991,  2.36947706,  1.        ,  0.5408854 ,
1.62282287,
         3.        ,  3.5859985 ,  2.97667191,  4.97568319,
2.26273243,
         3.52435975,  4.74409392, -0.50395152,  5.78009506,
6.25458539,
         0.96152317],
       [16.28022177,  1.02721775,  1.        ,  2.92837821,
1.51772964,
         3.        ,  5.2499411 ,  3.281773  ,  4.78601687,
1.78317075,
         4.83404111,  4.51432283,  2.86166225,  8.5950134 ,
8.60164929,
         8.06112104],
       [20.33120456,  2.79123751,  1.        ,  1.22194708,
1.53914014,
         3.        ,  5.02796751,  5.2007834 ,  3.73749311,
2.58510652,
         5.60646668,  1.54677129, 13.11600891,  6.33708084,
7.72344985,
         4.01752577],
```

```
        [15.06461527,  2.84201238,  1.        ,  0.79327123,
1.84148031,
         3.        ,  2.75424341,  2.9795186 ,  2.80914529,
0.88858875,
         1.05055073,  2.65560464, -0.52304032,  5.72583447,
6.94702979,
         2.22261724]])
```

```
Fail_3_residual = Y_test.select_dtypes(include='float64').dropna()-
Fail_3_reg_predict
Fail_3_residual
```

```
          age      Medu  Fedu  traveltime  studytime  failures
famrel  \
149 -1.348884 -0.629695   0.0    0.029517  -0.067545       0.0 -
0.625770
150 -0.320180 -1.369477   0.0    0.459115  -0.622823       0.0 -
1.585998
157  1.719778 -0.027218   0.0    0.071622  -0.517730       0.0 -
0.249941
247  1.668795  0.208762   0.0   -0.221947  -0.539140       0.0 -
0.027968
146 -0.064615  0.157988   0.0    0.206729   0.158520       0.0
0.245757

     freetime     goout      Dalc      Walc    health  absences
G1  \
149  0.898603  0.667690  0.522441  0.148621 -0.121865 -1.772635 -
0.727205
150  0.023328  0.024317 -0.262732  1.475640 -0.744094  0.503952
0.219905
157 -1.281773  0.213983 -0.783171  0.165959 -0.514323  3.138338
0.404987
247 -1.200783  1.262507 -0.585107 -0.606467 -0.546771  2.883991 -
0.337081
146  0.020481 -0.809145  0.111411 -0.050551  0.344395  0.523040
0.274166

          G2        G3
149  0.365184 -0.439713
150 -1.254585 -0.961523
157 -0.601649  1.938879
247  0.276550  3.982474
146  0.052970 -2.222617
```

```
sns.displot(Fail_3_residual,kind='kde')
```

```
C:\Users\apurv\AppData\Local\Temp\ipykernel_3484\1181344936.py:1:
UserWarning: Dataset has 0 variance; skipping density estimate. Pass
```

```
`warn_singular=False` to disable this warning.
  sns.displot(Fail_3_residual,kind='kde')

<seaborn.axisgrid.FacetGrid at 0x20294fb57c0>
```
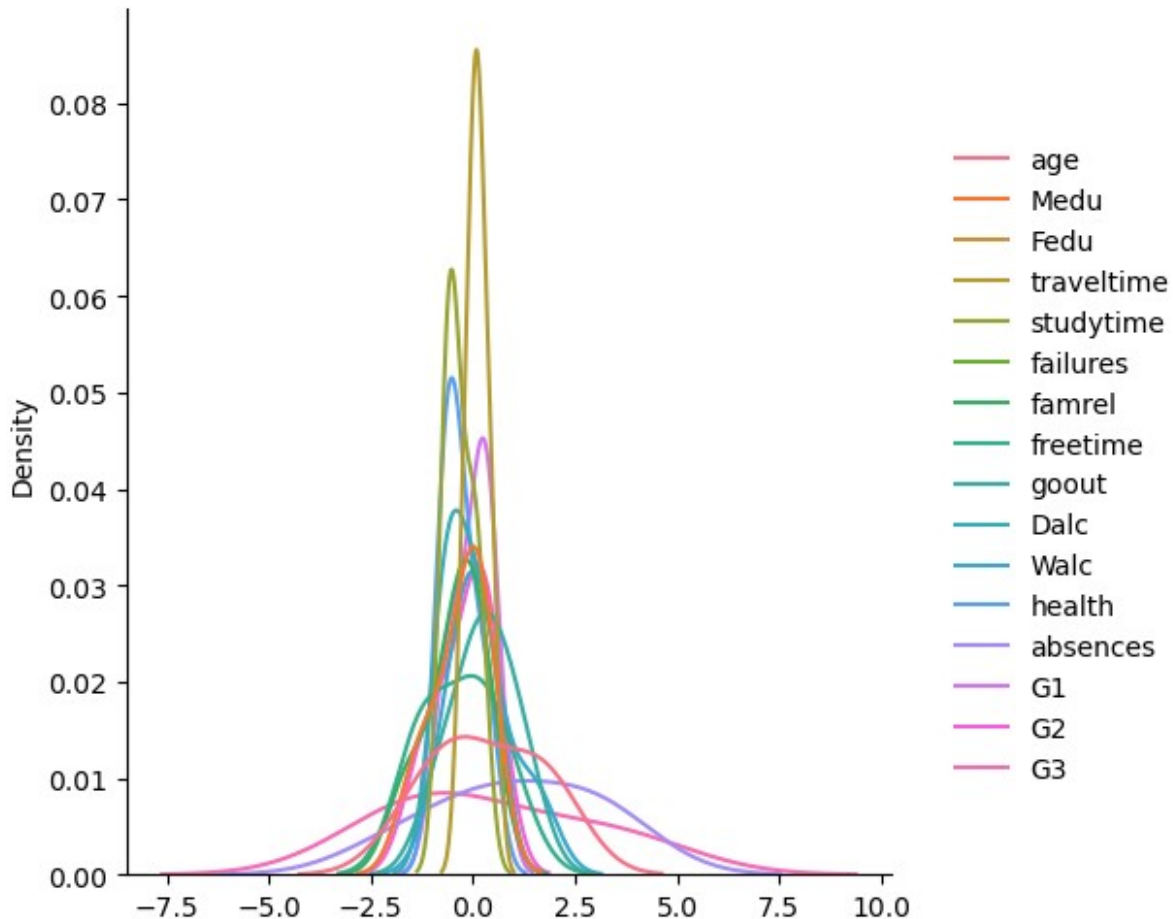


```
from sklearn.metrics import
mean_squared_error,mean_absolute_error,r2_score
print("Mean Squared Error
",mean_squared_error(Y_test.select_dtypes(include="float64").dropna(),
Fail_3_reg_predict))
print("Mean Absolute Error
",mean_absolute_error(Y_test.select_dtypes(include="float64").dropna()
,Fail_3_reg_predict))
print("R Square Score
",r2_score(Y_test.select_dtypes(include="float64").dropna(),Fail_3_reg
_predict))

Mean Squared Error  0.9588775361268577
Mean Absolute Error  0.6176639393540492
R Square Score  0.638615805960614
```

Above We have seen that there are students who failed in some subjects and some don't now
we are trying to figure out the reasons due to which they failed and they don't

```
Fail_0

     school sex    age address famsize Pstatus  Medu  Fedu       Mjob
Fjob  \
0        GP   F  18.0       U     GT3       A   4.0   4.0    at_home
teacher
1        GP   F  17.0       U     GT3       T   1.0   1.0    at_home
other
3        GP   F  15.0       U     GT3       T   4.0   2.0     health
services
4        GP   F  16.0       U     GT3       T   3.0   3.0      other
other
5        GP   M  16.0       U     LE3       T   4.0   3.0   services
other
..      ...  ..   ...     ...     ...     ...   ...   ...        ...
...
386      MS   F  18.0       R     GT3       T   4.0   4.0    teacher
at_home
388      MS   F  18.0       U     LE3       T   3.0   1.0    teacher
services
391      MS   M  17.0       U     LE3       T   3.0   1.0   services
services
393      MS   M  18.0       R     LE3       T   3.0   2.0   services
other
394      MS   M  17.0       U     LE3       T   1.0   1.0      other
at_home

      ... famrel freetime  goout  Dalc  Walc health absences     G1
G2    G3
0     ...    4.0      3.0    4.0   1.0   1.0    3.0      5.0    5.0
11.0  11.0
1     ...    5.0      3.0    3.0   1.0   1.0    3.0      5.0    5.0
11.0  11.0
3     ...    3.0      2.0    2.0   1.0   1.0    5.0      5.0   15.0
14.0  15.0
4     ...    4.0      3.0    2.0   1.0   2.0    5.0      5.0    6.0
10.0  10.0
5     ...    5.0      4.0    2.0   1.0   2.0    5.0      5.0   15.0
15.0  15.0
..    ...    ...      ...    ...   ...   ...    ...      ...    ..  ..
.    ...
386   ...    4.0      4.0    3.0   2.0   2.0    5.0      NaN    6.0
11.0  11.0
388   ...    4.0      3.0    4.0   1.0   1.0    1.0      NaN    7.0
9.0   8.0
391   ...    4.0      4.0    5.0   3.0   4.0    2.0      NaN   14.0
16.0  16.0
```

```
393  ...     4.0      4.0    1.0   3.0   4.0   5.0      NaN   11.0
12.0  10.0
394  ...     3.0      2.0    3.0   3.0   3.0   5.0      NaN    8.0
9.0    9.0

[312 rows x 33 columns]

Fail_0_No_Fees = Fail_0.where(data['paid']=='no').dropna()
Fail_0_No_Fees

    school sex   age address famsize Pstatus  Medu  Fedu       Mjob
Fjob  \
0       GP   F  18.0       U     GT3       A   4.0   4.0    at_home
teacher
1       GP   F  17.0       U     GT3       T   1.0   1.0    at_home
other
6       GP   M  16.0       U     LE3       T   2.0   2.0      other
other
7       GP   F  17.0       U     GT3       A   4.0   4.0      other
teacher
11      GP   F  15.0       U     GT3       T   2.0   1.0   services
other
14      GP   M  15.0       U     GT3       A   2.0   2.0      other
other
15      GP   F  16.0       U     GT3       T   4.0   4.0     health
other
17      GP   F  16.0       U     GT3       T   3.0   3.0      other
other
20      GP   M  15.0       U     GT3       T   4.0   3.0    teacher
other
22      GP   M  16.0       U     LE3       T   4.0   2.0    teacher
other
23      GP   M  16.0       U     LE3       T   2.0   2.0      other
other

    ... famrel freetime  goout  Dalc  Walc health absences    G1    G2
G3
0   ...    4.0      3.0    4.0   1.0   1.0    3.0      5.0   5.0  11.0
11.0
1   ...    5.0      3.0    3.0   1.0   1.0    3.0      5.0   5.0  11.0
11.0
6   ...    4.0      4.0    4.0   1.0   1.0    3.0      5.0  12.0  12.0
11.0
7   ...    4.0      3.0    4.0   1.0   1.0    1.0      5.0   6.0  11.0
11.0
11  ...    5.0      2.0    2.0   1.0   1.0    4.0      5.0  10.0  12.0
12.0
14  ...    4.0      5.0    2.0   1.0   1.0    3.0      5.0  14.0  16.0
16.0
15  ...    4.0      4.0    4.0   1.0   2.0    2.0      2.0  14.0  14.0
```

```
14.0
17   ...      5.0        3.0     2.0   1.0   1.0     4.0       2.0   8.0  10.0
10.0
20   ...      4.0        4.0     1.0   1.0   1.0     1.0       1.0  13.0  14.0
15.0
22   ...      4.0        5.0     1.0   1.0   3.0     5.0       1.0  15.0  15.0
16.0
23   ...      5.0        4.0     4.0   2.0   4.0     5.0       1.0  13.0  13.0
12.0

[11 rows x 33 columns]

Fail_0_No_Fees['school']

0      GP
1      GP
6      GP
7      GP
11     GP
14     GP
15     GP
17     GP
20     GP
22     GP
23     GP
Name: school, dtype: object
```

So Most of the students who haven't paid the fees and also not failed in any subject are from the same school

```
sns.displot(Fail_0_No_Fees['sex'])

<seaborn.axisgrid.FacetGrid at 0x2029c8f0d40>
```
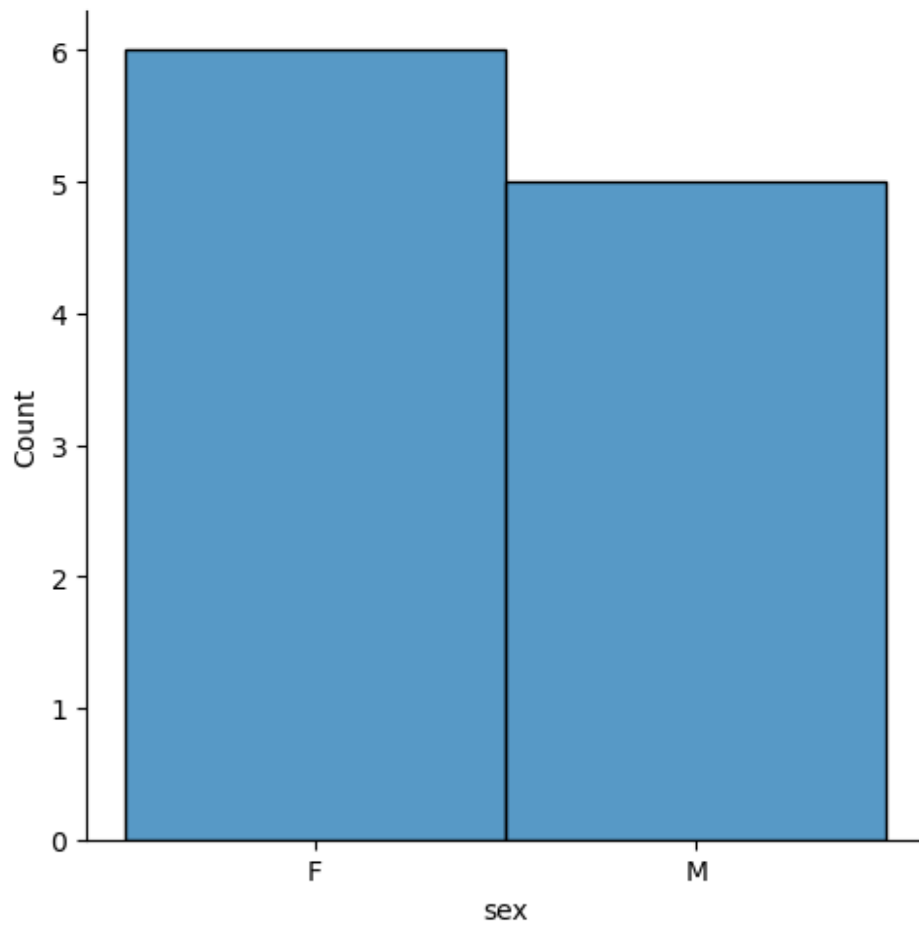
This is the sex ratio of students for the particular category

