

ULL Preliminary Analysis

January 13, 2025

1 Imports

```
[66]: import os
import copy
import math
import json
import collections

from IPython.display import display, HTML

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.preprocessing import OrdinalEncoder
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split, cross_validate, \
    ↳StratifiedKFold
from sklearn.metrics import get_scorer
from sklearn.preprocessing import KBinsDiscretizer

from aif360.algorithms.inprocessing import MetaFairClassifier, \
    ↳PrejudiceRemover, GerryFairClassifier
from fairlearn import metrics
```

2 General Settings

```
[67]: data_path = os.path.join("/", "home", "data")
ull_path = os.path.join(data_path, "ull")
ull_intermediate_path = os.path.join(ull_path, "intermediate")
ull_final_path = os.path.join(ull_path, "final")
```

3 Data Collection

```
[68]: df = pd.read_csv(os.path.join(ull_path, "original.csv"), low_memory=False)
df = df.set_index("id_questionnaire")
print(df.shape)
df
```

(80174, 159)

```
[68]:
```

	id_student_original	id_year	id_grade	id_class_group	\
id_questionnaire					
1	33613.0	2016	3	A	
2	19294.0	2016	3	A	
3	19587.0	2016	3	NaN	
4	29985.0	2016	3	A	
5	6982.0	2016	3	A	
...	
83853	2500.0	2019	6	B	
83854	5784.0	2019	6	A	
83855	7708.0	2019	6	X	
83856	18965.0	2019	6	B	
83857	7652.0	2019	6	A	

	id_school	id_student_16_19	id_school_16_19	score_MAT	\
id_questionnaire					
1	2415.0	14374.0	569.0	564.8700	
2	1842.0	8001.0	273.0	388.3400	
3	1432.0	8142.0	82.0	386.5900	
4	2280.0	12800.0	505.0	487.7600	
5	2040.0	2606.0	390.0	709.7900	
...	
83853	1278.0	NaN	271.0	400.8625	
83854	1079.0	NaN	78.0	597.0243	
83855	1486.0	3036.0	459.0	707.9254	
83856	1036.0	9913.0	36.0	522.8511	
83857	1173.0	5539.0	169.0	519.1200	

	level_MAT	score_LEN	...	\
id_questionnaire			...	
1	3.0	535.1500	...	
2	1.0	293.7000	...	
3	1.0	514.8100	...	
4	2.0	449.2500	...	
5	4.0	598.7200	...	
...	
83853	2.0	446.6522	...	
83854	3.0	632.6043	...	
83855	4.0	400.2761	...	

83856	3.0	656.1601	...
83857	3.0	426.1871	...

	t_extent_of_student_involvement_during_class \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.555556
5	0.666667
...	...
83853	0.555556
83854	0.666667
83855	NaN
83856	0.666667
83857	0.666667

	t_extent_of_teaching_methods_variety \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.666667
5	0.733333
...	...
83853	0.800000
83854	0.733333
83855	NaN
83856	0.666667
83857	0.666667

	t_extent_of_opinion_on_school	t_extent_of_class_behaviour \
id_questionnaire		
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	1.000000	0.500000
5	1.000000	0.833333
...
83853	1.000000	0.500000
83854	0.666667	0.666667
83855	NaN	NaN
83856	0.666667	0.666667
83857	0.666667	0.666667

	t_extent_of_resource_variety \
id_questionnaire	

1	NaN
2	NaN
3	NaN
4	0.458333
5	0.625000
...	...
83853	0.416667
83854	0.500000
83855	NaN
83856	0.458333
83857	0.375000

	t_extent_of_good_work_by_non_teachers \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.555556
5	0.888889
...	...
83853	1.000000
83854	1.000000
83855	NaN
83856	0.625000
83857	0.937500

	t_number_of_individual_training_topics \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	NaN
5	1.0
...	...
83853	1.0
83854	1.0
83855	NaN
83856	3.0
83857	2.0

	t_number_of_subjects_taught \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	NaN
5	NaN

```

...
83853      NaN
83854      NaN
83855      NaN
83856      NaN
83857      NaN

      t_extent_of_satisfaction_job_and_school \
id_questionnaire
1      NaN
2      NaN
3      NaN
4      NaN
5      NaN
...
83853      1.000000
83854      0.722222
83855      NaN
83856      0.611111
83857      0.666667

      t_extent_of_results_satisfaction
id_questionnaire
1      NaN
2      NaN
3      NaN
4      0.666667
5      1.000000
...
83853      0.166667
83854      0.500000
83855      NaN
83856      0.500000
83857      0.500000

```

[80174 rows x 159 columns]

4 Data Selection

```

[69]: df.groupby(by=["id_year", "id_grade"])[["id_year", "id_grade"]].count()
# df.groupby(by=["id_year", "id_grade"])[["id_year", "id_grade"]].count().plot.
     ↪ barh()

```

```

[69]:
id_year id_grade
2016    3      19466      19466

```

	6	19928	19928
2017	3	3056	3056
	4	2931	2931
	6	3110	3110
2018	3	3238	3238
	4	3106	3106
	6	3277	3277
2019	4	3199	3199
	6	18863	18863

```
[70]: df[["id_year", "id_grade", "s_gender"]].groupby(by=["id_year",
↳ "id_grade"])["s_gender"].count()
```

```
[70]: id_year  id_grade
2016      3      18582
        6      18550
2017      3       2991
        4       2930
        6       2973
2018      3      3133
        4      3106
        6      3110
2019      4       3191
        6      18658
Name: s_gender, dtype: int64
```

```
[71]: df[["id_year", "id_grade", "f_ESCS"]].groupby(by=["id_year",
↳ "id_grade"])["f_ESCS"].count()
```

```
[71]: id_year  id_grade
2016      3      13964
        6      14702
2017      3       1549
        4       2889
        6       1872
2018      3      2929
        4      3030
        6      2922
2019      4       3116
        6      11784
Name: f_ESCS, dtype: int64
```

```
[72]: df[["id_year", "id_grade", "level_MAT"]].groupby(by=["id_year",
↳ "id_grade"])["level_MAT"].count()
```

```
[72]: id_year  id_grade
2016      3      17806
```

```

        6          17916
2017    3          2995
        4          2362
        6          3031
2018    3          3102
        4          2397
        6          3186
2019    4          2385
        6          18019
Name: level_MAT, dtype: int64

```

```

[73]: df_sel = df[(df["id_year"] == 2016) & (df["id_grade"] == 3)]
df_sel = df_sel[df_sel['level_MAT'].notna()]
print(df_sel.shape)
print(df_sel["s_gender"].isna().sum())
print(df_sel["f_ESCS"].isna().sum())

```

```

(17806, 159)
1
4903

```

```

[74]: df_sel["f_ESCS"].describe()

```

```

[74]: count    12903.000000
      mean         0.018232
      std         0.994479
      min        -2.706281
      25%        -0.735529
      50%         0.016366
      75%         0.638482
      max         3.151773
      Name: f_ESCS, dtype: float64

```

5 Feature Selection

5.1 Utils

```

[75]: def filter_columns_starting_with(temp_df, pattern):
      df = copy.deepcopy(temp_df)
      if pattern is not None:
          return df[[col for col in df.columns if col.startswith(pattern)] +
                    ↪ [col for col in df.columns if col.startswith("level_")]]
      else:
          return df

```

```

[76]: def get_missing_values(temp_df):
      df = copy.deepcopy(temp_df)
      return df.isna().sum()/df.shape[0]

```

```
[77]: def discretize_columns(temp_df):
    df = copy.deepcopy(temp_df)
    categorical_features = df.select_dtypes(include=['object', 'category']).
    ↪columns

    # Apply Ordinal Encoding
    encoder = OrdinalEncoder()
    df[categorical_features] = encoder.fit_transform(df[categorical_features])
    return df
```

```
[78]: def compute_correlation_matrix(temp_df, annot=True):
    df = copy.deepcopy(temp_df)
    correlation_matrix = df.corr()
    ax = sns.heatmap(correlation_matrix, annot=annot, cmap='coolwarm', fmt="."
    ↪2f")
    ax.collections[0].set_clim(-1, 1)
    plt.title("Correlation Matrix Heatmap")
    plt.show()
    return correlation_matrix
```

```
[79]: def compute_exploratory_analysis(df, column_pattern, missing_threshold=0.09,
    ↪annot=True):
    df_filtered = filter_columns_starting_with(df, column_pattern)
    missings = get_missing_values(df_filtered)
    good_missings = missings[missings < missing_threshold]
    print(f"GOOD FEATURES: {good_missings.shape[0]}")
    print(good_missings)
    print()
    print("BAD FEATURES")
    bad_missings = missings[missings > missing_threshold]
    print(bad_missings)
    print()
    df_without_bad_cols = df_filtered.drop(bad_missings.index, axis=1)
    df_discretized = discretize_columns(df_without_bad_cols)
    correlation_matrix = compute_correlation_matrix(df_discretized, annot)
    display(df_without_bad_cols)
    display(correlation_matrix)
    file_names = {
        "id_": "general",
        "s_": "students",
        "t_": "teachers",
        "f_": "families",
        "p_": "principals"
    }
    if column_pattern is not None:
        df_without_bad_cols.to_csv(os.path.join(ull_intermediate_path,
    ↪f"{file_names[column_pattern]}.csv"))
```



```
return df_without_bad_cols, correlation_matrix
```

5.2 All Students

```
[80]: df_general, correlation_matrix_general = compute_exploratory_analysis(df=df, ↵  
    ↪column_pattern="id_", missing_threshold=1.0, annot=True)
```

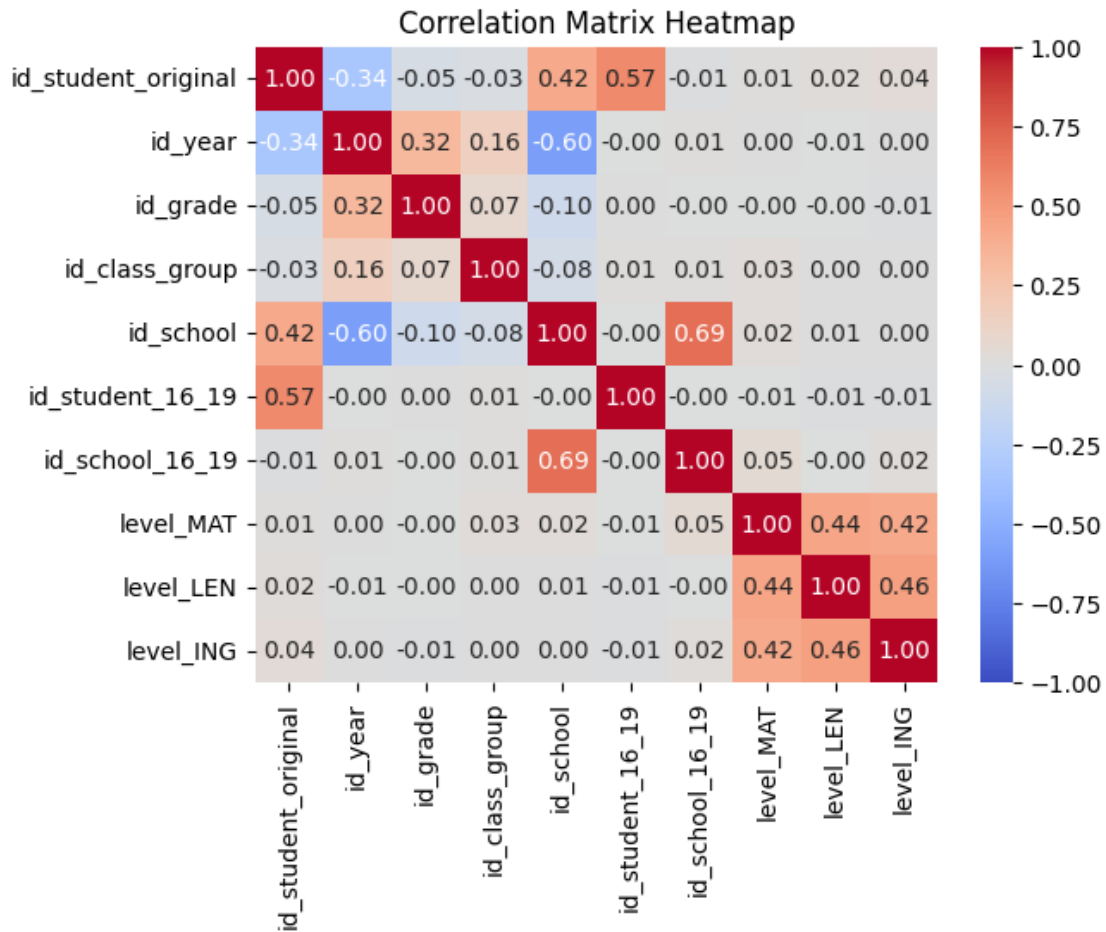
GOOD FEATURES: 10

id_student_original	0.010053
id_year	0.000000
id_grade	0.000000
id_class_group	0.113329
id_school	0.001110
id_student_16_19	0.582308
id_school_16_19	0.305286
level_MAT	0.086998
level_LEN	0.067491
level_ING	0.370681

dtype: float64

BAD FEATURES

Series([], dtype: float64)



	id_student_original	id_year	id_grade	id_class_group	\
id_questionnaire					
1	33613.0	2016	3	A	
2	19294.0	2016	3	A	
3	19587.0	2016	3	NaN	
4	29985.0	2016	3	A	
5	6982.0	2016	3	A	
...	
83853	2500.0	2019	6	B	
83854	5784.0	2019	6	A	
83855	7708.0	2019	6	X	
83856	18965.0	2019	6	B	
83857	7652.0	2019	6	A	

	id_school	id_student_16_19	id_school_16_19	level_MAT	\
id_questionnaire					
1	2415.0	14374.0	569.0	3.0	
2	1842.0	8001.0	273.0	1.0	

3	1432.0	8142.0	82.0	1.0
4	2280.0	12800.0	505.0	2.0
5	2040.0	2606.0	390.0	4.0
...
83853	1278.0	NaN	271.0	2.0
83854	1079.0	NaN	78.0	3.0
83855	1486.0	3036.0	459.0	4.0
83856	1036.0	9913.0	36.0	3.0
83857	1173.0	5539.0	169.0	3.0

	level_LEN	level_ING
id_questionnaire		
1	3.0	NaN
2	1.0	NaN
3	3.0	NaN
4	2.0	NaN
5	3.0	NaN
...
83853	2.0	1.0
83854	4.0	4.0
83855	2.0	2.0
83856	4.0	3.0
83857	2.0	2.0

[80174 rows x 10 columns]

	id_student_original	id_year	id_grade	id_class_group	\
id_student_original	1.000000	-0.336732	-0.052204	-0.031246	
id_year	-0.336732	1.000000	0.324603	0.160981	
id_grade	-0.052204	0.324603	1.000000	0.073813	
id_class_group	-0.031246	0.160981	0.073813	1.000000	
id_school	0.417052	-0.595819	-0.103873	-0.077164	
id_student_16_19	0.574448	-0.000049	0.000538	0.010418	
id_school_16_19	-0.007898	0.010441	-0.002465	0.007873	
level_MAT	0.014170	0.000758	-0.003471	0.027510	
level_LEN	0.020464	-0.006925	-0.002249	0.004177	
level_ING	0.038553	0.002740	-0.014849	0.001292	

	id_school	id_student_16_19	id_school_16_19	level_MAT	\
id_student_original	0.417052	0.574448	-0.007898	0.014170	
id_year	-0.595819	-0.000049	0.010441	0.000758	
id_grade	-0.103873	0.000538	-0.002465	-0.003471	
id_class_group	-0.077164	0.010418	0.007873	0.027510	
id_school	1.000000	-0.003389	0.688965	0.023545	
id_student_16_19	-0.003389	1.000000	-0.004426	-0.006985	
id_school_16_19	0.688965	-0.004426	1.000000	0.053949	
level_MAT	0.023545	-0.006985	0.053949	1.000000	
level_LEN	0.005499	-0.013151	-0.004213	0.438571	

level_ING	0.003349	-0.011873	0.015791	0.415471
-----------	----------	-----------	----------	----------

	level_LEN	level_ING
id_student_original	0.020464	0.038553
id_year	-0.006925	0.002740
id_grade	-0.002249	-0.014849
id_class_group	0.004177	0.001292
id_school	0.005499	0.003349
id_student_16_19	-0.013151	-0.011873
id_school_16_19	-0.004213	0.015791
level_MAT	0.438571	0.415471
level_LEN	1.000000	0.462179
level_ING	0.462179	1.000000

5.3 General Info

```
[81]: df_sel_gen, correlation_matrix_sel_gen =   
      ↪compute_exploratory_analysis(df=df_sel, column_pattern="id_",   
      ↪missing_threshold=1.0, annot=True)
```

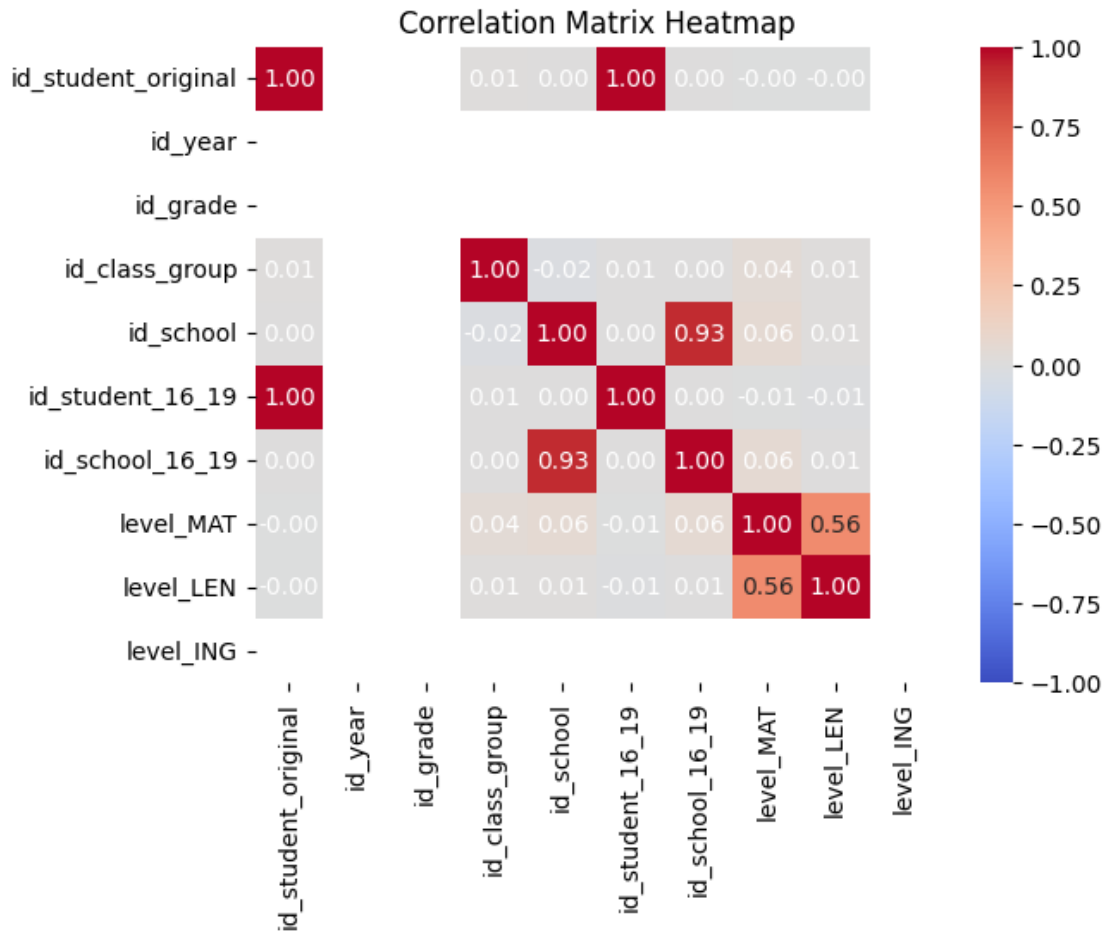
GOOD FEATURES: 9

id_student_original	0.000000
id_year	0.000000
id_grade	0.000000
id_class_group	0.154667
id_school	0.000000
id_student_16_19	0.103898
id_school_16_19	0.062956
level_MAT	0.000000
level_LEN	0.009772

dtype: float64

BAD FEATURES

Series([], dtype: float64)



	id_student_original	id_year	id_grade	id_class_group	\
id_questionnaire					
1	33613.0	2016	3	A	
2	19294.0	2016	3	A	
3	19587.0	2016	3	NaN	
4	29985.0	2016	3	A	
5	6982.0	2016	3	A	
...	
20426	5794.0	2016	3	NaN	
20427	3000.0	2016	3	A	
20428	36202.0	2016	3	B	
20429	21532.0	2016	3	A	
20430	3094.0	2016	3	NaN	

	id_school	id_student_16_19	id_school_16_19	level_MAT	\
id_questionnaire					
1	2415.0	14374.0	569.0	3.0	
2	1842.0	8001.0	273.0	1.0	

3	1432.0	8142.0	82.0	1.0
4	2280.0	12800.0	505.0	2.0
5	2040.0	2606.0	390.0	4.0
...
20426	1863.0	NaN	NaN	1.0
20427	2203.0	877.0	468.0	2.0
20428	1368.0	15545.0	51.0	2.0
20429	2373.0	8990.0	549.0	3.0
20430	1907.0	912.0	327.0	2.0

	level_LEN	level_ING
id_questionnaire		
1	3.0	NaN
2	1.0	NaN
3	3.0	NaN
4	2.0	NaN
5	3.0	NaN
...
20426	1.0	NaN
20427	3.0	NaN
20428	2.0	NaN
20429	3.0	NaN
20430	3.0	NaN

[17806 rows x 10 columns]

	id_student_original	id_year	id_grade	id_class_group	\
id_student_original	1.000000	NaN	NaN	0.008156	
id_year	NaN	NaN	NaN	NaN	
id_grade	NaN	NaN	NaN	NaN	
id_class_group	0.008156	NaN	NaN	1.000000	
id_school	0.000439	NaN	NaN	-0.020820	
id_student_16_19	0.999983	NaN	NaN	0.005759	
id_school_16_19	0.000417	NaN	NaN	0.000943	
level_MAT	-0.003242	NaN	NaN	0.042456	
level_LEN	-0.004734	NaN	NaN	0.011373	
level_ING	NaN	NaN	NaN	NaN	

	id_school	id_student_16_19	id_school_16_19	level_MAT	\
id_student_original	0.000439	0.999983	0.000417	-0.003242	
id_year	NaN	NaN	NaN	NaN	
id_grade	NaN	NaN	NaN	NaN	
id_class_group	-0.020820	0.005759	0.000943	0.042456	
id_school	1.000000	0.000108	0.929463	0.061969	
id_student_16_19	0.000108	1.000000	0.000665	-0.006903	
id_school_16_19	0.929463	0.000665	1.000000	0.059139	
level_MAT	0.061969	-0.006903	0.059139	1.000000	
level_LEN	0.012964	-0.010989	0.006073	0.562532	

level_ING	NaN	NaN	NaN	NaN
	level_LEN	level_ING		
id_student_original	-0.004734	NaN		
id_year	NaN	NaN		
id_grade	NaN	NaN		
id_class_group	0.011373	NaN		
id_school	0.012964	NaN		
id_student_16_19	-0.010989	NaN		
id_school_16_19	0.006073	NaN		
level_MAT	0.562532	NaN		
level_LEN	1.000000	NaN		
level_ING	NaN	NaN		