```
sns.displot(Fail_0_No_Fees['age'])
```

```
<seaborn.axisgrid.FacetGrid at 0x2029c1fa600>
```
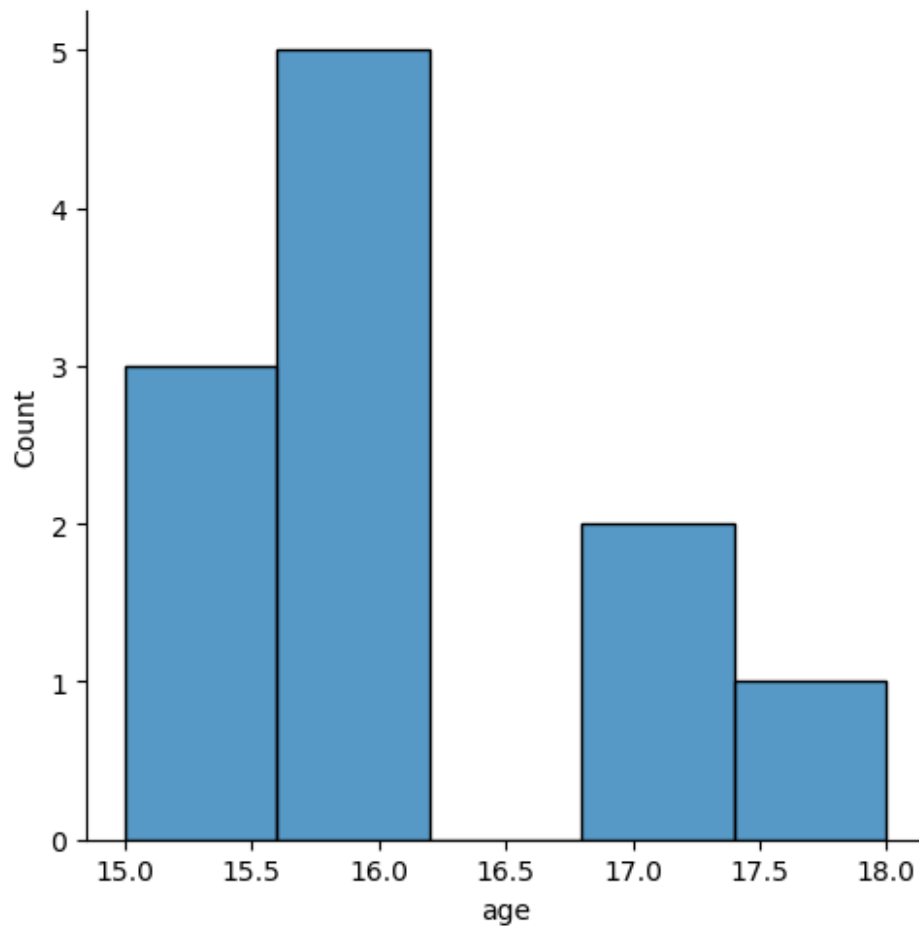
Age Group Distribution among the students

```
sns.displot(Fail_0_No_Fees['age'].where(Fail_0_No_Fees['sex']=='F'))
<seaborn.axisgrid.FacetGrid at 0x2029b8b9520>
```
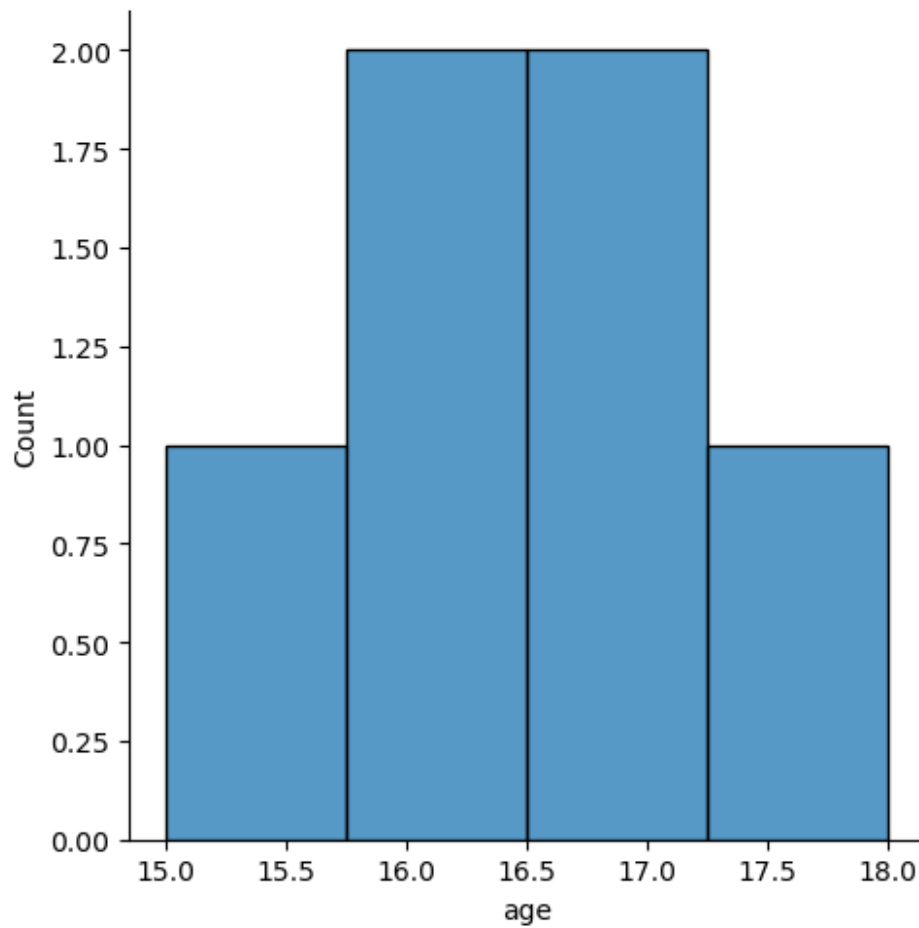
Age Group of Female students in the particular category

```
sns.displot(Fail_0_No_Fees['age'].where(Fail_0_No_Fees['sex']=='M'))
<seaborn.axisgrid.FacetGrid at 0x2029eaa66f0>
```
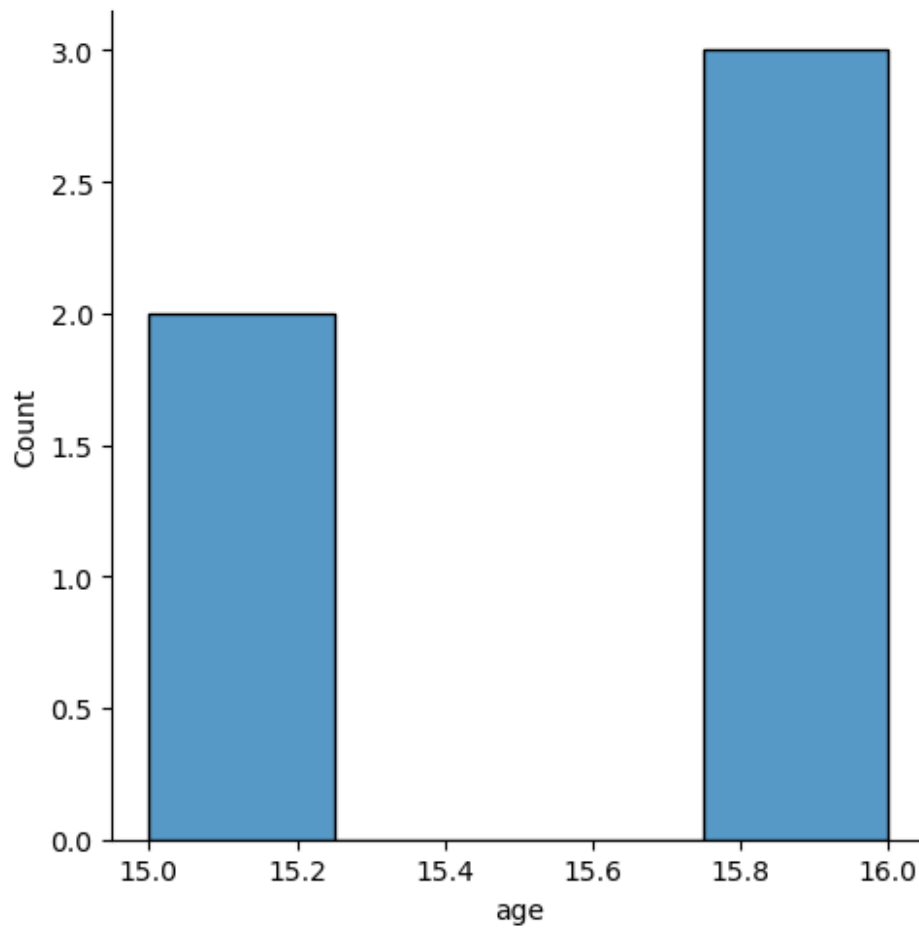
Distribution Among Age Group of Male students in the given category

```
sns.displot(Fail_0_No_Fees['address'])
<seaborn.axisgrid.FacetGrid at 0x2029ea85040>
```
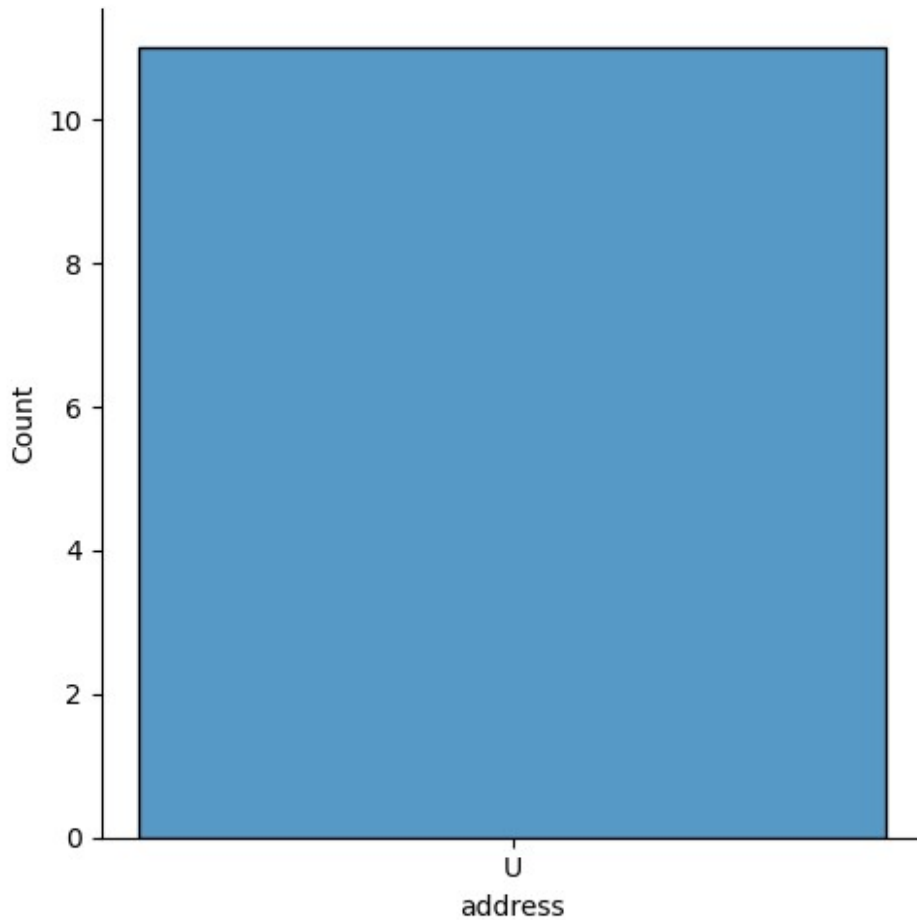
From this we can see that all of the students of this category belongs to Urban

```
Fail_0_No_Fees['famsize'].value_counts()

famsize
GT3    8
LE3    3
Name: count, dtype: int64
```

Family Size of distribution of the particular category

```
sns.displot(Fail_0_No_Fees['famsize'].where(Fail_0_No_Fees['sex']=='F'
))

<seaborn.axisgrid.FacetGrid at 0x202945beba0>
```
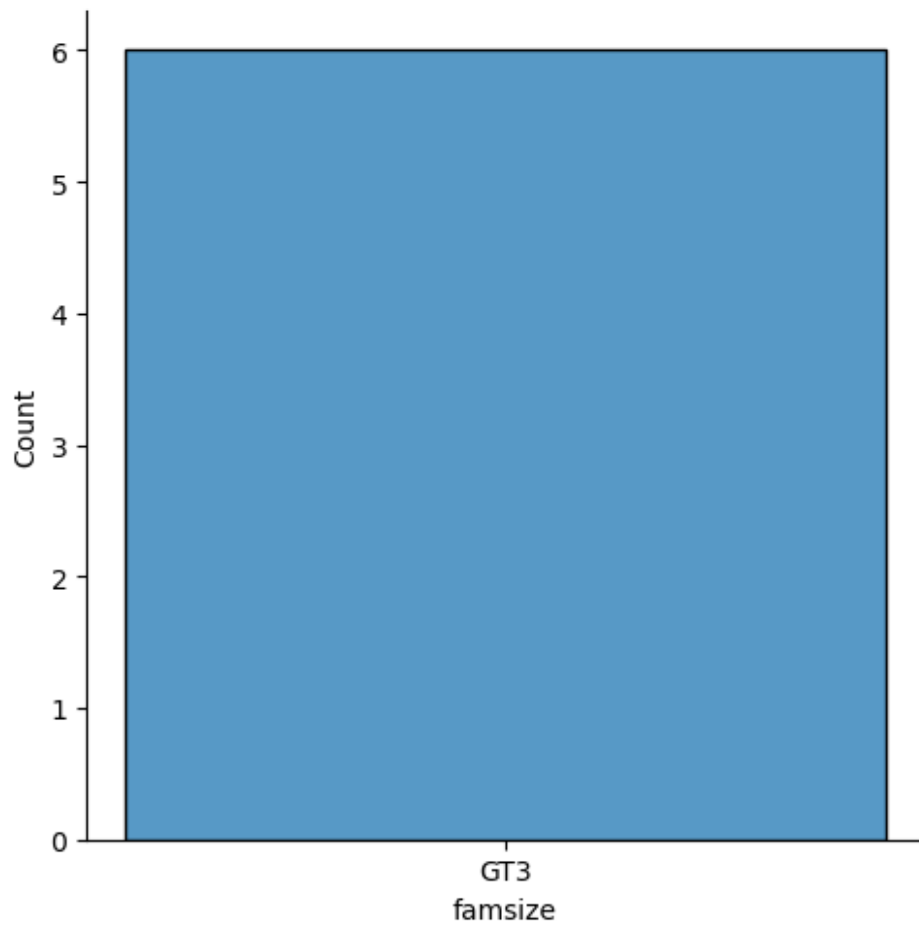
Family Size of Female Students of the following group

```
sns.displot(Fail_0_No_Fees['famsize'].where(Fail_0_No_Fees['sex']=='M'
))
```

```
<seaborn.axisgrid.FacetGrid at 0x2029eaa5c40>
```
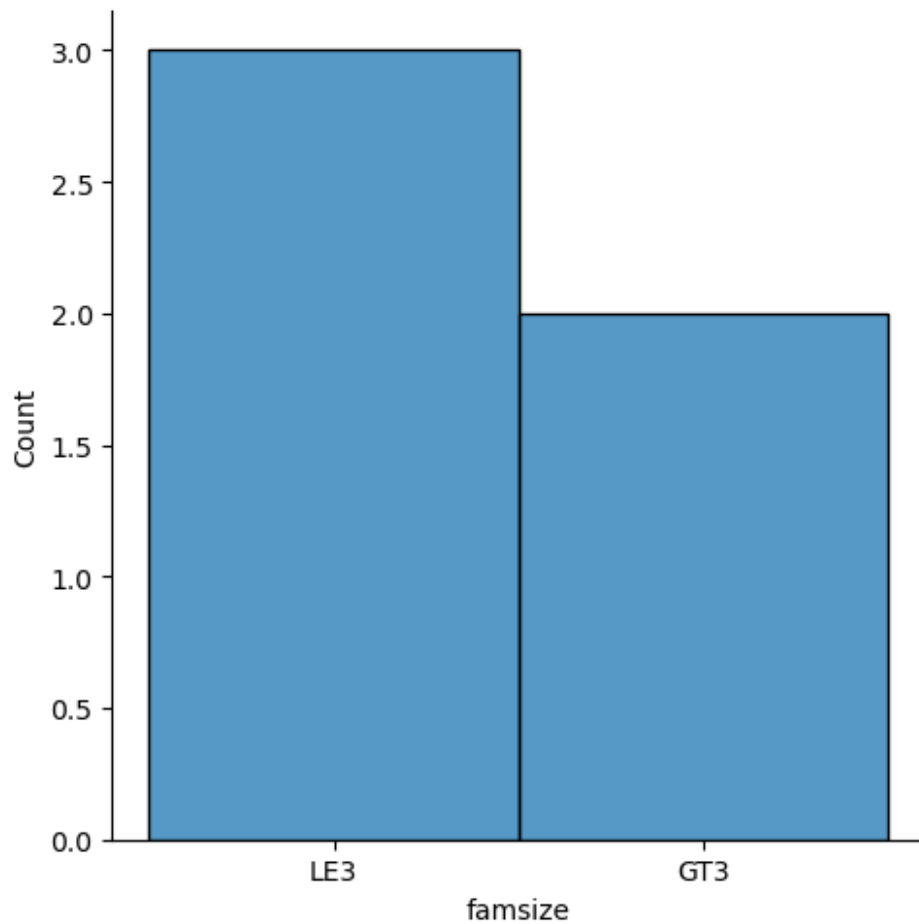
Family Size of Male Students of the Particular Group

```
Fail_0_No_Fees['Pstatus'].value_counts()

Pstatus
T    8
A    3
Name: count, dtype: int64
```

Count of Parents who are together and apart

```
sns.displot(Fail_0_No_Fees['Pstatus'].where(Fail_0_No_Fees['sex']=='F'
))

<seaborn.axisgrid.FacetGrid at 0x2029f397560>
```
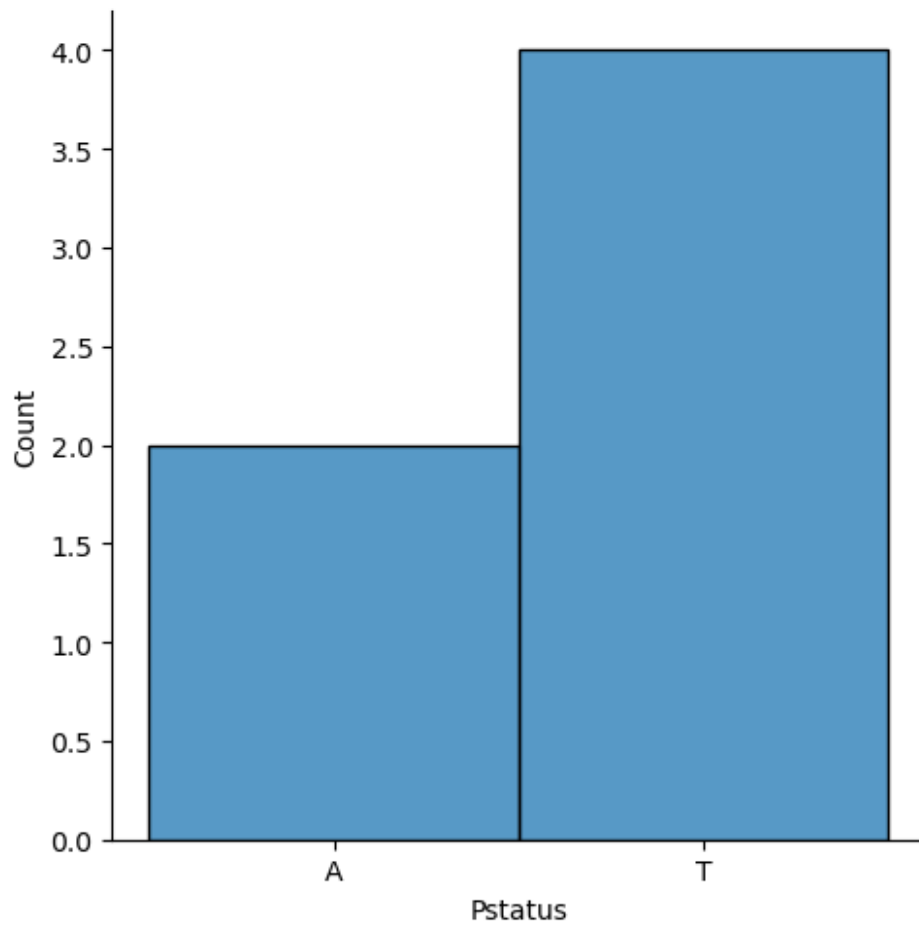
Parents Status for the female students of the following group

```
sns.displot(Fail_0_No_Fees['Pstatus'].where(Fail_0_No_Fees['sex']=='M'
))

<seaborn.axisgrid.FacetGrid at 0x2029e7c9ca0>
```
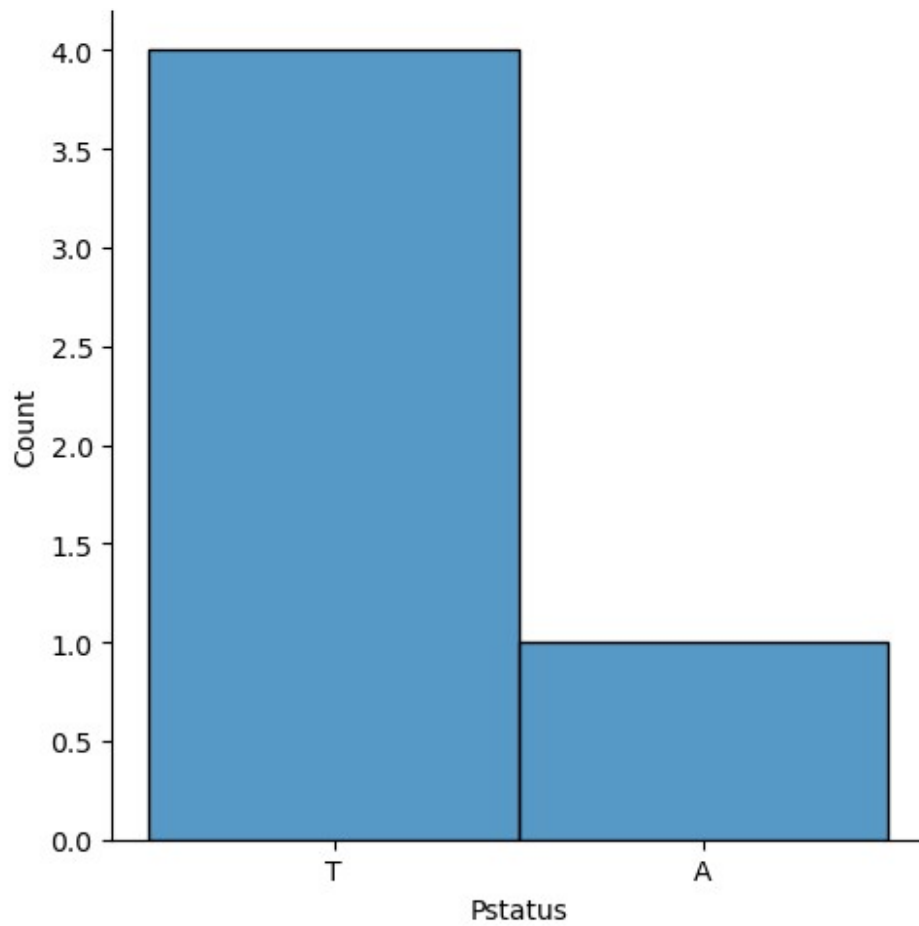
Parents Status for the Male students of the following group

```
sns.displot(Fail_0_No_Fees['traveltime'])
<seaborn.axisgrid.FacetGrid at 0x2029c8bf3e0>
```
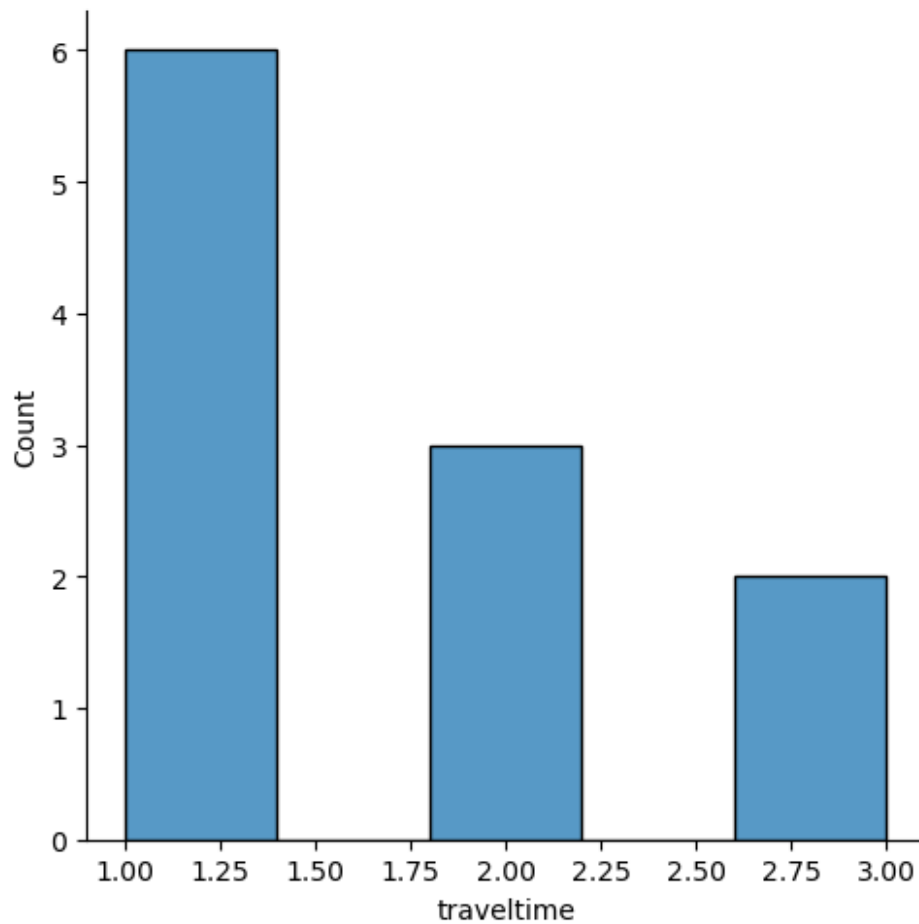
Distribution of Travel Time of Students

```
sns.displot(Fail_0_No_Fees['traveltime'].where((Fail_0_No_Fees['sex']=
='F') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['Pstatus']=='T')))
```

```
<seaborn.axisgrid.FacetGrid at 0x2029f49e3c0>
```
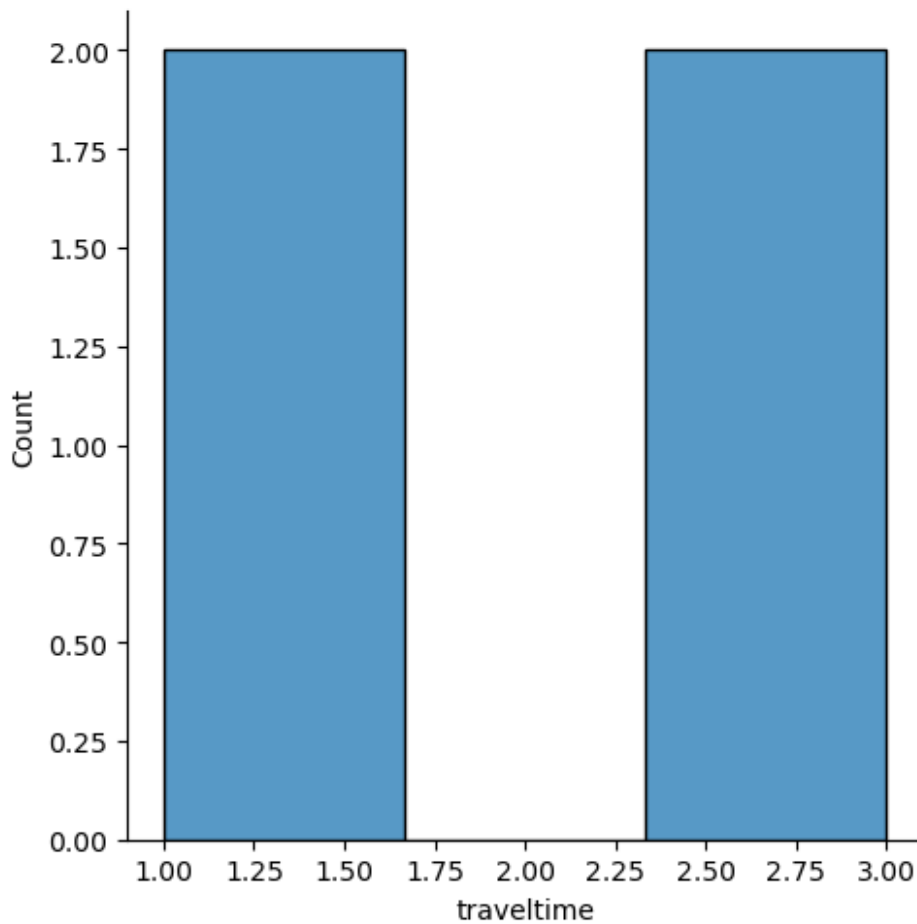
Distribution of Travel Time of Female With Family Size Greater Than 3 and Parents Status Together

```
plt.xlabel('Traveltime')
plt.ylabel('Count')

plt.hist(Fail_0_No_Fees['traveltime'].where((Fail_0_No_Fees['sex']=='F
') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['Pstatus']=='A')))

(array([0., 0., 0., 0., 0., 2., 0., 0., 0., 0.]),
 array([1.5, 1.6, 1.7, 1.8, 1.9, 2. , 2.1, 2.2, 2.3, 2.4, 2.5]),
 <BarContainer object of 10 artists>)
```

Distribution of Travel Time of Female With Family Size Greater Than 3 and Parents Status Apart

```python
plt.xlabel('Traveltime')
plt.ylabel('Count')
plt.hist(Fail_0_No_Fees['traveltime'].where((Fail_0_No_Fees['sex']=='M') & (Fail_0_No_Fees['famsize']=='GT3') & (Fail_0_No_Fees['Pstatus']=='T')))