5.4 Student Questionnaire

```
[82]: df_sel_students, correlation_matrix_sel_students = \
    ↪compute_exploratory_analysis(df=df_sel, column_pattern="s_", \
    ↪missing_threshold=0.095, annot=False)
```

GOOD FEATURES: 18

s_gender	0.000056
s_birth_year	0.050376
s_is_living_with_mother	0.071661
s_has_repeated	0.081152
s_frequency_of_skips	0.060261
s_homeworks_day_per_week	0.046894
s_extent_of_teachers_affinity	0.076772
s_frequency_of_computer_usage	0.056049
s_frequency_of_internet_usage	0.055824
s_frequency_of_work_with_teachers	0.056329
s_frequency_of_materials_in_class	0.059811
s_frequency_of_evaluations	0.058632
s_extent_of_teacher_performance	0.059025
s_extent_of_class_env	0.059250
s_extent_of_classmates_affinity	0.060036
s_extent_of_school_satisfaction	0.057396
level_MAT	0.000000
level_LEN	0.009772

dtype: float64

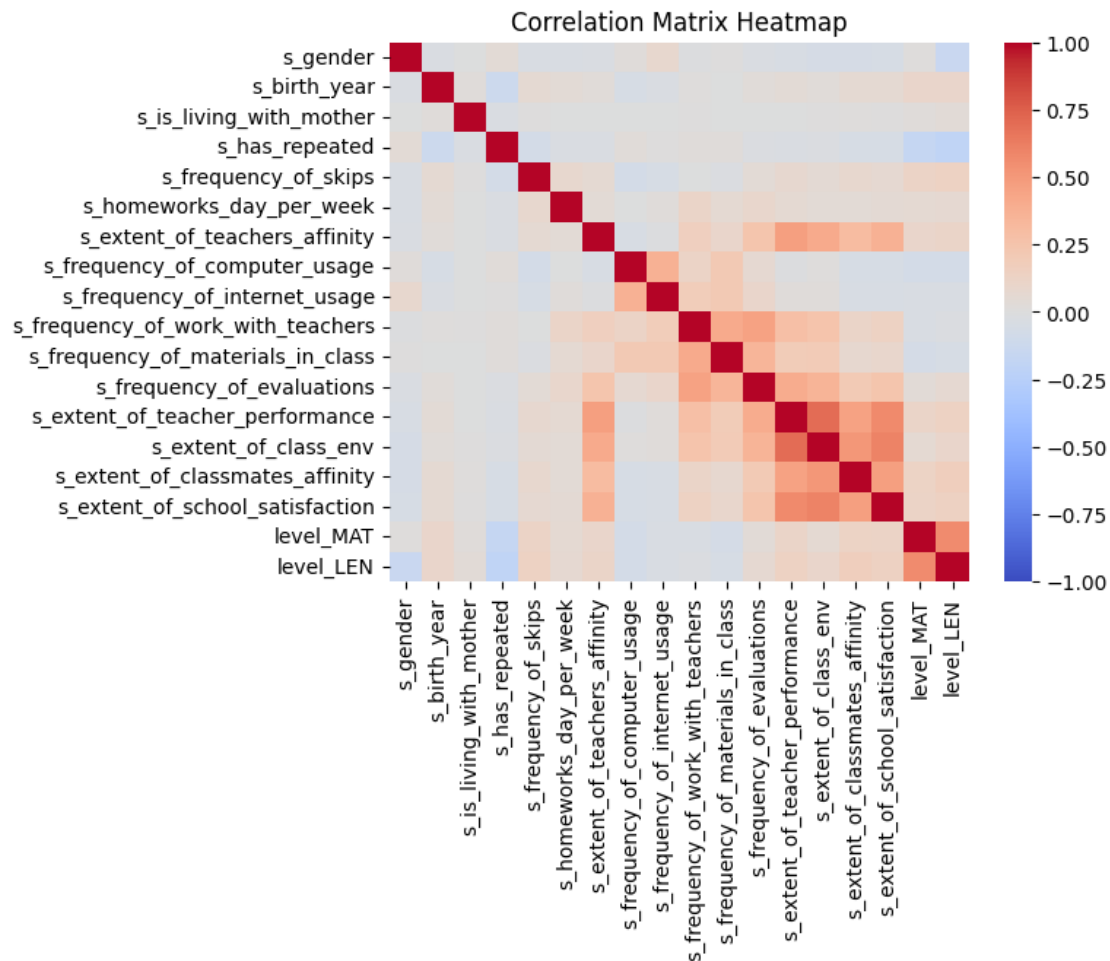
BAD FEATURES

s_is_living_with_father	0.137594
s_is_living_with_siblings	0.387229
s_is_living_with_other_relatives	0.837190
s_birth_country	1.000000
s_nationality_country	1.000000
s_weight	1.000000

```

s_extent_of_math_affinity      1.000000
s_extent_of_reading_affinity    1.000000
level_ING                      1.000000
dtype: float64

```



id_questionnaire	s_gender	s_birth_year	s_is_living_with_mother	\
1	FEMALE	2007.0	True	
2	MALE	2007.0	True	
3	FEMALE	2007.0	True	
4	MALE	2007.0	True	
5	FEMALE	2007.0	True	
...	
20426	FEMALE	2001.0	True	
20427	FEMALE	2007.0	True	
20428	MALE	2001.0	True	
20429	MALE	2007.0	True	

20430	FEMALE	2007.0	True
-------	--------	--------	------

	s_has_repeated	s_frequency_of_skips \
id_questionnaire		
1	False	1.000000
2	True	1.000000
3	False	1.000000
4	False	1.000000
5	False	1.000000
...
20426	NaN	0.000000
20427	False	1.000000
20428	NaN	0.333333
20429	False	1.000000
20430	False	1.000000

	s_homeworks_day_per_week	s_extent_of_teachers_affinity \
id_questionnaire		
1	0.0	1.0
2	3.0	1.0
3	2.0	NaN
4	3.0	1.0
5	1.0	1.0
...
20426	2.0	NaN
20427	3.0	1.0
20428	4.0	NaN
20429	3.0	1.0
20430	2.0	1.0

	s_frequency_of_computer_usage \
id_questionnaire	
1	0.111111
2	0.333333
3	0.666667
4	0.333333
5	0.000000
...	...
20426	NaN
20427	0.000000
20428	0.222222
20429	0.000000
20430	0.111111

	s_frequency_of_internet_usage \
id_questionnaire	
1	0.066667
2	0.600000

3	0.666667
4	0.266667
5	0.133333
...	...
20426	0.333333
20427	0.133333
20428	0.933333
20429	0.533333
20430	0.400000

	s_frequency_of_work_with_teachers \
id_questionnaire	
1	0.733333
2	0.533333
3	0.833333
4	0.400000
5	0.566667
...	...
20426	0.333333
20427	0.666667
20428	0.518519
20429	0.633333
20430	0.629630

	s_frequency_of_materials_in_class \
id_questionnaire	
1	0.380952
2	0.714286
3	1.000000
4	NaN
5	0.666667
...	...
20426	NaN
20427	0.476190
20428	0.777778
20429	0.142857
20430	0.476190

	s_frequency_of_evaluations	s_extent_of_teacher_performance \
id_questionnaire		
1	0.740741	0.833333
2	0.888889	0.800000
3	0.750000	0.800000
4	0.541667	0.800000
5	0.888889	1.000000
...
20426	0.333333	0.333333
20427	0.740741	0.900000

20428	0.740741	0.566667
20429	0.740741	1.000000
20430	0.481481	0.733333

	s_extent_of_class_env	s_extent_of_classmates_affinity \
id_questionnaire		
1	0.791667	0.714286
2	0.791667	0.714286
3	0.833333	0.833333
4	0.750000	0.857143
5	1.000000	0.952381
...
20426	0.333333	0.333333
20427	0.916667	0.857143
20428	0.958333	0.666667
20429	0.875000	0.714286
20430	0.708333	0.619048

	s_extent_of_school_satisfaction	level_MAT	level_LEN
id_questionnaire			
1	0.944444	3.0	3.0
2	1.000000	1.0	1.0
3	0.888889	1.0	3.0
4	0.722222	2.0	2.0
5	1.000000	4.0	3.0
...
20426	1.000000	1.0	1.0
20427	1.000000	2.0	3.0
20428	0.833333	2.0	2.0
20429	1.000000	3.0	3.0
20430	0.722222	2.0	3.0

[17806 rows x 18 columns]

	s_gender	s_birth_year \
s_gender	1.000000	-0.025734
s_birth_year	-0.025734	1.000000
s_is_living_with_mother	-0.004612	0.016656
s_has_repeated	0.040699	-0.118357
s_frequency_of_skips	-0.032429	0.054848
s_homeworks_day_per_week	-0.031404	0.043845
s_extent_of_teachers_affinity	-0.028543	0.028999
s_frequency_of_computer_usage	0.019260	-0.054042
s_frequency_of_internet_usage	0.084017	-0.026979
s_frequency_of_work_with_teachers	-0.013262	0.012553
s_frequency_of_materials_in_class	0.000793	-0.003437
s_frequency_of_evaluations	-0.028531	0.030408
s_extent_of_teacher_performance	-0.044767	0.045375

s_extent_of_class_env	-0.055349	0.029812
s_extent_of_classmates_affinity	-0.055421	0.055133
s_extent_of_school_satisfaction	-0.054659	0.057673
level_MAT	0.015317	0.094234
level_LEN	-0.144875	0.101092

	s_is_living_with_mother	s_has_repeated	\
s_gender	-0.004612	0.040699	
s_birth_year	0.016656	-0.118357	
s_is_living_with_mother	1.000000	-0.026872	
s_has_repeated	-0.026872	1.000000	
s_frequency_of_skips	0.007770	-0.076749	
s_homeworks_day_per_week	-0.000438	-0.030218	
s_extent_of_teachers_affinity	-0.002969	-0.024342	
s_frequency_of_computer_usage	-0.006162	0.023778	
s_frequency_of_internet_usage	-0.004206	0.007180	
s_frequency_of_work_with_teachers	0.002344	0.022051	
s_frequency_of_materials_in_class	-0.000594	0.020193	
s_frequency_of_evaluations	-0.000971	-0.017233	
s_extent_of_teacher_performance	-0.000717	-0.030589	
s_extent_of_class_env	0.000941	-0.021100	
s_extent_of_classmates_affinity	0.014985	-0.049876	
s_extent_of_school_satisfaction	0.007443	-0.054415	
level_MAT	0.022023	-0.164473	
level_LEN	0.036011	-0.201384	

	s_frequency_of_skips	\
s_gender	-0.032429	
s_birth_year	0.054848	
s_is_living_with_mother	0.007770	
s_has_repeated	-0.076749	
s_frequency_of_skips	1.000000	
s_homeworks_day_per_week	0.083523	
s_extent_of_teachers_affinity	0.057898	
s_frequency_of_computer_usage	-0.077787	
s_frequency_of_internet_usage	-0.052809	
s_frequency_of_work_with_teachers	-0.003545	
s_frequency_of_materials_in_class	-0.018838	
s_frequency_of_evaluations	0.045844	
s_extent_of_teacher_performance	0.077858	
s_extent_of_class_env	0.054547	
s_extent_of_classmates_affinity	0.079323	
s_extent_of_school_satisfaction	0.068743	
level_MAT	0.122328	
level_LEN	0.135020	

	s_homeworks_day_per_week	\
s_gender	-0.031404	

s_birth_year	0.043845
s_is_living_with_mother	-0.000438
s_has_repeated	-0.030218
s_frequency_of_skips	0.083523
s_homeworks_day_per_week	1.000000
s_extent_of_teachers_affinity	0.042352
s_frequency_of_computer_usage	-0.001209
s_frequency_of_internet_usage	0.020738
s_frequency_of_work_with_teachers	0.110804
s_frequency_of_materials_in_class	0.059834
s_frequency_of_evaluations	0.091831
s_extent_of_teacher_performance	0.057171
s_extent_of_class_env	0.046462
s_extent_of_classmates_affinity	0.042416
s_extent_of_school_satisfaction	0.049517
level_MAT	0.058022
level_LEN	0.069309

	s_extent_of_teachers_affinity \
s_gender	-0.028543
s_birth_year	0.028999
s_is_living_with_mother	-0.002969
s_has_repeated	-0.024342
s_frequency_of_skips	0.057898
s_homeworks_day_per_week	0.042352
s_extent_of_teachers_affinity	1.000000
s_frequency_of_computer_usage	-0.039669
s_frequency_of_internet_usage	-0.008730
s_frequency_of_work_with_teachers	0.161501
s_frequency_of_materials_in_class	0.094192
s_frequency_of_evaluations	0.248668
s_extent_of_teacher_performance	0.469525
s_extent_of_class_env	0.409045
s_extent_of_classmates_affinity	0.308001
s_extent_of_school_satisfaction	0.370532
level_MAT	0.093792
level_LEN	0.116560

	s_frequency_of_computer_usage \
s_gender	0.019260
s_birth_year	-0.054042
s_is_living_with_mother	-0.006162
s_has_repeated	0.023778
s_frequency_of_skips	-0.077787
s_homeworks_day_per_week	-0.001209
s_extent_of_teachers_affinity	-0.039669
s_frequency_of_computer_usage	1.000000
s_frequency_of_internet_usage	0.371735

s_frequency_of_work_with_teachers	0.120296
s_frequency_of_materials_in_class	0.215274
s_frequency_of_evaluations	0.064127
s_extent_of_teacher_performance	-0.011863
s_extent_of_class_env	0.014833
s_extent_of_classmates_affinity	-0.052212
s_extent_of_school_satisfaction	-0.051609
level_MAT	-0.072532
level_LEN	-0.075436

	s_frequency_of_internet_usage \
s_gender	0.084017
s_birth_year	-0.026979
s_is_living_with_mother	-0.004206
s_has_repeated	0.007180
s_frequency_of_skips	-0.052809
s_homeworks_day_per_week	0.020738
s_extent_of_teachers_affinity	-0.008730
s_frequency_of_computer_usage	0.371735
s_frequency_of_internet_usage	1.000000
s_frequency_of_work_with_teachers	0.184811
s_frequency_of_materials_in_class	0.215844
s_frequency_of_evaluations	0.097839
s_extent_of_teacher_performance	0.020481
s_extent_of_class_env	0.022687
s_extent_of_classmates_affinity	-0.043809
s_extent_of_school_satisfaction	-0.039423
level_MAT	-0.036665
level_LEN	-0.037401

	s_frequency_of_work_with_teachers \
s_gender	-0.013262
s_birth_year	0.012553
s_is_living_with_mother	0.002344
s_has_repeated	0.022051
s_frequency_of_skips	-0.003545
s_homeworks_day_per_week	0.110804
s_extent_of_teachers_affinity	0.161501
s_frequency_of_computer_usage	0.120296
s_frequency_of_internet_usage	0.184811
s_frequency_of_work_with_teachers	1.000000
s_frequency_of_materials_in_class	0.407241
s_frequency_of_evaluations	0.459210
s_extent_of_teacher_performance	0.284243
s_extent_of_class_env	0.253826
s_extent_of_classmates_affinity	0.115582
s_extent_of_school_satisfaction	0.140080
level_MAT	-0.038320

level_LEN	-0.016607
-----------	-----------

	s_frequency_of_materials_in_class \
s_gender	0.000793
s_birth_year	-0.003437
s_is_living_with_mother	-0.000594
s_has_repeated	0.020193
s_frequency_of_skips	-0.018838
s_homeworks_day_per_week	0.059834
s_extent_of_teachers_affinity	0.094192
s_frequency_of_computer_usage	0.215274
s_frequency_of_internet_usage	0.215844
s_frequency_of_work_with_teachers	0.407241
s_frequency_of_materials_in_class	1.000000
s_frequency_of_evaluations	0.346366
s_extent_of_teacher_performance	0.192643
s_extent_of_class_env	0.198315
s_extent_of_classmates_affinity	0.075477
s_extent_of_school_satisfaction	0.088556
level_MAT	-0.066594
level_LEN	-0.048885

	s_frequency_of_evaluations \
s_gender	-0.028531
s_birth_year	0.030408
s_is_living_with_mother	-0.000971
s_has_repeated	-0.017233
s_frequency_of_skips	0.045844
s_homeworks_day_per_week	0.091831
s_extent_of_teachers_affinity	0.248668
s_frequency_of_computer_usage	0.064127
s_frequency_of_internet_usage	0.097839
s_frequency_of_work_with_teachers	0.459210
s_frequency_of_materials_in_class	0.346366
s_frequency_of_evaluations	1.000000
s_extent_of_teacher_performance	0.405473
s_extent_of_class_env	0.353524
s_extent_of_classmates_affinity	0.208706
s_extent_of_school_satisfaction	0.246491
level_MAT	0.033004
level_LEN	0.062949

	s_extent_of_teacher_performance \
s_gender	-0.044767
s_birth_year	0.045375
s_is_living_with_mother	-0.000717
s_has_repeated	-0.030589
s_frequency_of_skips	0.077858

s_homeworks_day_per_week	0.057171
s_extent_of_teachers_affinity	0.469525
s_frequency_of_computer_usage	-0.011863
s_frequency_of_internet_usage	0.020481
s_frequency_of_work_with_teachers	0.284243
s_frequency_of_materials_in_class	0.192643
s_frequency_of_evaluations	0.405473
s_extent_of_teacher_performance	1.000000
s_extent_of_class_env	0.699774
s_extent_of_classmates_affinity	0.455324
s_extent_of_school_satisfaction	0.572676
level_MAT	0.110412
level_LEN	0.138279

	s_extent_of_class_env \
s_gender	-0.055349
s_birth_year	0.029812
s_is_living_with_mother	0.000941
s_has_repeated	-0.021100
s_frequency_of_skips	0.054547
s_homeworks_day_per_week	0.046462
s_extent_of_teachers_affinity	0.409045
s_frequency_of_computer_usage	0.014833
s_frequency_of_internet_usage	0.022687
s_frequency_of_work_with_teachers	0.253826
s_frequency_of_materials_in_class	0.198315
s_frequency_of_evaluations	0.353524
s_extent_of_teacher_performance	0.699774
s_extent_of_class_env	1.000000
s_extent_of_classmates_affinity	0.512753
s_extent_of_school_satisfaction	0.603268
level_MAT	0.059059
level_LEN	0.098072

	s_extent_of_classmates_affinity \
s_gender	-0.055421
s_birth_year	0.055133
s_is_living_with_mother	0.014985
s_has_repeated	-0.049876
s_frequency_of_skips	0.079323
s_homeworks_day_per_week	0.042416
s_extent_of_teachers_affinity	0.308001
s_frequency_of_computer_usage	-0.052212
s_frequency_of_internet_usage	-0.043809
s_frequency_of_work_with_teachers	0.115582
s_frequency_of_materials_in_class	0.075477
s_frequency_of_evaluations	0.208706
s_extent_of_teacher_performance	0.455324

s_extent_of_class_env	0.512753
s_extent_of_classmates_affinity	1.000000
s_extent_of_school_satisfaction	0.471544
level_MAT	0.132252
level_LEN	0.166773

	s_extent_of_school_satisfaction	level_MAT	\
s_gender	-0.054659	0.015317	
s_birth_year	0.057673	0.094234	
s_is_living_with_mother	0.007443	0.022023	
s_has_repeated	-0.054415	-0.164473	
s_frequency_of_skips	0.068743	0.122328	
s_homeworks_day_per_week	0.049517	0.058022	
s_extent_of_teachers_affinity	0.370532	0.093792	
s_frequency_of_computer_usage	-0.051609	-0.072532	
s_frequency_of_internet_usage	-0.039423	-0.036665	
s_frequency_of_work_with_teachers	0.140080	-0.038320	
s_frequency_of_materials_in_class	0.088556	-0.066594	
s_frequency_of_evaluations	0.246491	0.033004	
s_extent_of_teacher_performance	0.572676	0.110412	
s_extent_of_class_env	0.603268	0.059059	
s_extent_of_classmates_affinity	0.471544	0.132252	
s_extent_of_school_satisfaction	1.000000	0.123259	
level_MAT	0.123259	1.000000	
level_LEN	0.146688	0.562532	

	level_LEN
s_gender	-0.144875
s_birth_year	0.101092
s_is_living_with_mother	0.036011
s_has_repeated	-0.201384
s_frequency_of_skips	0.135020
s_homeworks_day_per_week	0.069309
s_extent_of_teachers_affinity	0.116560
s_frequency_of_computer_usage	-0.075436
s_frequency_of_internet_usage	-0.037401
s_frequency_of_work_with_teachers	-0.016607
s_frequency_of_materials_in_class	-0.048885
s_frequency_of_evaluations	0.062949
s_extent_of_teacher_performance	0.138279
s_extent_of_class_env	0.098072
s_extent_of_classmates_affinity	0.166773
s_extent_of_school_satisfaction	0.146688
level_MAT	0.562532
level_LEN	1.000000

5.5 Family Questionnaire

```
[83]: df_sel_families, correlation_matrix_sel_families =  
      ↪compute_exploratory_analysis(df=df_sel, column_pattern="f_",  
      ↪missing_threshold=0.3, annot=False)
```