(array([0., 0., 0., 0., 0., 1., 0., 0., 0., 0.]),
 array([0.5, 0.6, 0.7, 0.8, 0.9, 1. , 1.1, 1.2, 1.3, 1.4, 1.5]),
 <BarContainer object of 10 artists>)
```
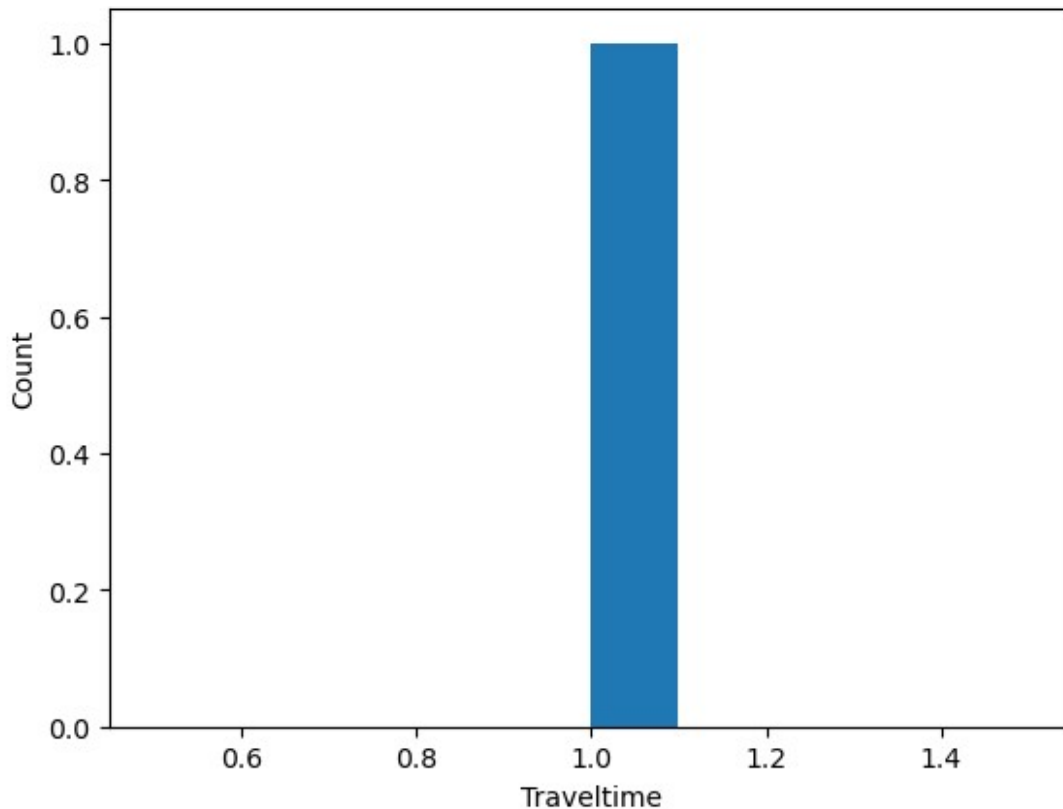
Distribution of Travel Time of Male With Family Size Greater Than 3 and Parents Status Together

```
plt.xlabel('Traveltime')
plt.ylabel('Count')
plt.hist(Fail_0_No_Fees['traveltime'].where((Fail_0_No_Fees['sex']=='M
') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['Pstatus']=='A')))

(array([0., 0., 0., 0., 0., 1., 0., 0., 0., 0.]),
 array([0.5, 0.6, 0.7, 0.8, 0.9, 1. , 1.1, 1.2, 1.3, 1.4, 1.5]),
 <BarContainer object of 10 artists>)
```
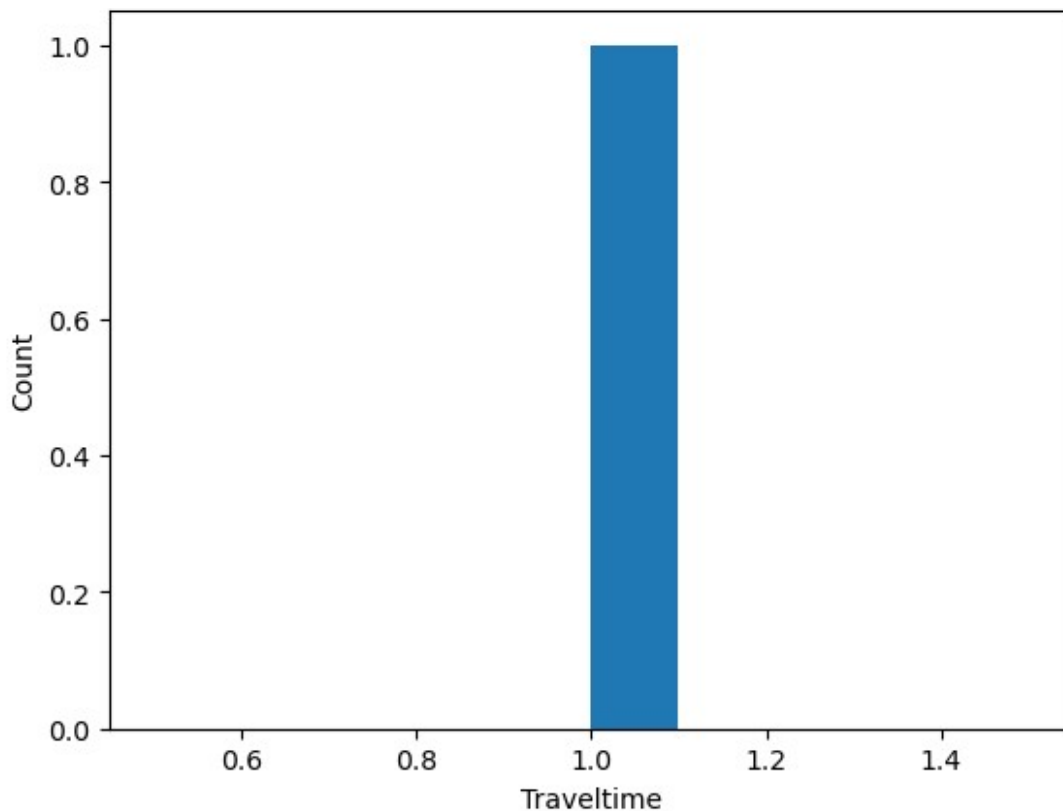
Distribution of Travel Time of Male With Family Size Greater Than 3 and Parents Status Apart

```
plt.xlabel('Traveltime')
plt.ylabel('Count')
plt.hist(Fail_0_No_Fees['traveltime'].where((Fail_0_No_Fees['sex']=='M
') & (Fail_0_No_Fees['famsize']=='LE3') &
(Fail_0_No_Fees['Pstatus']=='T')))

(array([2., 0., 0., 0., 0., 0., 0., 0., 0., 1.]),
 array([1. , 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2. ]),
 <BarContainer object of 10 artists>)
```
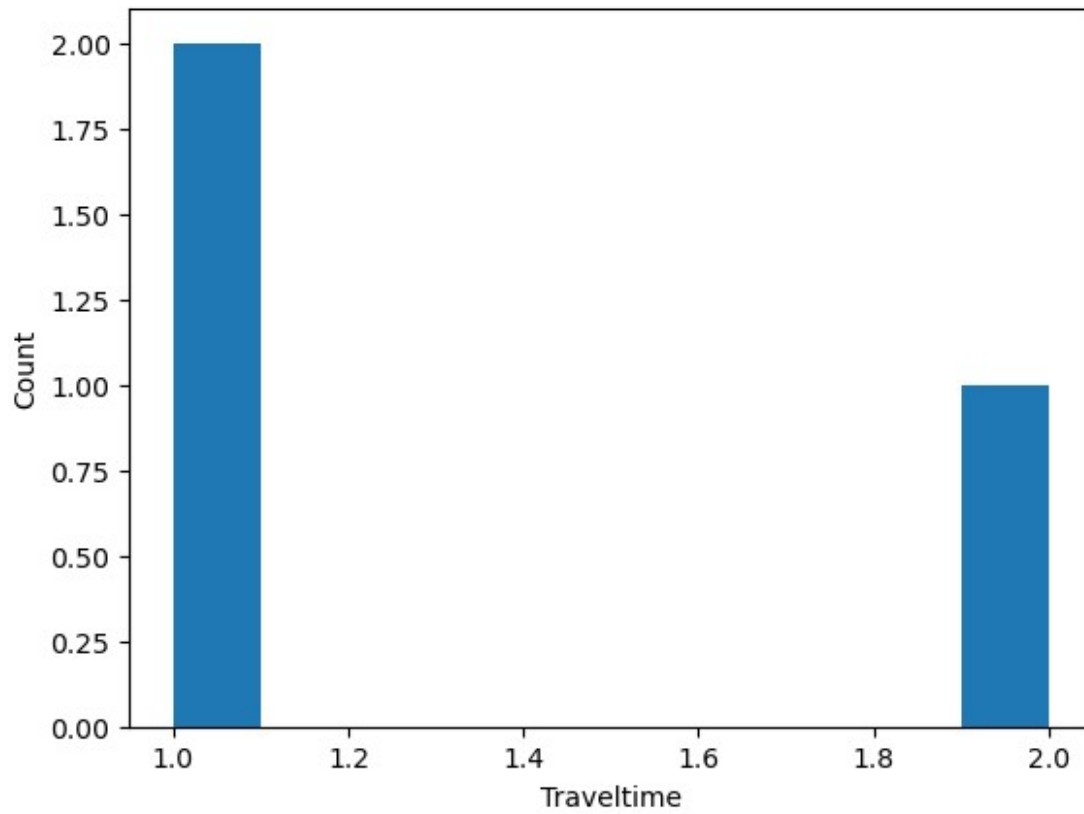
Distribution of Travel Time of Male With Family Size Less Than 3 and Parents Status Together

```
sns.displot(Fail_0_No_Fees['studytime'])
```

```
<seaborn.axisgrid.FacetGrid at 0x2029e66aba0>
```
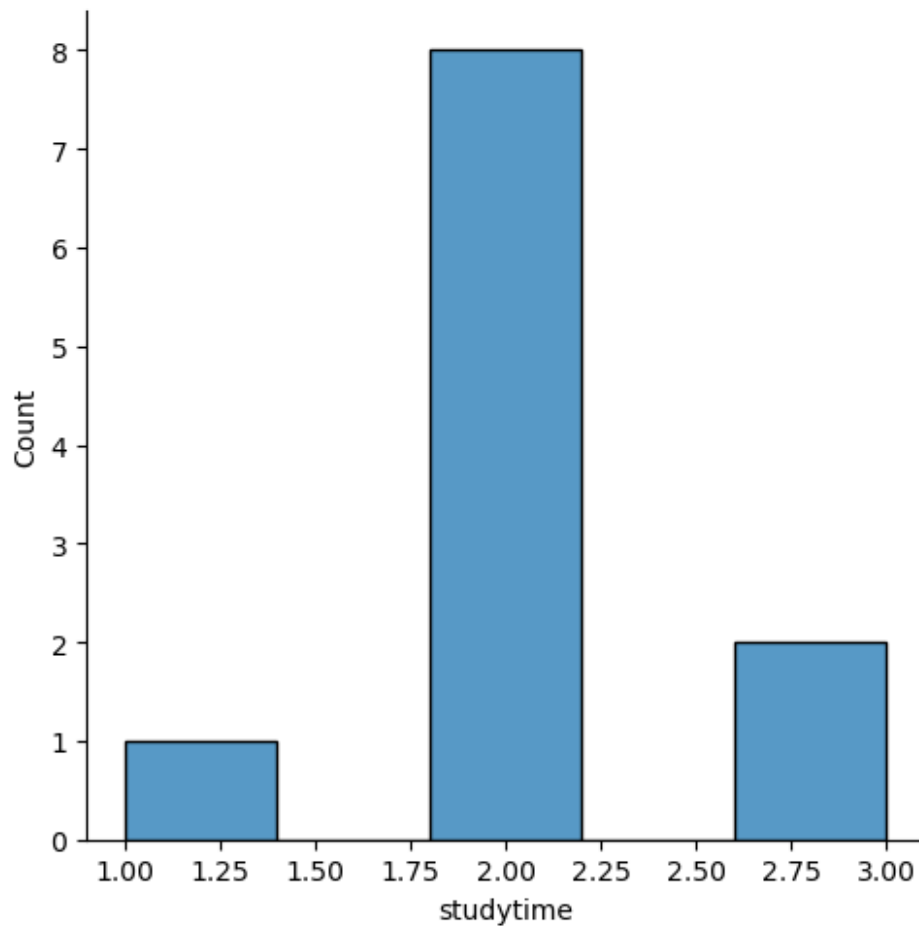
Distribution of Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where(Fail_0_No_Fees['sex']=='F'))
```

```
<seaborn.axisgrid.FacetGrid at 0x2029f5bd490>
```
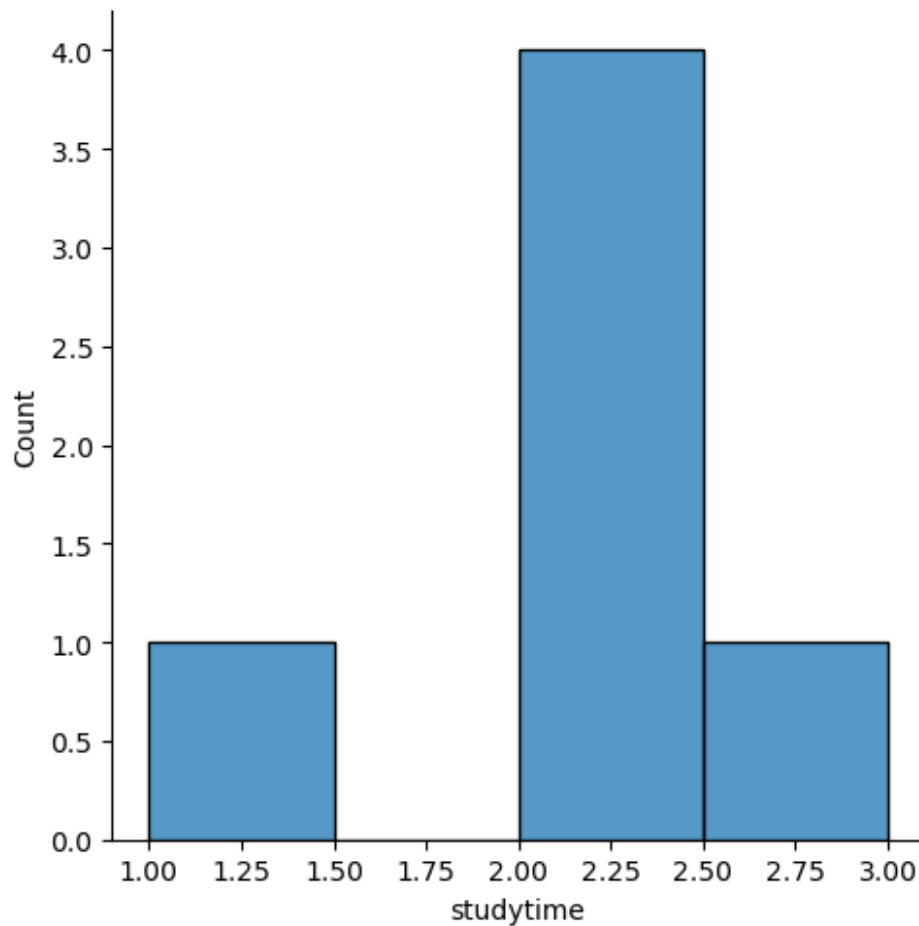
Distribution of Female student Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'F') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['schoolsup']=='yes') &
          (Fail_0_No_Fees['famsup']=='yes')))