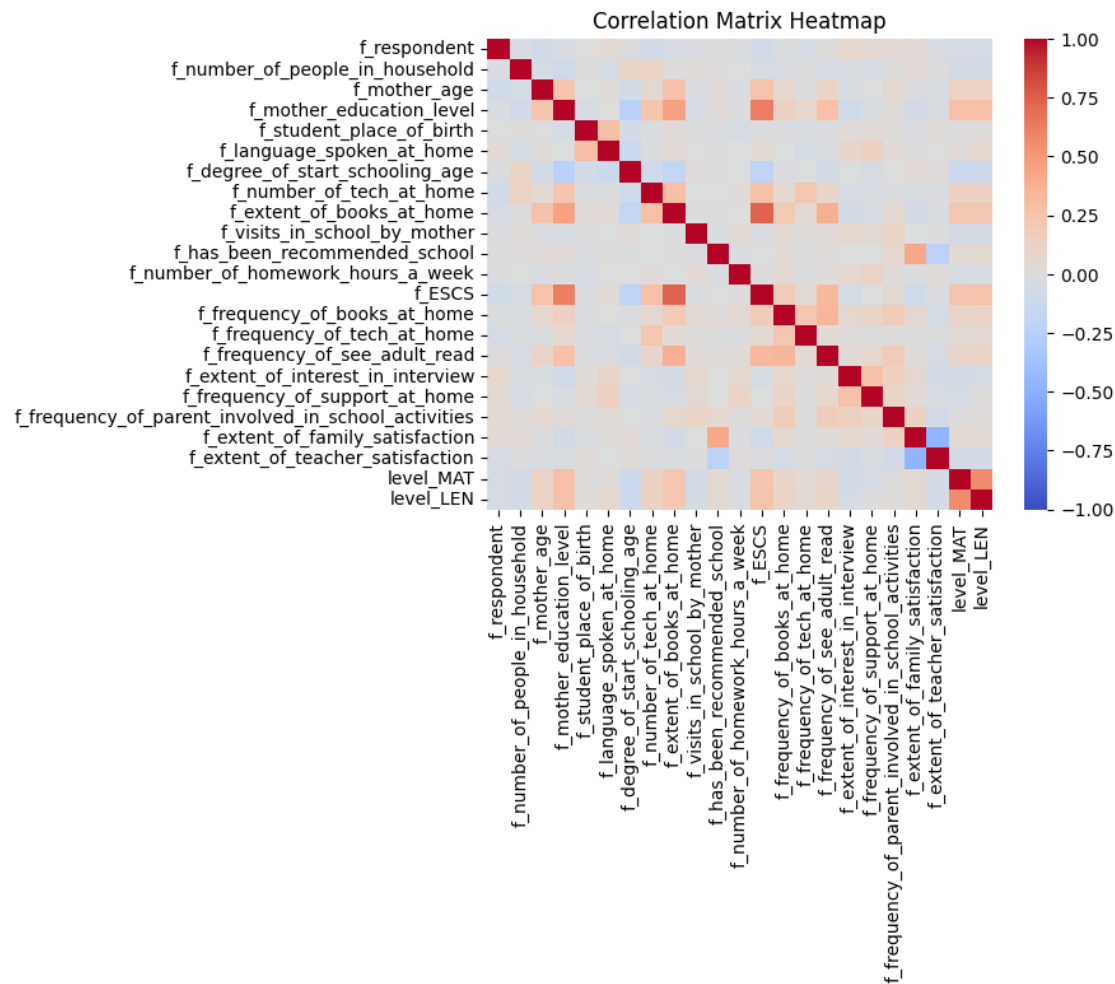
GOOD FEATURES: 23

f_respondent	0.288217
f_number_of_people_in_household	0.299562
f_mother_age	0.281534
f_mother_education_level	0.291082
f_student_place_of_birth	0.287768
f_language_spoken_at_home	0.287937
f_degree_of_start_schooling_age	0.293665
f_number_of_tech_at_home	0.281984
f_extent_of_books_at_home	0.274907
f_visits_in_school_by_mother	0.289790
f_has_been_recommended_school	0.282882
f_number_of_homework_hours_a_week	0.278445
f_ESCS	0.275357
f_frequency_of_books_at_home	0.290295
f_frequency_of_tech_at_home	0.291531
f_frequency_of_see_adult_read	0.274739
f_extent_of_interest_in_interview	0.278052
f_frequency_of_support_at_home	0.275862
f_frequency_of_parent_involved_in_school_activities	0.278502
f_extent_of_family_satisfaction	0.277547
f_extent_of_teacher_satisfaction	0.277940
level_MAT	0.000000
level_LEN	0.009772
dtype: float64	

BAD FEATURES

f_father_age	0.331405
f_father_education_level	0.348085
f_mother_employment_status	0.304504
f_father_employment_status	0.358194
f_mother_place_of_birth	0.327530
f_father_place_of_birth	0.369145
f_degree_of_years_in_spanish_education_system	0.690554
f_visits_in_school_by_father	0.376783
f_parental_education_expectations	0.370156
f_mother_occupation	0.302145
f_father_occupation	0.352128
f_number_of_children_in_household	1.000000
f_type_of_family_unit	1.000000
f_monthly_household_income	1.000000
level_ING	1.000000

dtype: float64



	f_respondent	f_number_of_people_in_household	f_mother_age	\
id_questionnaire				
1	MOTHER		3.0	43.0
2	NaN		NaN	NaN
3	MOTHER		4.0	45.0
4	MOTHER		4.0	39.0
5	FATHER		4.0	39.0
...
20426	FATHER		3.0	31.0
20427	MOTHER		3.0	37.0
20428	NaN		NaN	NaN
20429	NaN		NaN	NaN
20430	NaN		NaN	NaN

id_questionnaire	f_mother_education_level	f_student_place_of_birth \
1	7.0	CANARY_ISLANDS
2	NaN	NaN
3	3.0	CANARY_ISLANDS
4	5.0	CANARY_ISLANDS
5	9.0	CANARY_ISLANDS
...
20426	3.0	CANARY_ISLANDS
20427	5.0	CANARY_ISLANDS
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

id_questionnaire	f_language_spoken_at_home	f_degree_of_start_schooling_age \
1	SPANISH	1.0
2	NaN	NaN
3	SPANISH	1.0
4	SPANISH	1.0
5	SPANISH	1.0
...
20426	SPANISH	2.0
20427	SPANISH	2.0
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

id_questionnaire	f_number_of_tech_at_home	f_extent_of_books_at_home \
1	3.0	0.75
2	NaN	NaN
3	4.0	0.25
4	3.0	0.75
5	7.0	1.00
...
20426	4.0	0.00
20427	5.0	0.50
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

id_questionnaire	f_visits_in_school_by_mother ... \
1	ONCE_PER_WEEK ...
2	NaN ...
3	ONCE_PER_MONTH ...
4	ONCE_PER_WEEK ...

5	ONCE_PER_MONTH	...
...
20426	NaN	...
20427	SOMETIMES	...
20428	NaN	...
20429	NaN	...
20430	NaN	...

	f_frequency_of_books_at_home	f_frequency_of_tech_at_home \
id_questionnaire		
1	1.000000	1.000000
2	NaN	NaN
3	0.444444	0.333333
4	0.333333	0.333333
5	0.555556	0.500000
...
20426	0.333333	1.000000
20427	0.444444	0.666667
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

	f_frequency_of_see_adult_read \
id_questionnaire	
1	0.625
2	NaN
3	0.250
4	0.500
5	0.625
...	...
20426	0.000
20427	0.500
20428	NaN
20429	NaN
20430	NaN

	f_extent_of_interest_in_interview \
id_questionnaire	
1	1.000000
2	NaN
3	1.000000
4	1.000000
5	NaN
...	...
20426	0.444444
20427	0.500000
20428	NaN
20429	NaN

20430	NaN
-------	-----

	f_frequency_of_support_at_home \
id_questionnaire	
1	1.000000
2	NaN
3	1.000000
4	1.000000
5	0.888889
...	...
20426	0.800000
20427	0.888889
20428	NaN
20429	NaN
20430	NaN

	f_frequency_of_parent_involved_in_school_activities \
id_questionnaire	
1	0.333333
2	NaN
3	0.166667
4	0.250000
5	0.250000
...	...
20426	0.000000
20427	0.000000
20428	NaN
20429	NaN
20430	NaN

	f_extent_of_family_satisfaction \
id_questionnaire	
1	0.700000
2	NaN
3	0.291667
4	1.000000
5	0.833333
...	...
20426	0.791667
20427	0.291667
20428	NaN
20429	NaN
20430	NaN

	f_extent_of_teacher_satisfaction	level_MAT	level_LEN
id_questionnaire			
1	0.15	3.0	3.0
2	NaN	1.0	1.0

3	0.50	1.0	3.0
4	0.20	2.0	2.0
5	0.35	4.0	3.0
...
20426	0.25	1.0	1.0
20427	0.40	2.0	3.0
20428	NaN	2.0	2.0
20429	NaN	3.0	3.0
20430	NaN	2.0	3.0

[17806 rows x 23 columns]

	f_respondent \
f_respondent	1.000000
f_number_of_people_in_household	-0.027957
f_mother_age	-0.081211
f_mother_education_level	-0.027144
f_student_place_of_birth	-0.006793
f_language_spoken_at_home	0.041598
f_degree_of_start_schooling_age	-0.014239
f_number_of_tech_at_home	-0.084977
f_extent_of_books_at_home	-0.038491
f_visits_in_school_by_mother	-0.017420
f_has_been_recommended_school	0.002542
f_number_of_homework_hours_a_week	-0.015882
f_ESCS	-0.082319
f_frequency_of_books_at_home	-0.008995
f_frequency_of_tech_at_home	-0.053953
f_frequency_of_see_adult_read	0.015517
f_extent_of_interest_in_interview	0.074843
f_frequency_of_support_at_home	0.054338
f_frequency_of_parent_involved_in_school_activi...	0.046602
f_extent_of_family_satisfaction	0.037620
f_extent_of_teacher_satisfaction	-0.032144
level_MAT	-0.040162
level_LEN	-0.044839

f_number_of_people_in_household \	
f_respondent	-0.
027957	
f_number_of_people_in_household	1.
000000	
f_mother_age	-0.
053139	
f_mother_education_level	-0.
088100	

f_student_place_of_birth	0.
↪010731	
f_language_spoken_at_home	-0.
↪052196	
f_degree_of_start_schooling_age	0.
↪108799	
f_number_of_tech_at_home	0.
↪123348	
f_extent_of_books_at_home	0.
↪011982	
f_visits_in_school_by_mother	0.
↪016202	
f_has_been_recommended_school	0.
↪014522	
f_number_of_homework_hours_a_week	-0.
↪005058	
f_ESCS	-0.
↪040514	
f_frequency_of_books_at_home	-0.
↪025456	
f_frequency_of_tech_at_home	-0.
↪010474	
f_frequency_of_see_adult_read	-0.
↪031906	
f_extent_of_interest_in_interview	-0.
↪026799	
f_frequency_of_support_at_home	-0.
↪042393	
f_frequency_of_parent_involved_in_school_activi...	0.009101
f_extent_of_family_satisfaction	0.
↪031290	
f_extent_of_teacher_satisfaction	-0.
↪011043	
level_MAT	-0.
↪067898	
level_LEN	-0.
↪076081	

	f_mother_age \
f_respondent	-0.081211
f_number_of_people_in_household	-0.053139
f_mother_age	1.000000
f_mother_education_level	0.250886
f_student_place_of_birth	-0.002140
f_language_spoken_at_home	0.022580
f_degree_of_start_schooling_age	-0.072617

f_number_of_tech_at_home	0.058864
f_extent_of_books_at_home	0.268057
f_visits_in_school_by_mother	-0.015672
f_has_been_recommended_school	0.023831
f_number_of_homework_hours_a_week	-0.019828
f_ESCS	0.261448
f_frequency_of_books_at_home	0.072558
f_frequency_of_tech_at_home	-0.011576
f_frequency_of_see_adult_read	0.110363
f_extent_of_interest_in_interview	-0.029604
f_frequency_of_support_at_home	-0.002376
f_frequency_of_parent_involved_in_school_activi...	0.060719
f_extent_of_family_satisfaction	-0.013724
f_extent_of_teacher_satisfaction	-0.024134
level_MAT	0.126018
level_LEN	0.129967

	f_mother_education_level \
f_respondent	-0.027144
f_number_of_people_in_household	-0.088100
f_mother_age	0.250886
f_mother_education_level	1.000000
f_student_place_of_birth	-0.049596
f_language_spoken_at_home	-0.002979
f_degree_of_start_schooling_age	-0.248467
f_number_of_tech_at_home	0.254568
f_extent_of_books_at_home	0.457494
f_visits_in_school_by_mother	-0.042295
f_has_been_recommended_school	0.027079
f_number_of_homework_hours_a_week	-0.008756
f_ESCS	0.627503
f_frequency_of_books_at_home	0.146603
f_frequency_of_tech_at_home	0.074593
f_frequency_of_see_adult_read	0.284140
f_extent_of_interest_in_interview	-0.089370
f_frequency_of_support_at_home	-0.030900
f_frequency_of_parent_involved_in_school_activi...	0.020266
f_extent_of_family_satisfaction	-0.065329
f_extent_of_teacher_satisfaction	-0.018038
level_MAT	0.284138
level_LEN	0.283595

	f_student_place_of_birth \
f_respondent	-0.006793
f_number_of_people_in_household	0.010731
f_mother_age	-0.002140
f_mother_education_level	-0.049596
f_student_place_of_birth	1.000000

f_language_spoken_at_home	0.289808
f_degree_of_start_schooling_age	-0.072643
f_number_of_tech_at_home	-0.015643
f_extent_of_books_at_home	0.025610
f_visits_in_school_by_mother	-0.012835
f_has_been_recommended_school	-0.015724
f_number_of_homework_hours_a_week	-0.042928
f_ESCS	-0.012139
f_frequency_of_books_at_home	-0.017911
f_frequency_of_tech_at_home	-0.020781
f_frequency_of_see_adult_read	-0.012698
f_extent_of_interest_in_interview	0.026882
f_frequency_of_support_at_home	0.044377
f_frequency_of_parent_involved_in_school_activi...	0.011125
f_extent_of_family_satisfaction	-0.020163
f_extent_of_teacher_satisfaction	0.021383
level_MAT	0.000980
level_LEN	0.020267

	f_language_spoken_at_home \
f_respondent	0.041598
f_number_of_people_in_household	-0.052196
f_mother_age	0.022580
f_mother_education_level	-0.002979
f_student_place_of_birth	0.289808
f_language_spoken_at_home	1.000000
f_degree_of_start_schooling_age	-0.112067
f_number_of_tech_at_home	-0.008967
f_extent_of_books_at_home	0.035558
f_visits_in_school_by_mother	-0.004309
f_has_been_recommended_school	0.015072
f_number_of_homework_hours_a_week	-0.001131
f_ESCS	0.045499
f_frequency_of_books_at_home	-0.006676
f_frequency_of_tech_at_home	-0.049509
f_frequency_of_see_adult_read	-0.021653
f_extent_of_interest_in_interview	0.091814
f_frequency_of_support_at_home	0.141668
f_frequency_of_parent_involved_in_school_activi...	0.035891
f_extent_of_family_satisfaction	0.014853
f_extent_of_teacher_satisfaction	-0.003008
level_MAT	0.025272
level_LEN	0.066332

↩f_degree_of_start_schooling_age \

f_respondent	-0.
↩014239	

f_number_of_people_in_household	0.
↪108799	
f_mother_age	-0.
↪072617	
f_mother_education_level	-0.
↪248467	
f_student_place_of_birth	-0.
↪072643	
f_language_spoken_at_home	-0.
↪112067	
f_degree_of_start_schooling_age	1.
↪000000	
f_number_of_tech_at_home	-0.
↪098786	
f_extent_of_books_at_home	-0.
↪170949	
f_visits_in_school_by_mother	0.
↪027965	
f_has_been_recommended_school	-0.
↪010893	
f_number_of_homework_hours_a_week	-0.
↪007506	
f_ESCS	-0.
↪210289	
f_frequency_of_books_at_home	-0.
↪021553	
f_frequency_of_tech_at_home	-0.
↪002300	
f_frequency_of_see_adult_read	-0.
↪076068	
f_extent_of_interest_in_interview	-0.
↪004336	
f_frequency_of_support_at_home	-0.
↪034542	
f_frequency_of_parent_involved_in_school_activi...	-0.006201
f_extent_of_family_satisfaction	0.
↪036661	
f_extent_of_teacher_satisfaction	0.
↪015961	
level_MAT	-0.
↪115373	
level_LEN	-0.
↪108522	
	f_number_of_tech_at_home \
f_respondent	-0.084977

f_number_of_people_in_household	0.123348
f_mother_age	0.058864
f_mother_education_level	0.254568
f_student_place_of_birth	-0.015643
f_language_spoken_at_home	-0.008967
f_degree_of_start_schooling_age	-0.098786
f_number_of_tech_at_home	1.000000
f_extent_of_books_at_home	0.283857
f_visits_in_school_by_mother	-0.007627
f_has_been_recommended_school	-0.001399
f_number_of_homework_hours_a_week	0.020601
f_ESCS	0.278867
f_frequency_of_books_at_home	0.063726
f_frequency_of_tech_at_home	0.241223
f_frequency_of_see_adult_read	0.100015
f_extent_of_interest_in_interview	-0.036621
f_frequency_of_support_at_home	-0.026188
f_frequency_of_parent_involved_in_school_activi...	-0.011387
f_extent_of_family_satisfaction	-0.030832
f_extent_of_teacher_satisfaction	-0.012458
level_MAT	0.143473
level_LEN	0.146762

	f_extent_of_books_at_home \
f_respondent	-0.038491
f_number_of_people_in_household	0.011982
f_mother_age	0.268057
f_mother_education_level	0.457494
f_student_place_of_birth	0.025610
f_language_spoken_at_home	0.035558
f_degree_of_start_schooling_age	-0.170949
f_number_of_tech_at_home	0.283857
f_extent_of_books_at_home	1.000000
f_visits_in_school_by_mother	-0.011430
f_has_been_recommended_school	-0.007821
f_number_of_homework_hours_a_week	-0.005190
f_ESCS	0.735864
f_frequency_of_books_at_home	0.209352
f_frequency_of_tech_at_home	0.043936
f_frequency_of_see_adult_read	0.385506
f_extent_of_interest_in_interview	-0.062068
f_frequency_of_support_at_home	-0.033528
f_frequency_of_parent_involved_in_school_activi...	0.066643
f_extent_of_family_satisfaction	-0.066426
f_extent_of_teacher_satisfaction	-0.022848
level_MAT	0.229863
level_LEN	0.231046

↪ \

f_respondent	-0.017420
f_number_of_people_in_household	0.016202
f_mother_age	-0.015672
f_mother_education_level	-0.042295
f_student_place_of_birth	-0.012835
f_language_spoken_at_home	-0.004309
f_degree_of_start_schooling_age	0.027965
f_number_of_tech_at_home	-0.007627
f_extent_of_books_at_home	-0.011430
f_visits_in_school_by_mother	1.000000
f_has_been_recommended_school	-0.022420
f_number_of_homework_hours_a_week	0.041970
f_ESCS	-0.026805
f_frequency_of_books_at_home	0.048104
f_frequency_of_tech_at_home	0.012176
f_frequency_of_see_adult_read	0.030557
f_extent_of_interest_in_interview	0.063744
f_frequency_of_support_at_home	0.054227
f_frequency_of_parent_involved_in_school_activi...	0.101717
f_extent_of_family_satisfaction	-0.002028
f_extent_of_teacher_satisfaction	-0.008348
level_MAT	-0.056037
level_LEN	-0.055724

	...	\
f_respondent	...	
f_number_of_people_in_household	...	
f_mother_age	...	
f_mother_education_level	...	
f_student_place_of_birth	...	
f_language_spoken_at_home	...	
f_degree_of_start_schooling_age	...	
f_number_of_tech_at_home	...	
f_extent_of_books_at_home	...	
f_visits_in_school_by_mother	...	
f_has_been_recommended_school	...	
f_number_of_homework_hours_a_week	...	
f_ESCS	...	
f_frequency_of_books_at_home	...	
f_frequency_of_tech_at_home	...	
f_frequency_of_see_adult_read	...	
f_extent_of_interest_in_interview	...	
f_frequency_of_support_at_home	...	
f_frequency_of_parent_involved_in_school_activi...	...	
f_extent_of_family_satisfaction	...	

f_extent_of_teacher_satisfaction	...
level_MAT	...
level_LEN	...

f_frequency_of_books_at_home

\	
f_respondent	-0.008995
f_number_of_people_in_household	-0.025456
f_mother_age	0.072558
f_mother_education_level	0.146603
f_student_place_of_birth	-0.017911
f_language_spoken_at_home	-0.006676
f_degree_of_start_schooling_age	-0.021553
f_number_of_tech_at_home	0.063726
f_extent_of_books_at_home	0.209352
f_visits_in_school_by_mother	0.048104
f_has_been_recommended_school	0.023591
f_number_of_homework_hours_a_week	0.057714
f_ESCS	0.175323
f_frequency_of_books_at_home	1.000000
f_frequency_of_tech_at_home	0.237772
f_frequency_of_see_adult_read	0.345352
f_extent_of_interest_in_interview	0.073173
f_frequency_of_support_at_home	0.106625
f_frequency_of_parent_involved_in_school_activi...	0.179154
f_extent_of_family_satisfaction	0.057873
f_extent_of_teacher_satisfaction	-0.049062
level_MAT	0.097342
level_LEN	0.113689

f_frequency_of_tech_at_home

\	
f_respondent	-0.053953
f_number_of_people_in_household	-0.010474
f_mother_age	-0.011576
f_mother_education_level	0.074593
f_student_place_of_birth	-0.020781
f_language_spoken_at_home	-0.049509
f_degree_of_start_schooling_age	-0.002300
f_number_of_tech_at_home	0.241223
f_extent_of_books_at_home	0.043936
f_visits_in_school_by_mother	0.012176
f_has_been_recommended_school	0.004843
f_number_of_homework_hours_a_week	-0.010289
f_ESCS	0.047837
f_frequency_of_books_at_home	0.237772
f_frequency_of_tech_at_home	1.000000

f_frequency_of_see_adult_read	0.040688
f_extent_of_interest_in_interview	0.011785
f_frequency_of_support_at_home	0.002813
f_frequency_of_parent_involved_in_school_activi...	0.010881
f_extent_of_family_satisfaction	0.007259
f_extent_of_teacher_satisfaction	-0.010616
level_MAT	0.043599
level_LEN	0.048300

↪f_frequency_of_see_adult_read \

f_respondent	0.015517
f_number_of_people_in_household	-0.031906
f_mother_age	0.110363
f_mother_education_level	0.284140
f_student_place_of_birth	-0.012698
f_language_spoken_at_home	-0.021653
f_degree_of_start_schooling_age	-0.076068
f_number_of_tech_at_home	0.100015
f_extent_of_books_at_home	0.385506
f_visits_in_school_by_mother	0.030557
f_has_been_recommended_school	0.017556
f_number_of_homework_hours_a_week	0.024970
f_ESCS	0.330894
f_frequency_of_books_at_home	0.345352
f_frequency_of_tech_at_home	0.040688
f_frequency_of_see_adult_read	1.000000
f_extent_of_interest_in_interview	0.023952
f_frequency_of_support_at_home	0.077578
f_frequency_of_parent_involved_in_school_activi...	0.177247
f_extent_of_family_satisfaction	0.018249
f_extent_of_teacher_satisfaction	-0.037240
level_MAT	0.088369
level_LEN	0.106418