<seaborn.axisgrid.FacetGrid at 0x2029e446ed0>
```
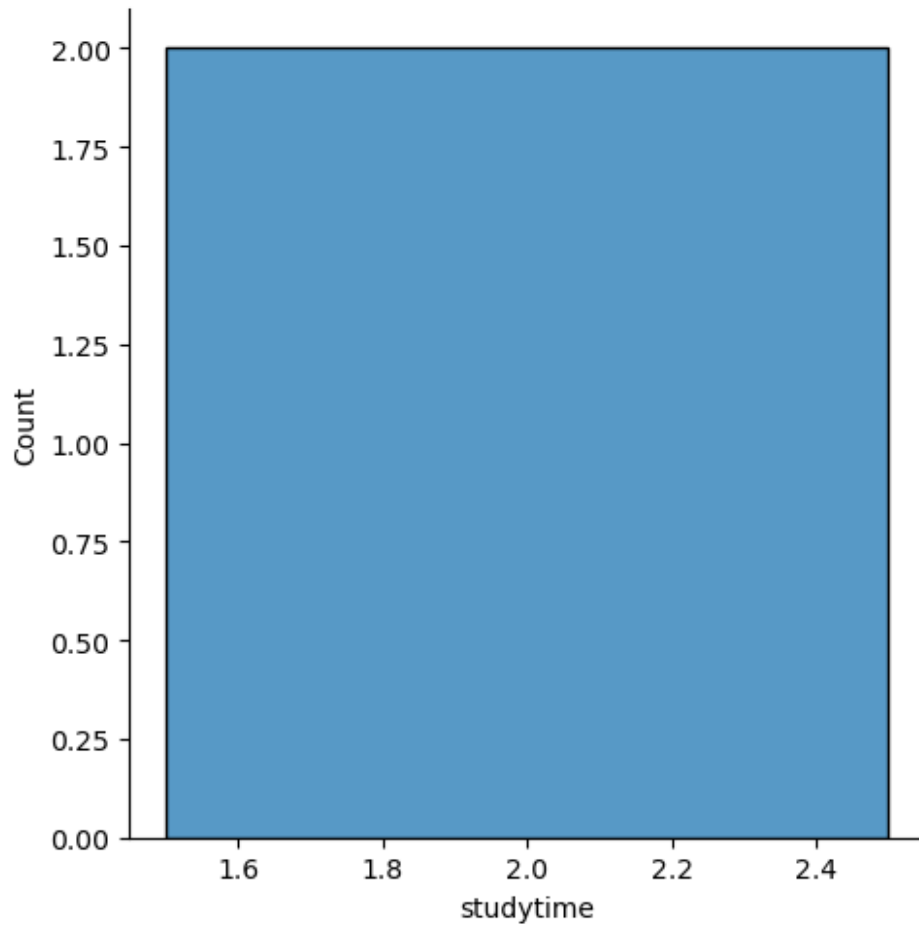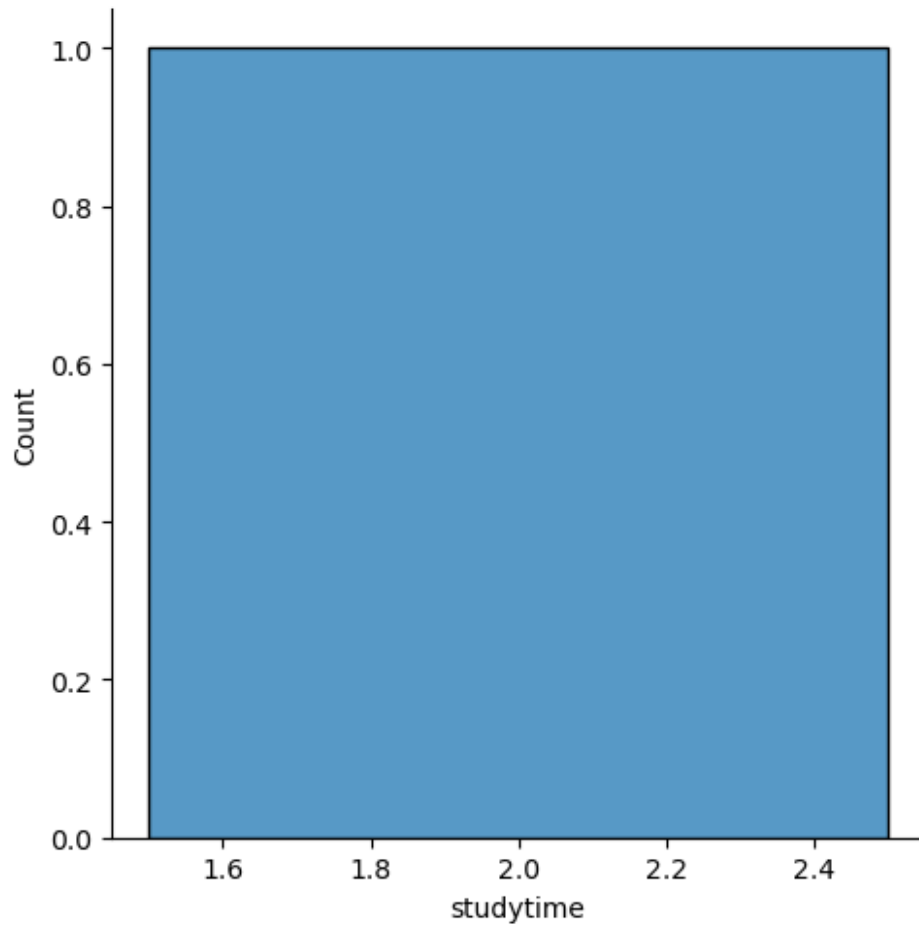
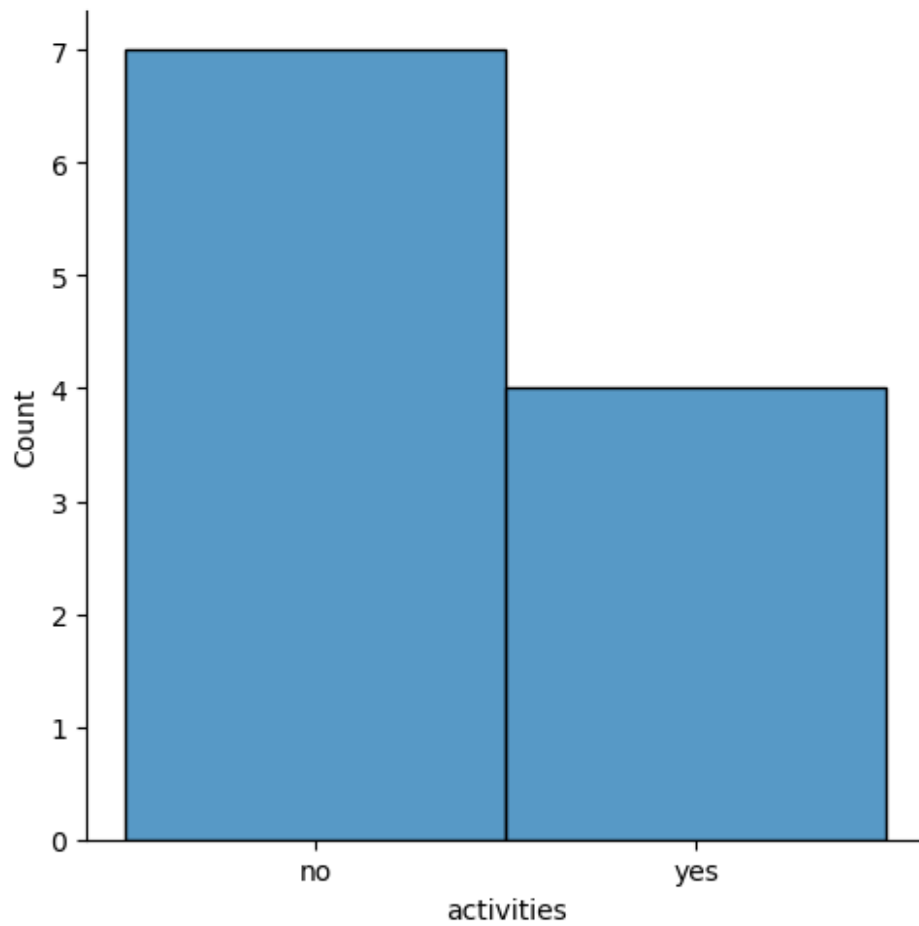**Distribution of Female student With Family Size More than 3 with School and Family Support Study Time**

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'F') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['schoolsup']=='yes') &
          (Fail_0_No_Fees['famsup']=='no')))

<seaborn.axisgrid.FacetGrid at 0x202a09bb260>
```
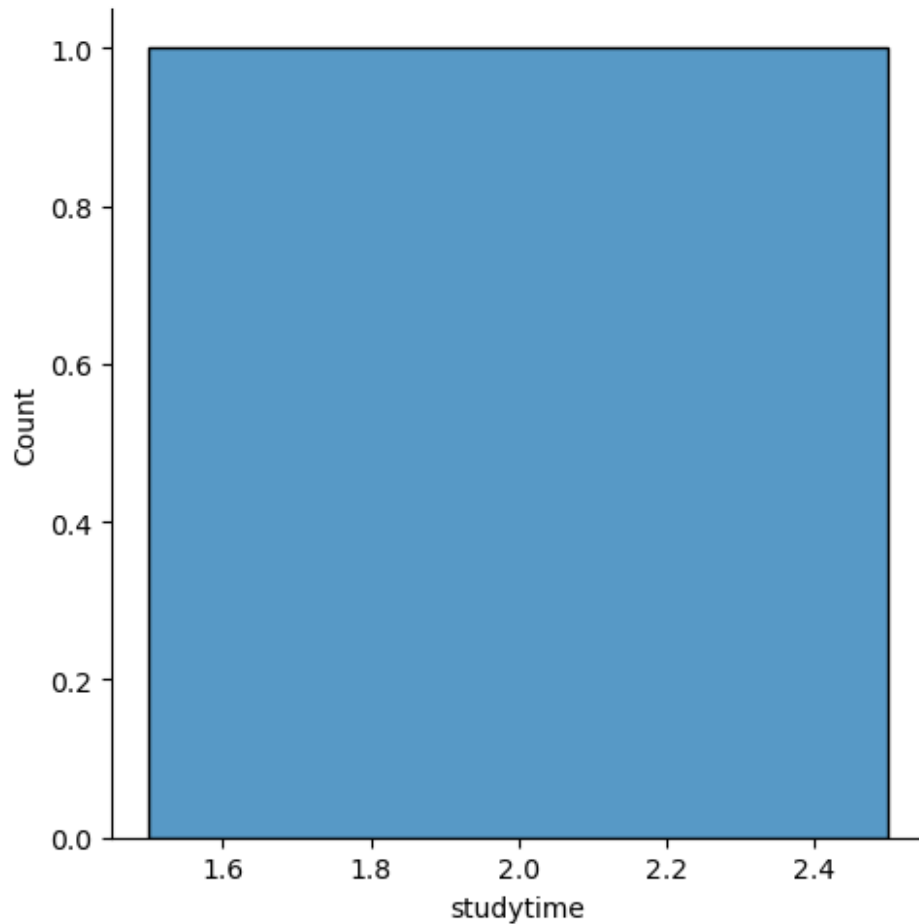
Distribution of Female student With Family Size More than 3 with School Support but without Family Support Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'F') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['schoolsup']=='no') &
            (Fail_0_No_Fees['famsup']=='yes')))

<seaborn.axisgrid.FacetGrid at 0x2029e68c140>
```
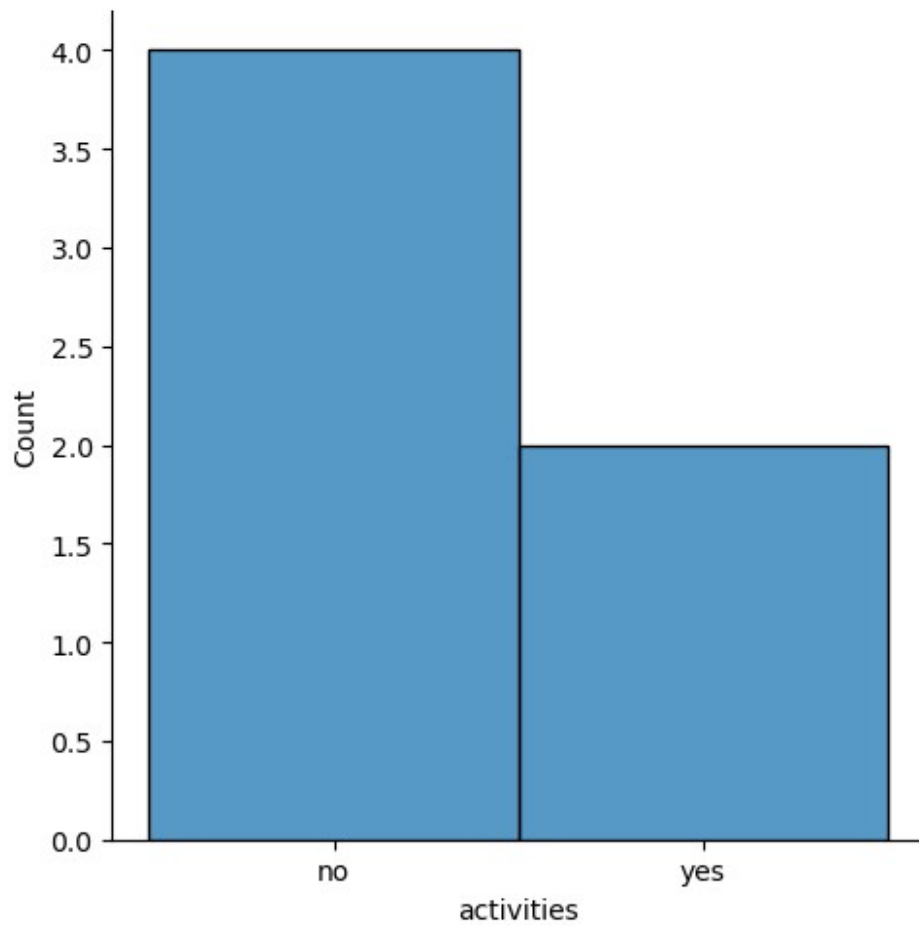
Distribution of Female student With Family Size More than 3 without School Support and with Family Support Study Time
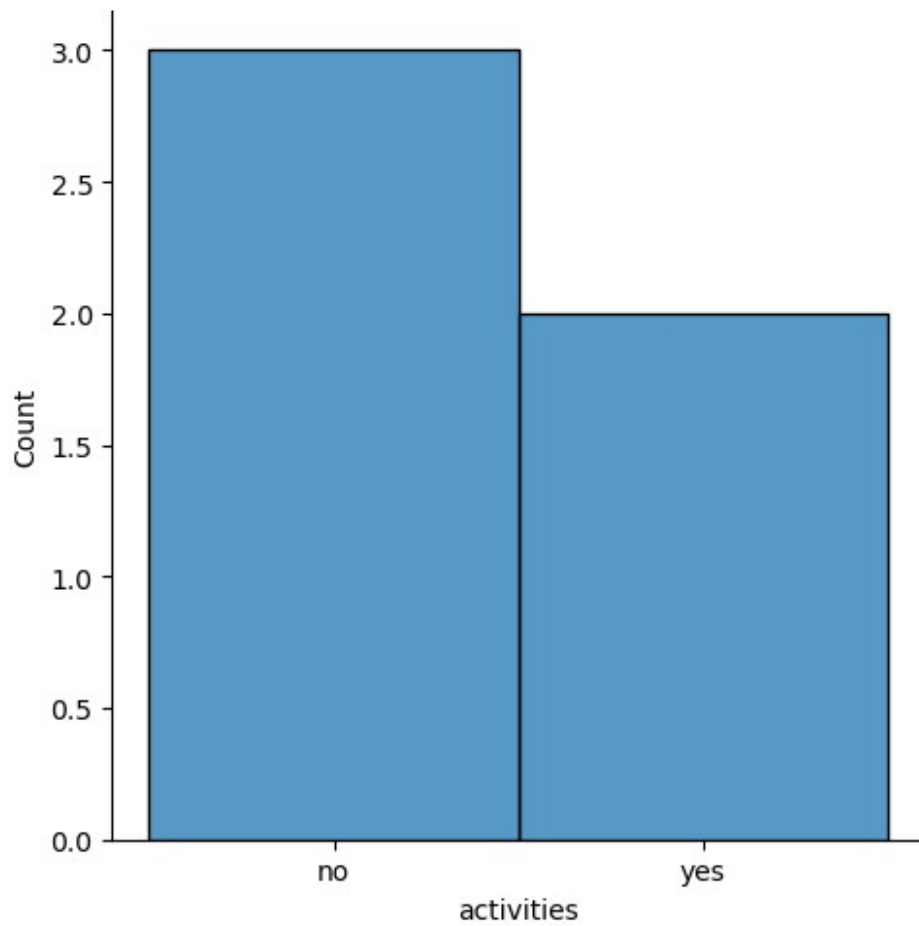
```
sns.displot(Fail_0_No_Fees['studytime'].where(Fail_0_No_Fees['sex']=='M'))