↪f_extent_of_interest_in_interview \

f_respondent	0.
↪074843	
f_number_of_people_in_household	-0.
↪026799	
f_mother_age	-0.
↪029604	
f_mother_education_level	-0.
↪089370	
f_student_place_of_birth	0.
↪026882	

f_language_spoken_at_home	0.
↵091814	
f_degree_of_start_schooling_age	-0.
↵004336	
f_number_of_tech_at_home	-0.
↵036621	
f_extent_of_books_at_home	-0.
↵062068	
f_visits_in_school_by_mother	0.
↵063744	
f_has_been_recommended_school	-0.
↵000693	
f_number_of_homework_hours_a_week	0.
↵068049	
f_ESCS	-0.
↵052693	
f_frequency_of_books_at_home	0.
↵073173	
f_frequency_of_tech_at_home	0.
↵011785	
f_frequency_of_see_adult_read	0.
↵023952	
f_extent_of_interest_in_interview	1.
↵000000	
f_frequency_of_support_at_home	0.
↵266364	
f_frequency_of_parent_involved_in_school_activi...	0.
↵139872	
f_extent_of_family_satisfaction	0.
↵068859	
f_extent_of_teacher_satisfaction	-0.
↵060370	
level_MAT	-0.
↵072618	
level_LEN	-0.
↵052626	
□	
↵f_frequency_of_support_at_home \	
f_respondent	0.
↵054338	
f_number_of_people_in_household	-0.
↵042393	
f_mother_age	-0.
↵002376	

f_mother_education_level	-0.
↪030900	
f_student_place_of_birth	0.
↪044377	
f_language_spoken_at_home	0.
↪141668	
f_degree_of_start_schooling_age	-0.
↪034542	
f_number_of_tech_at_home	-0.
↪026188	
f_extent_of_books_at_home	-0.
↪033528	
f_visits_in_school_by_mother	0.
↪054227	
f_has_been_recommended_school	-0.
↪004020	
f_number_of_homework_hours_a_week	0.
↪119507	
f_ESCS	-0.
↪002452	
f_frequency_of_books_at_home	0.
↪106625	
f_frequency_of_tech_at_home	0.
↪002813	
f_frequency_of_see_adult_read	0.
↪077578	
f_extent_of_interest_in_interview	0.
↪266364	
f_frequency_of_support_at_home	1.
↪000000	
f_frequency_of_parent_involved_in_school_activi...	0.104213
f_extent_of_family_satisfaction	0.
↪037971	
f_extent_of_teacher_satisfaction	-0.
↪026951	
level_MAT	-0.
↪050041	
level_LEN	0.
↪001853	
↪f_frequency_of_parent_involved_in_school_activities \	
f_respondent	
↪0.046602	
f_number_of_people_in_household	
↪0.009101	

f_mother_age		□
↪	0.060719	
f_mother_education_level		□
↪	0.020266	
f_student_place_of_birth		□
↪	0.011125	
f_language_spoken_at_home		□
↪	0.035891	
f_degree_of_start_schooling_age		□
↪	-0.006201	
f_number_of_tech_at_home		□
↪	-0.011387	
f_extent_of_books_at_home		□
↪	0.066643	
f_visits_in_school_by_mother		□
↪	0.101717	
f_has_been_recommended_school		□
↪	0.073856	
f_number_of_homework_hours_a_week		□
↪	0.015347	
f_ESCS		□
↪	0.047424	
f_frequency_of_books_at_home		□
↪	0.179154	
f_frequency_of_tech_at_home		□
↪	0.010881	
f_frequency_of_see_adult_read		□
↪	0.177247	
f_extent_of_interest_in_interview		□
↪	0.139872	
f_frequency_of_support_at_home		□
↪	0.104213	
f_frequency_of_parent_involved_in_school_activi...		□
↪	1.000000	
f_extent_of_family_satisfaction		□
↪	0.136940	
f_extent_of_teacher_satisfaction		□
↪	-0.063903	
level_MAT		□
↪	0.002740	
level_LEN		□
↪	0.021056	
↪f_extent_of_family_satisfaction \		□

f_respondent	0.
↪037620	
f_number_of_people_in_household	0.
↪031290	
f_mother_age	-0.
↪013724	
f_mother_education_level	-0.
↪065329	
f_student_place_of_birth	-0.
↪020163	
f_language_spoken_at_home	0.
↪014853	
f_degree_of_start_schooling_age	0.
↪036661	
f_number_of_tech_at_home	-0.
↪030832	
f_extent_of_books_at_home	-0.
↪066426	
f_visits_in_school_by_mother	-0.
↪002028	
f_has_been_recommended_school	0.
↪426081	
f_number_of_homework_hours_a_week	-0.
↪019991	
f_ESCS	-0.
↪087052	
f_frequency_of_books_at_home	0.
↪057873	
f_frequency_of_tech_at_home	0.
↪007259	
f_frequency_of_see_adult_read	0.
↪018249	
f_extent_of_interest_in_interview	0.
↪068859	
f_frequency_of_support_at_home	0.
↪037971	
f_frequency_of_parent_involved_in_school_activi...	0.136940
f_extent_of_family_satisfaction	1.
↪000000	
f_extent_of_teacher_satisfaction	-0.
↪474855	
level_MAT	0.
↪050280	
level_LEN	0.
↪053399	

□

↪f_extent_of_teacher_satisfaction \	
f_respondent	-0.
↪032144	
f_number_of_people_in_household	-0.
↪011043	
f_mother_age	-0.
↪024134	
f_mother_education_level	-0.
↪018038	
f_student_place_of_birth	0.
↪021383	
f_language_spoken_at_home	-0.
↪003008	
f_degree_of_start_schooling_age	0.
↪015961	
f_number_of_tech_at_home	-0.
↪012458	
f_extent_of_books_at_home	-0.
↪022848	
f_visits_in_school_by_mother	-0.
↪008348	
f_has_been_recommended_school	-0.
↪229920	
f_number_of_homework_hours_a_week	0.
↪028860	
f_ESCS	-0.
↪007822	
f_frequency_of_books_at_home	-0.
↪049062	
f_frequency_of_tech_at_home	-0.
↪010616	
f_frequency_of_see_adult_read	-0.
↪037240	
f_extent_of_interest_in_interview	-0.
↪060370	
f_frequency_of_support_at_home	-0.
↪026951	
f_frequency_of_parent_involved_in_school_activi...	-0.
↪063903	
f_extent_of_family_satisfaction	-0.
↪474855	
f_extent_of_teacher_satisfaction	1.
↪000000	
level_MAT	-0.
↪055982	

```
level_LEN -0.067184
```

	level_MAT	level_LEN
f_respondent	-0.040162	-0.044839
f_number_of_people_in_household	-0.067898	-0.076081
f_mother_age	0.126018	0.129967
f_mother_education_level	0.284138	0.283595
f_student_place_of_birth	0.000980	0.020267
f_language_spoken_at_home	0.025272	0.066332
f_degree_of_start_schooling_age	-0.115373	-0.108522
f_number_of_tech_at_home	0.143473	0.146762
f_extent_of_books_at_home	0.229863	0.231046
f_visits_in_school_by_mother	-0.056037	-0.055724
f_has_been_recommended_school	0.050717	0.048489
f_number_of_homework_hours_a_week	-0.035409	-0.023550
f_ESCS	0.250165	0.265560
f_frequency_of_books_at_home	0.097342	0.113689
f_frequency_of_tech_at_home	0.043599	0.048300
f_frequency_of_see_adult_read	0.088369	0.106418
f_extent_of_interest_in_interview	-0.072618	-0.052626
f_frequency_of_support_at_home	-0.050041	0.001853
f_frequency_of_parent_involved_in_school_activi...	0.002740	0.021056
f_extent_of_family_satisfaction	0.050280	0.053399
f_extent_of_teacher_satisfaction	-0.055982	-0.067184
level_MAT	1.000000	0.562532
level_LEN	0.562532	1.000000

[23 rows x 23 columns]

5.6 Teacher Questionnaire

```
[84]: df_sel_teachers, correlation_matrix_sel_teachers = \
    compute_exploratory_analysis(df=df_sel, column_pattern="t_", \
    missing_threshold=0.4, annot=False)
```

GOOD FEATURES: 31

t_gender	0.273054
t_age	0.291643
t_number_of_years_as_teacher	0.276424
t_number_of_years_in_school	0.283444
t_has_taught_same_group_last_two_years	0.273110
t_number_of_students_in_group	0.275525
t_number_of_teaching_hours_per_week	0.286982
t_average_explanation_time	0.270920
t_degree_of_individual_training_incidence	0.347130
t_seat_configuration	0.273447
t_behaviour_problems_solution	0.275750

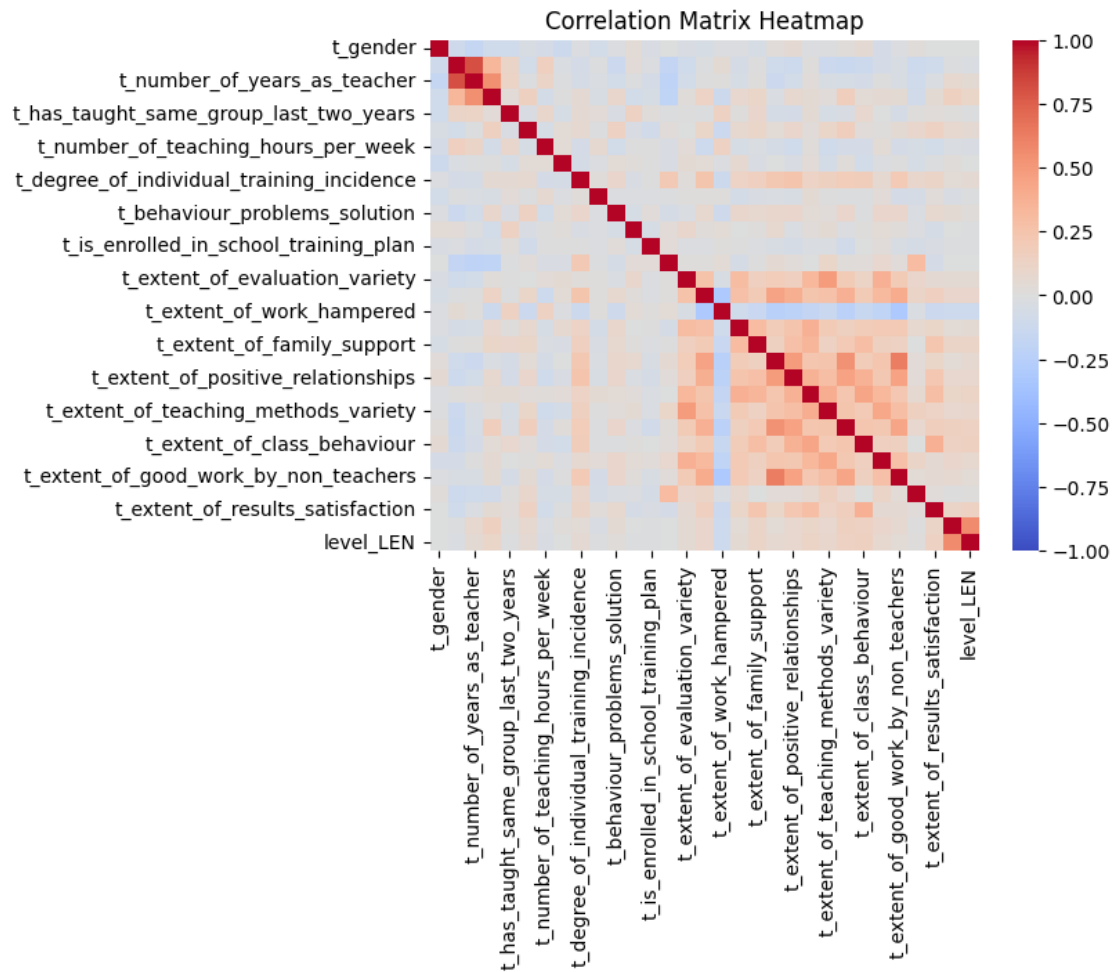
t_number_of_meetings_with_families	0.308267
t_is_enrolled_in_school_training_plan	0.273503
t_number_of_training_hours_last_six_years	0.331854
t_extent_of_evaluation_variety	0.271032
t_extent_of_pfc_incidence	0.335617
t_extent_of_work_hampered	0.274907
t_extent_of_family_interest	0.273503
t_extent_of_family_support	0.274402
t_extent_of_work_facilitated_by_management	0.277322
t_extent_of_positive_relationships	0.272436
t_extent_of_student_involvement_during_class	0.272155
t_extent_of_teaching_methods_variety	0.271032
t_extent_of_opinion_on_school	0.293890
t_extent_of_class_behaviour	0.273728
t_extent_of_resource_variety	0.271032
t_extent_of_good_work_by_non_teachers	0.283893
t_number_of_individual_training_topics	0.377457
t_extent_of_results_satisfaction	0.271987
level_MAT	0.000000
level_LEN	0.009772

dtype: float64

BAD FEATURES

t_number_of_students_disadvantaged_economic_situation	0.592216
t_number_of_foreign_students_speaking_spanish	0.471021
t_number_of_foreign_students_not_speaking_spanish	0.475458
t_number_of_training_hours_ceu_offer	1.000000
t_main_topic_of_pfc	1.000000
t_number_of_special_attention_students	0.431877
t_number_of_subjects_taught	1.000000
t_extent_of_satisfaction_job_and_school	1.000000
level_ING	1.000000

dtype: float64



	t_gender	t_age	t_number_of_years_as_teacher \
id_questionnaire			
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	FEMALE	56.0	29.0
5	FEMALE	44.0	20.0
...
20426	NaN	NaN	NaN
20427	FEMALE	48.0	27.0
20428	MALE	34.0	7.0
20429	NaN	NaN	NaN
20430	NaN	NaN	NaN

	t_number_of_years_in_school \
id_questionnaire	
1	NaN

2	NaN
3	NaN
4	18.0
5	17.0
...	...
20426	NaN
20427	11.0
20428	1.0
20429	NaN
20430	NaN

	t_has_taught_same_group_last_two_years \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	False
5	False
...	...
20426	NaN
20427	False
20428	False
20429	NaN
20430	NaN

	t_number_of_students_in_group \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	26.0
5	26.0
...	...
20426	NaN
20427	25.0
20428	26.0
20429	NaN
20430	NaN

	t_number_of_teaching_hours_per_week \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	15.0
5	18.0
...	...
20426	NaN