<seaborn.axisgrid.FacetGrid at 0x202a08d5820>
```
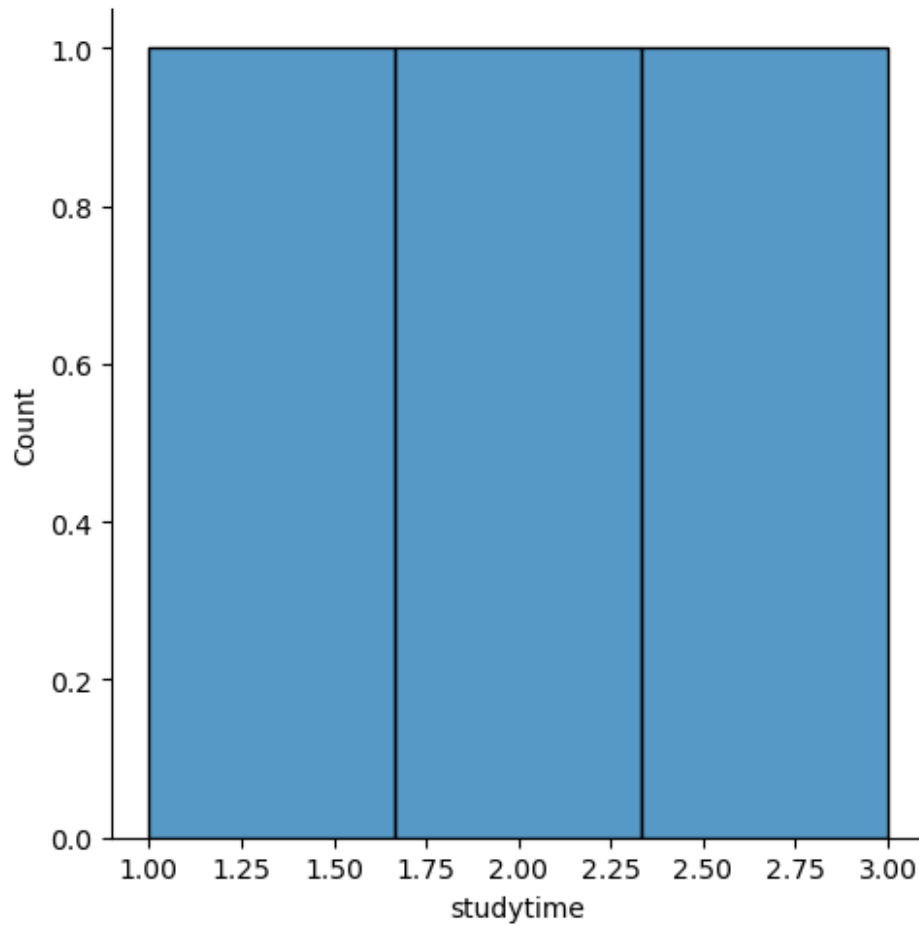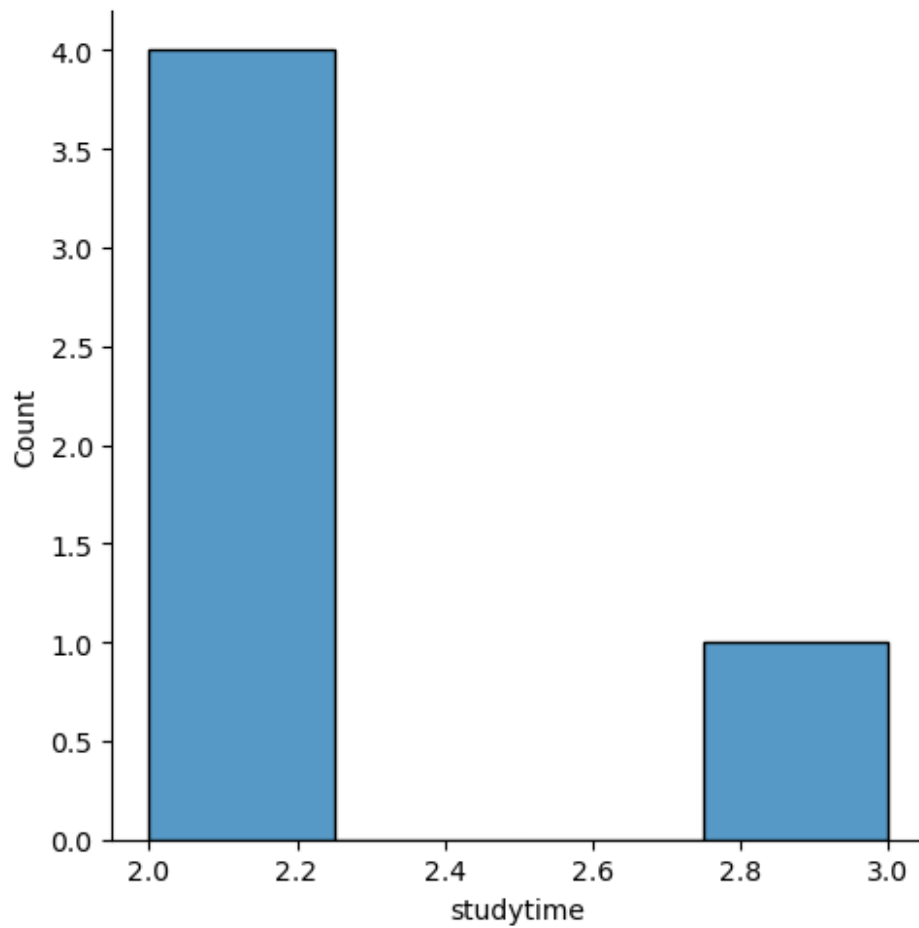
Distribution of Male student

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'M') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['schoolsup']=='no') &
        (Fail_0_No_Fees['famsup']=='no')))

<seaborn.axisgrid.FacetGrid at 0x202a0883080>
```

Distribution of Male student With Family Size More than 3 without School and Family Support Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'M') & (Fail_0_No_Fees['famsize']=='GT3') &
(Fail_0_No_Fees['schoolsup']=='no') &
          (Fail_0_No_Fees['famsup']=='yes')))

<seaborn.axisgrid.FacetGrid at 0x202a0b3d070>
```

Distribution of Male student With Family Size More than 3 without School Support and with Family Support Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'M') & (Fail_0_No_Fees['famsize']=='LE3') &
(Fail_0_No_Fees['schoolsup']=='no') &
        (Fail_0_No_Fees['famsup']=='no')))

<seaborn.axisgrid.FacetGrid at 0x202a0a3b350>
```

Distribution of Male student With Family Size Less than Equal to 3 without School and Family Support Study Time

```
sns.displot(Fail_0_No_Fees['studytime'].where((Fail_0_No_Fees['sex']==
'M') & (Fail_0_No_Fees['famsize']=='LE3') &
(Fail_0_No_Fees['schoolsup']=='no') &
        (Fail_0_No_Fees['famsup']=='yes')))

<seaborn.axisgrid.FacetGrid at 0x202a0bddd60>
```

Distribution of Male student With Family Size Less than Equal to 3 without School Support with Family Support Study Time

```
sns.displot(Fail_0_No_Fees['activities'])
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0b6eea0>
```

Distribution of Students involved in other activities along with studies

```
sns.displot(Fail_0_No_Fees['activities'].where(Fail_0_No_Fees['sex']==
'F'))
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0ce8da0>
```

Distribution of Female Students involved in other activities along with studies

```
sns.displot(Fail_0_No_Fees['activities'].where(Fail_0_No_Fees['sex']==
'M'))
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0cd6960>
```

Distribution of Male Students involved in other activities along with studies

```
sns.displot(Fail_0_No_Fees['Dalc'])
<seaborn.axisgrid.FacetGrid at 0x202a0e149b0>
```
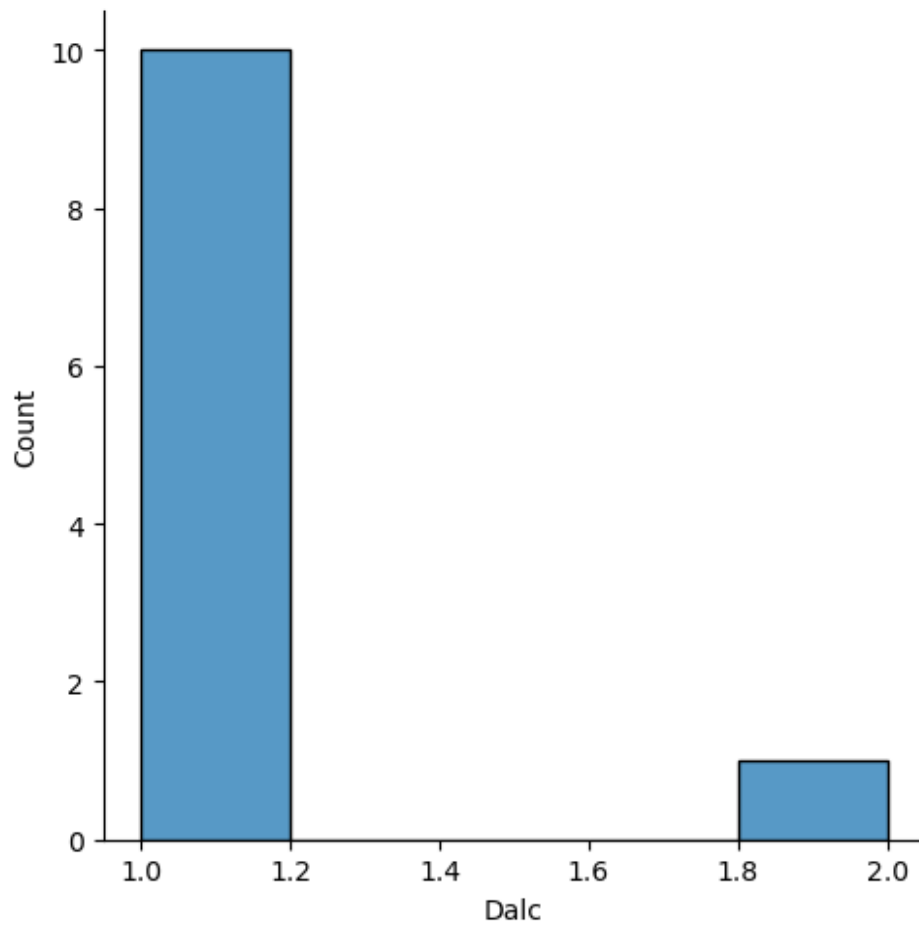
Distribution of Daily Alcohol Consumption by Students

```
sns.displot(Fail_0_No_Fees['Dalc'].where(Fail_0_No_Fees['sex']=='F'))
<seaborn.axisgrid.FacetGrid at 0x202a0c809e0>
```
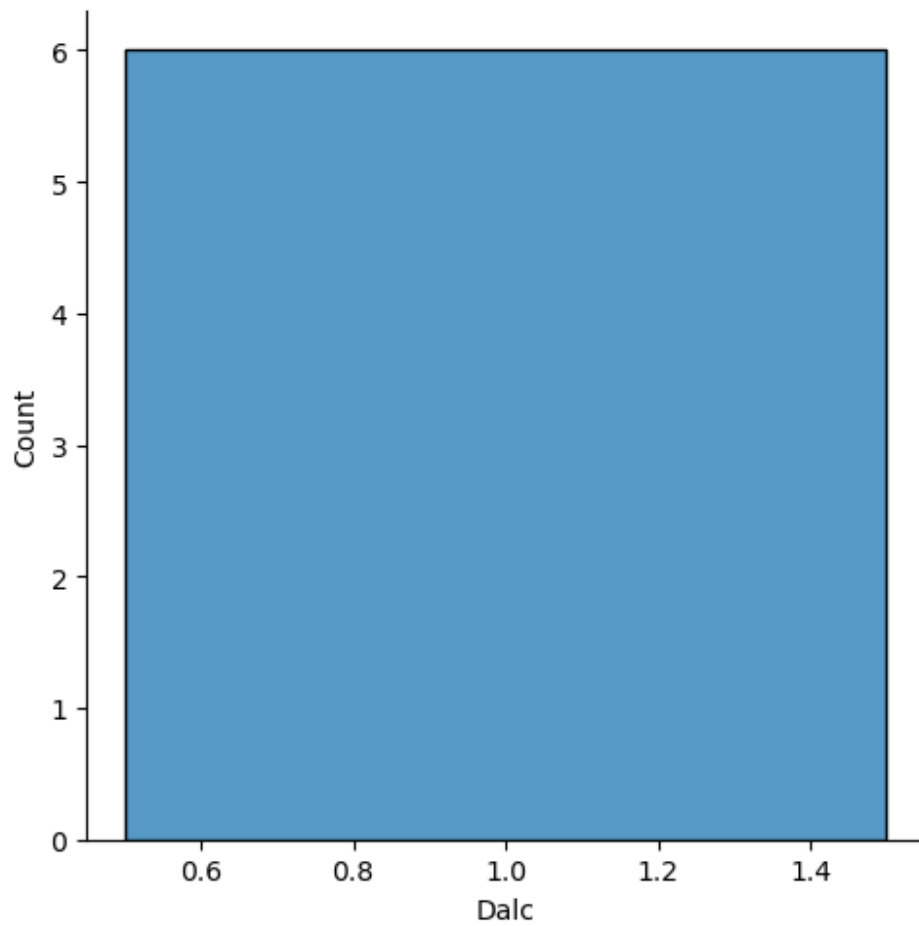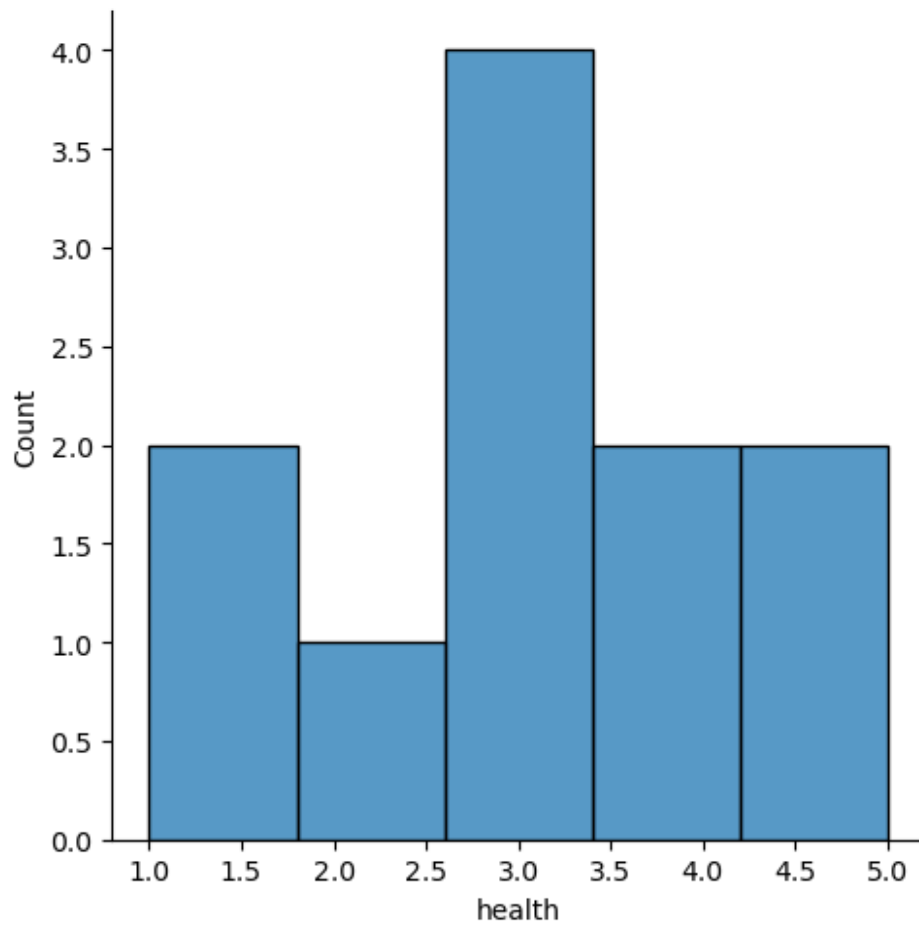
Distribution of Daily Alcohol Consumption by Female Students

```
sns.displot(Fail_0_No_Fees['Dalc'].where(Fail_0_No_Fees['sex']=='M'))
<seaborn.axisgrid.FacetGrid at 0x202a0aa1340>
```
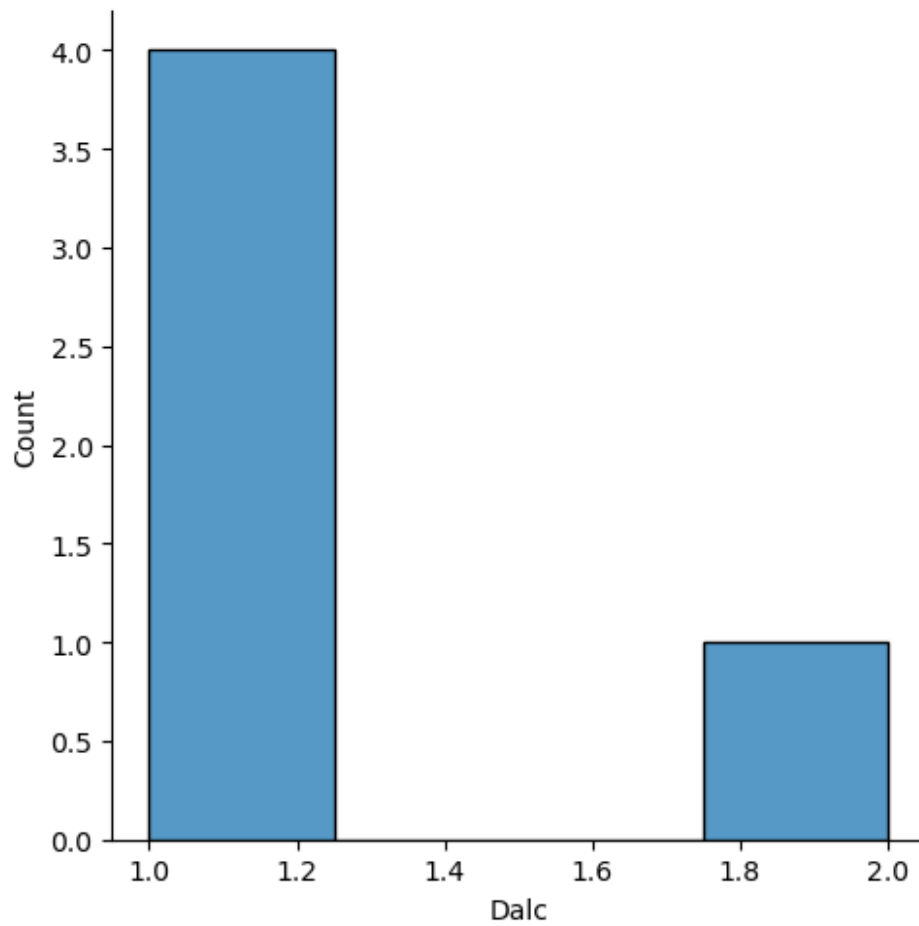
Distribution of Daily Alcohol Consumption by Male Students

```
sns.displot(Fail_0_No_Fees['Walc'])
<seaborn.axisgrid.FacetGrid at 0x202a0e8a840>
```
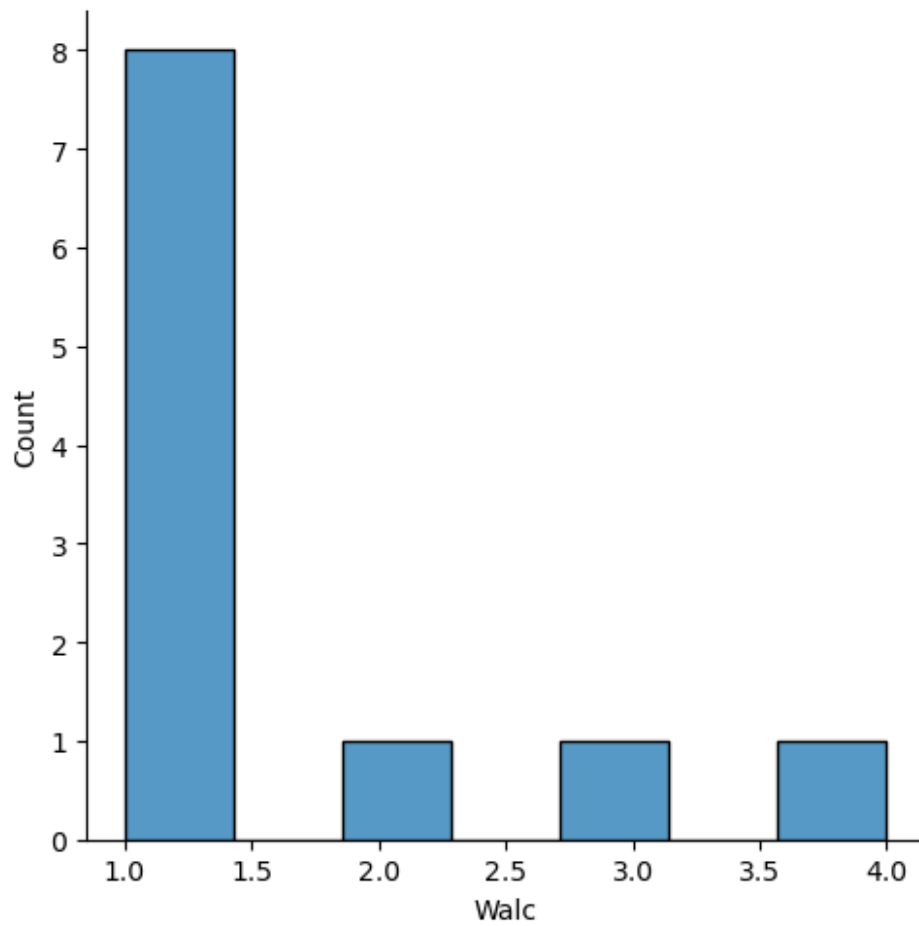
Distribution Weekly Alcohol Consumption By Students

```
sns.displot(Fail_0_No_Fees['Walc'].where(Fail_0_No_Fees['sex']=='F'))
<seaborn.axisgrid.FacetGrid at 0x2029d97af30>
```
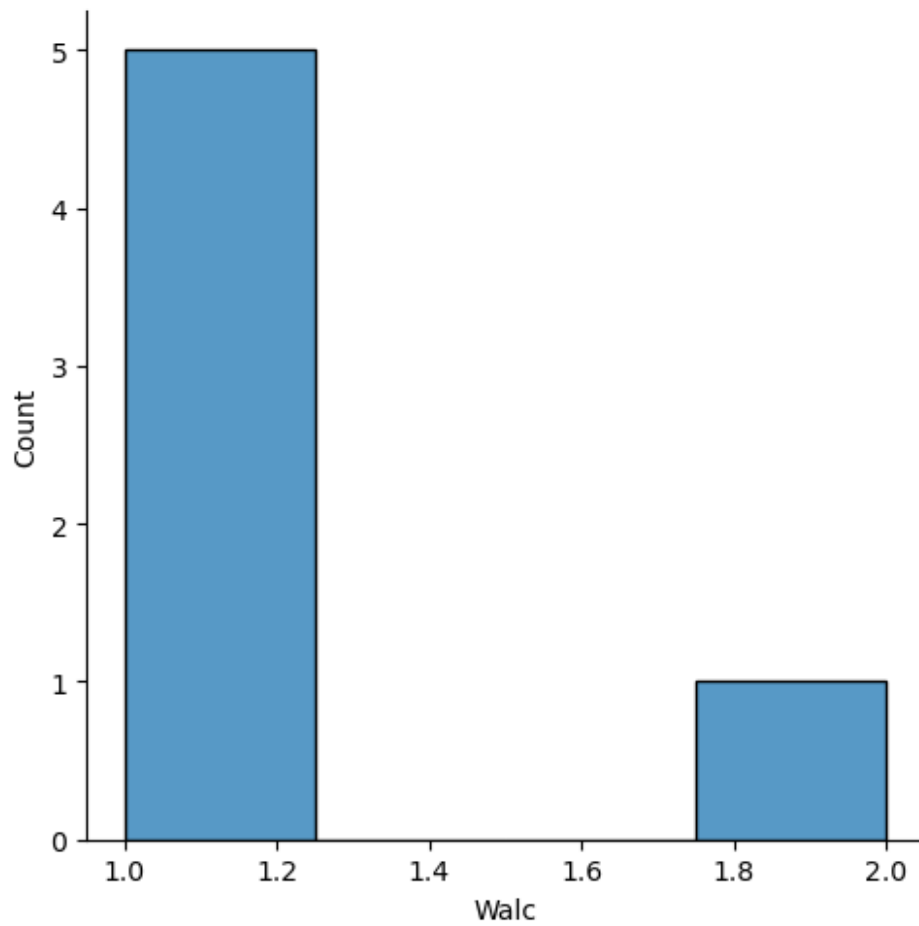
Distribution of Weekly Alcohol Consumption by Female Students
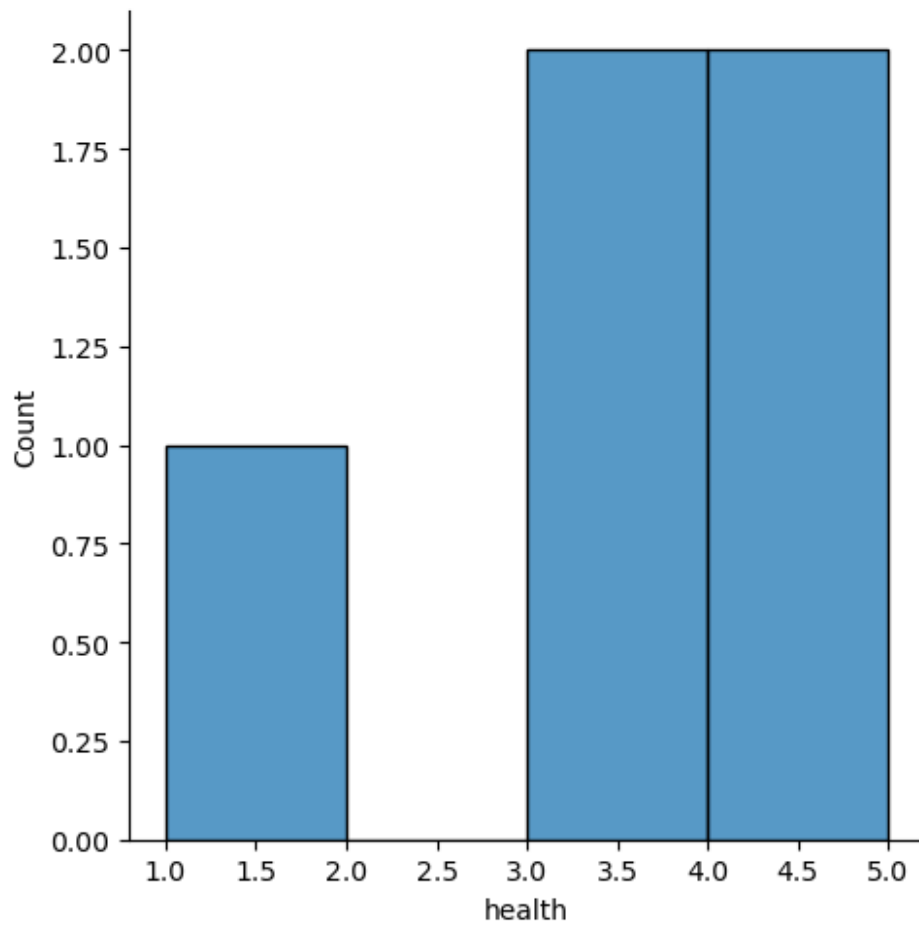
```
sns.displot(Fail_0_No_Fees['Walc'].where(Fail_0_No_Fees['sex']=='M'))
<seaborn.axisgrid.FacetGrid at 0x202a549c7a0>
```
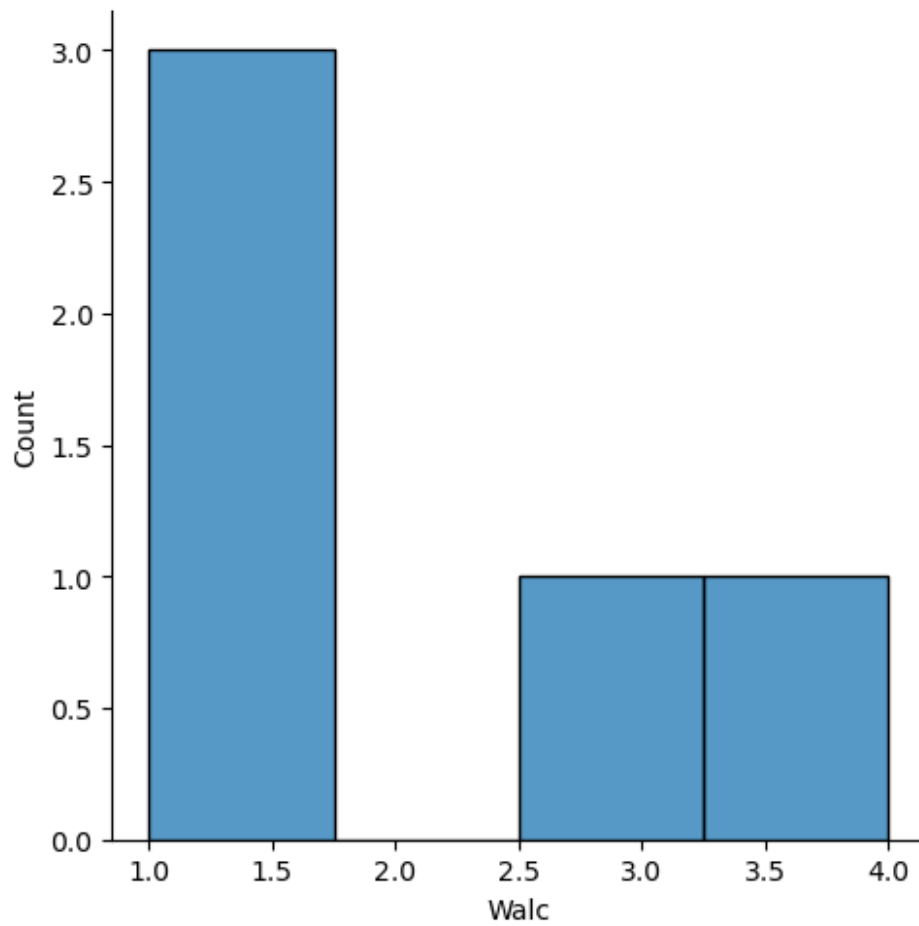
Distribution of Weekly Alcohol Comsumption by Male Students

```
sns.displot(Fail_0_No_Fees['health'])
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0f3a300>
```

Distribution of Students health Where 1 : Lowest 5 : Highest

```
sns.displot(Fail_0_No_Fees['health'].where(Fail_0_No_Fees['sex']=='F')
)

<seaborn.axisgrid.FacetGrid at 0x202a655c1d0>
```
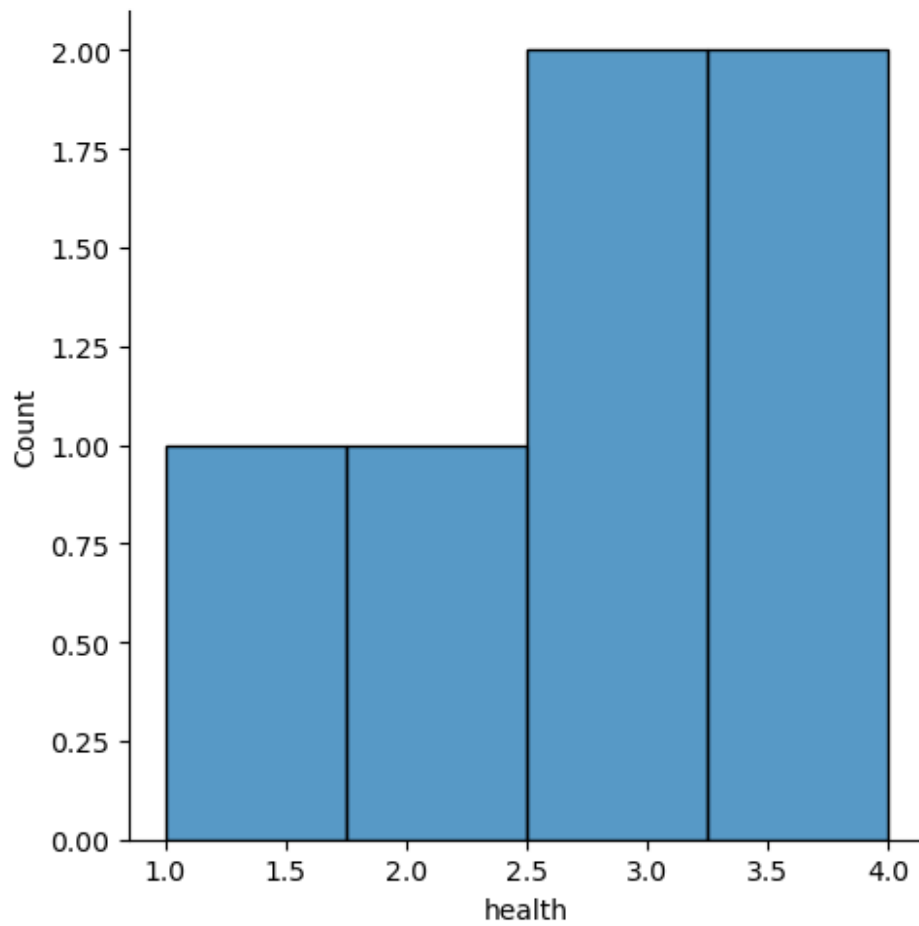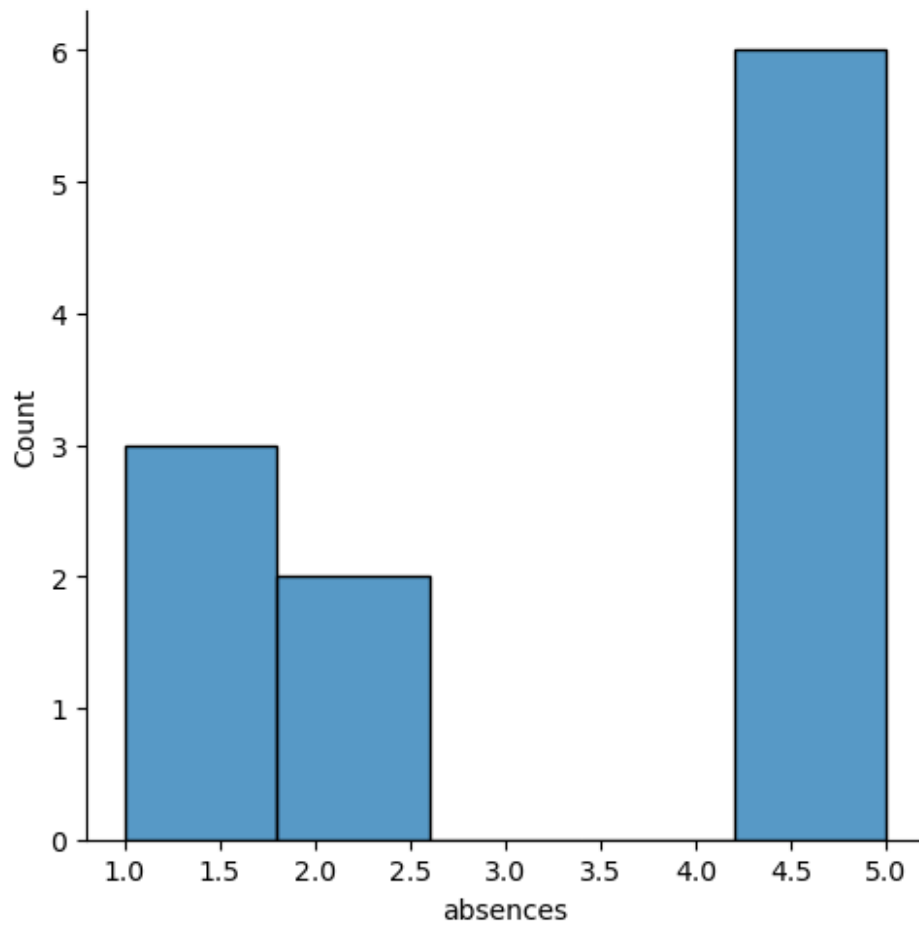
Distibution of Female Student Health

```
sns.displot(Fail_0_No_Fees['health'].where(Fail_0_No_Fees['sex']=='M')
)
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0b001a0>
```
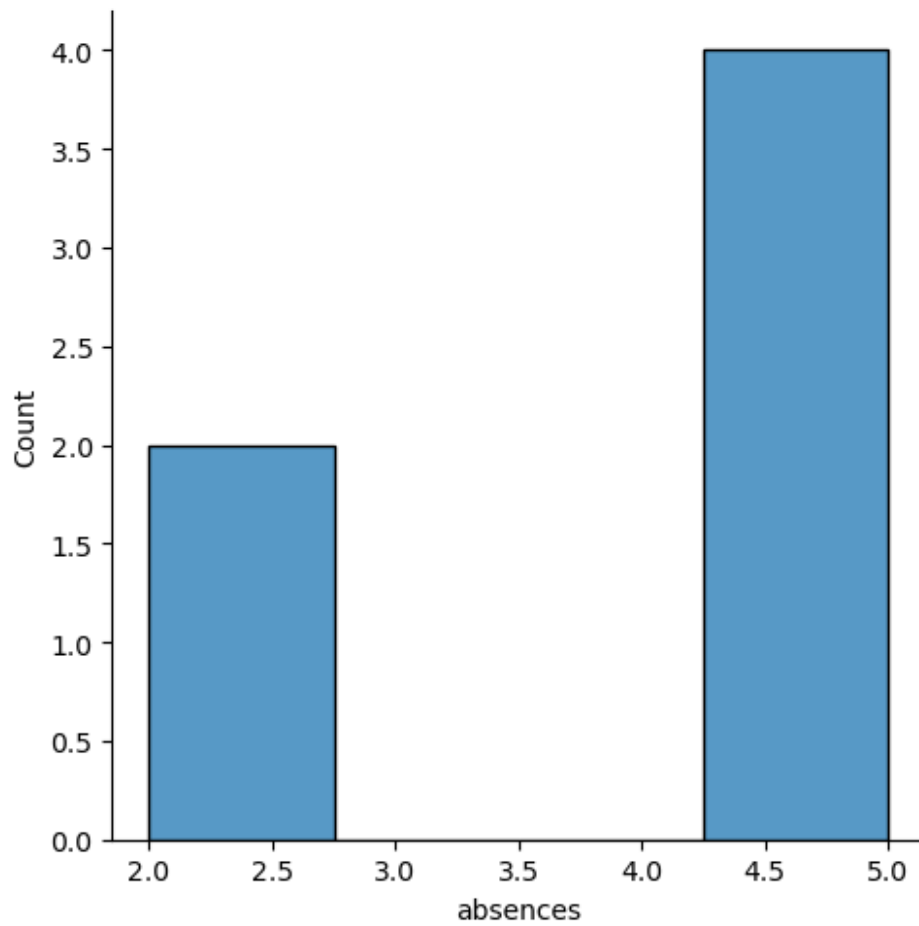
Distribution of Male Student health
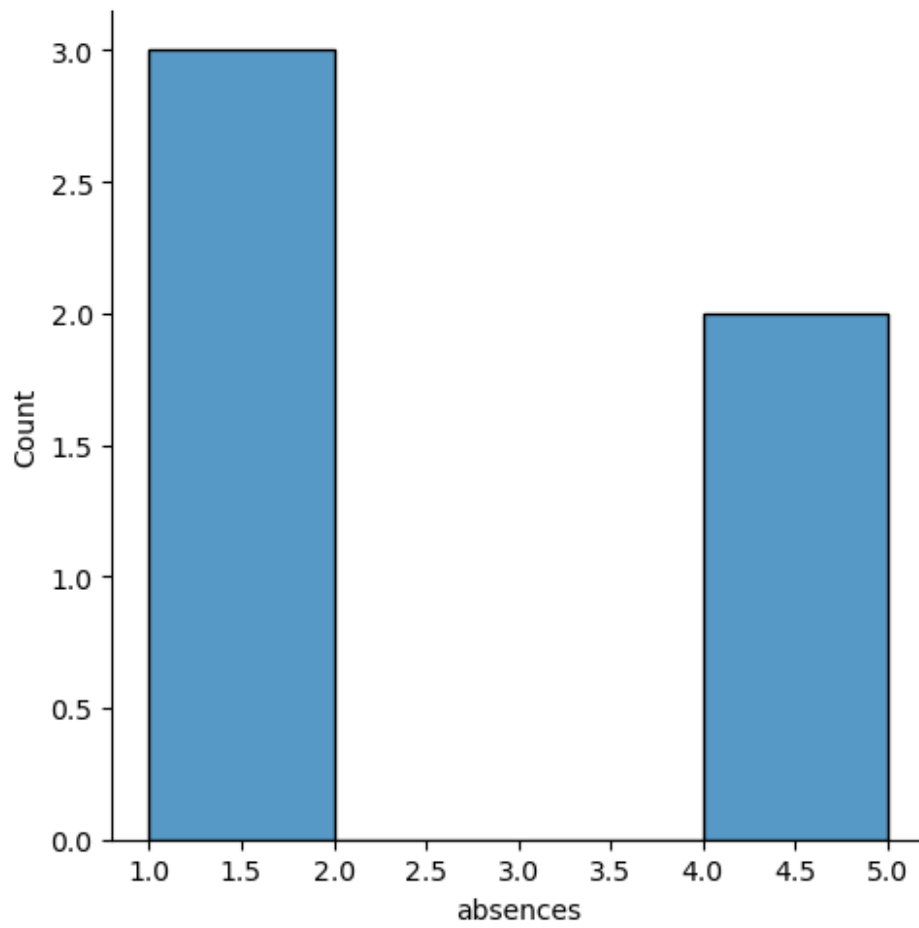
```
sns.displot(Fail_0_No_Fees['absences'])

<seaborn.axisgrid.FacetGrid at 0x202a65d3830>
```

Distribution of Absences of Students

```
sns.displot(Fail_0_No_Fees['absences'].where(Fail_0_No_Fees['sex']=='F
'))
```

```
<seaborn.axisgrid.FacetGrid at 0x202a0e8b8c0>
```

Distribution of Abesences of Female Students

```
sns.displot(Fail_0_No_Fees['absences'].where(Fail_0_No_Fees['sex']=='M
'))
```

```
<seaborn.axisgrid.FacetGrid at 0x202a673ad20>
```

Distribution of Absences of Male Students