20427	18.0
20428	12.0
20429	NaN
20430	NaN

id_questionnaire	t_average_explanation_time \
1	NaN
2	NaN
3	NaN
4	2.0
5	2.0
...	...
20426	NaN
20427	2.0
20428	2.0
20429	NaN
20430	NaN

id_questionnaire	t_degree_of_individual_training_incidence \
1	NaN
2	NaN
3	NaN
4	NaN
5	4.0
...	...
20426	NaN
20427	4.0
20428	NaN
20429	NaN
20430	NaN

id_questionnaire	t_seat_configuration ... \
1	NaN ...
2	NaN ...
3	NaN ...
4	5.0 ...
5	5.0 ...
...
20426	NaN ...
20427	4.0 ...
20428	2.0 ...
20429	NaN ...
20430	NaN ...

t_extent_of_student_involvement_during_class \

id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.555556
5	0.666667
...	...
20426	NaN
20427	0.600000
20428	0.500000
20429	NaN
20430	NaN

	t_extent_of_teaching_methods_variety \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.666667
5	0.733333
...	...
20426	NaN
20427	0.600000
20428	0.733333
20429	NaN
20430	NaN

	t_extent_of_opinion_on_school	t_extent_of_class_behaviour \
id_questionnaire		
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	1.000000	0.500000
5	1.000000	0.833333
...
20426	NaN	NaN
20427	1.000000	0.666667
20428	0.666667	0.666667
20429	NaN	NaN
20430	NaN	NaN

	t_extent_of_resource_variety \
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.458333
5	0.625000

...	...
20426	NaN
20427	0.571429
20428	0.583333
20429	NaN
20430	NaN

t_extent_of_good_work_by_non_teachers \	
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	0.555556
5	0.888889
...	...
20426	NaN
20427	0.888889
20428	0.777778
20429	NaN
20430	NaN

t_number_of_individual_training_topics \	
id_questionnaire	
1	NaN
2	NaN
3	NaN
4	NaN
5	1.0
...	...
20426	NaN
20427	2.0
20428	NaN
20429	NaN
20430	NaN

t_extent_of_results_satisfaction level_MAT level_LEN			
id_questionnaire			
1	NaN	3.0	3.0
2	NaN	1.0	1.0
3	NaN	1.0	3.0
4	0.666667	2.0	2.0
5	1.000000	4.0	3.0
...
20426	NaN	1.0	1.0
20427	0.666667	2.0	3.0
20428	0.666667	2.0	2.0
20429	NaN	3.0	3.0
20430	NaN	2.0	3.0

[17806 rows x 31 columns]

	t_gender	t_age \
t_gender	1.000000	-0.106988
t_age	-0.106988	1.000000
t_number_of_years_as_teacher	-0.164927	0.805214
t_number_of_years_in_school	-0.100052	0.342080
t_has_taught_same_group_last_two_years	-0.099470	0.112678
t_number_of_students_in_group	-0.039606	-0.084781
t_number_of_teaching_hours_per_week	-0.063430	0.161169
t_average_explanation_time	-0.120649	-0.026750
t_degree_of_individual_training_incidence	-0.018532	-0.070094
t_seat_configuration	-0.084971	-0.007977
t_behaviour_problems_solution	-0.017883	-0.128723
t_number_of_meetings_with_families	0.042797	0.035827
t_is_enrolled_in_school_training_plan	-0.036946	-0.019760
t_number_of_training_hours_last_six_years	-0.007311	-0.176008
t_extent_of_evaluation_variety	-0.038288	-0.095122
t_extent_of_pfc_incidence	-0.049936	-0.008681
t_extent_of_work_hampered	-0.021475	0.052627
t_extent_of_family_interest	-0.022273	-0.060418
t_extent_of_family_support	-0.052638	0.003515
t_extent_of_work_facilitated_by_management	0.026460	-0.097534
t_extent_of_positive_relationships	0.057162	-0.115030
t_extent_of_student_involvement_during_class	-0.020070	0.012214
t_extent_of_teaching_methods_variety	-0.009020	-0.125742
t_extent_of_opinion_on_school	0.015698	-0.135391
t_extent_of_class_behaviour	0.062479	-0.124766
t_extent_of_resource_variety	-0.066329	-0.065868
t_extent_of_good_work_by_non_teachers	-0.022120	-0.041972
t_number_of_individual_training_topics	0.017530	-0.112589
t_extent_of_results_satisfaction	-0.004949	-0.096093
level_MAT	-0.000143	-0.008295
level_LEN	-0.005552	-0.024233

	t_number_of_years_as_teacher \
t_gender	-0.164927
t_age	0.805214
t_number_of_years_as_teacher	1.000000
t_number_of_years_in_school	0.551866
t_has_taught_same_group_last_two_years	0.120687
t_number_of_students_in_group	-0.012364
t_number_of_teaching_hours_per_week	0.112551
t_average_explanation_time	-0.013882
t_degree_of_individual_training_incidence	-0.059970
t_seat_configuration	-0.027214
t_behaviour_problems_solution	-0.083623

t_number_of_meetings_with_families	-0.009686
t_is_enrolled_in_school_training_plan	-0.070146
t_number_of_training_hours_last_six_years	-0.210125
t_extent_of_evaluation_variety	-0.080935
t_extent_of_pfc_incidence	0.002415
t_extent_of_work_hampered	0.014978
t_extent_of_family_interest	-0.054371
t_extent_of_family_support	0.008892
t_extent_of_work_facilitated_by_management	-0.130148
t_extent_of_positive_relationships	-0.079614
t_extent_of_student_involvement_during_class	0.024339
t_extent_of_teaching_methods_variety	-0.046590
t_extent_of_opinion_on_school	-0.047873
t_extent_of_class_behaviour	-0.063738
t_extent_of_resource_variety	0.007308
t_extent_of_good_work_by_non_teachers	-0.042023
t_number_of_individual_training_topics	-0.098857
t_extent_of_results_satisfaction	-0.028828
level_MAT	0.075011
level_LEN	0.050907

	t_number_of_years_in_school \
t_gender	-0.100052
t_age	0.342080
t_number_of_years_as_teacher	0.551866
t_number_of_years_in_school	1.000000
t_has_taught_same_group_last_two_years	0.038472
t_number_of_students_in_group	0.148874
t_number_of_teaching_hours_per_week	-0.077058
t_average_explanation_time	0.002995
t_degree_of_individual_training_incidence	0.059977
t_seat_configuration	-0.040416
t_behaviour_problems_solution	0.086679
t_number_of_meetings_with_families	-0.079157
t_is_enrolled_in_school_training_plan	-0.047551
t_number_of_training_hours_last_six_years	-0.201768
t_extent_of_evaluation_variety	0.018906
t_extent_of_pfc_incidence	0.128609
t_extent_of_work_hampered	-0.123483
t_extent_of_family_interest	0.080315
t_extent_of_family_support	0.108870
t_extent_of_work_facilitated_by_management	-0.029098
t_extent_of_positive_relationships	0.074178
t_extent_of_student_involvement_during_class	0.047081
t_extent_of_teaching_methods_variety	0.024319
t_extent_of_opinion_on_school	0.168196
t_extent_of_class_behaviour	0.041879
t_extent_of_resource_variety	0.056558

t_extent_of_good_work_by_non_teachers	0.052474
t_number_of_individual_training_topics	-0.087610
t_extent_of_results_satisfaction	0.052541
level_MAT	0.151830
level_LEN	0.108267

↪t_has_taught_same_group_last_two_years \	
t_gender	-0.
↪099470	
t_age	0.
↪112678	
t_number_of_years_as_teacher	0.
↪120687	
t_number_of_years_in_school	0.
↪038472	
t_has_taught_same_group_last_two_years	1.
↪000000	
t_number_of_students_in_group	-0.
↪032484	
t_number_of_teaching_hours_per_week	0.
↪089767	
t_average_explanation_time	-0.
↪048343	
t_degree_of_individual_training_incidence	0.
↪055218	
t_seat_configuration	0.
↪038558	
t_behaviour_problems_solution	-0.
↪033473	
t_number_of_meetings_with_families	0.
↪141232	
t_is_enrolled_in_school_training_plan	0.
↪021192	
t_number_of_training_hours_last_six_years	-0.
↪012515	
t_extent_of_evaluation_variety	0.
↪007145	
t_extent_of_pfc_incidence	-0.
↪012599	
t_extent_of_work_hampered	0.
↪141869	
t_extent_of_family_interest	0.
↪011439	
t_extent_of_family_support	0.
↪098353	

t_extent_of_work_facilitated_by_management ↪062418	-0.
t_extent_of_positive_relationships ↪062956	-0.
t_extent_of_student_involvement_during_class ↪049167	0.
t_extent_of_teaching_methods_variety ↪050967	-0.
t_extent_of_opinion_on_school ↪037333	-0.
t_extent_of_class_behaviour ↪085873	0.
t_extent_of_resource_variety ↪000101	-0.
t_extent_of_good_work_by_non_teachers ↪060428	-0.
t_number_of_individual_training_topics ↪014706	0.
t_extent_of_results_satisfaction ↪029821	0.
level_MAT ↪007877	0.
level_LEN ↪012844	-0.

	t_number_of_students_in_group \
t_gender	-0.039606
t_age	-0.084781
t_number_of_years_as_teacher	-0.012364
t_number_of_years_in_school	0.148874
t_has_taught_same_group_last_two_years	-0.032484
t_number_of_students_in_group	1.000000
t_number_of_teaching_hours_per_week	-0.053343
t_average_explanation_time	0.040897
t_degree_of_individual_training_incidence	0.065211
t_seat_configuration	-0.009042
t_behaviour_problems_solution	0.145043
t_number_of_meetings_with_families	-0.033945
t_is_enrolled_in_school_training_plan	-0.086156
t_number_of_training_hours_last_six_years	0.027917
t_extent_of_evaluation_variety	-0.002029
t_extent_of_pfc_incidence	0.104163
t_extent_of_work_hampered	-0.127681
t_extent_of_family_interest	0.070680
t_extent_of_family_support	0.104725
t_extent_of_work_facilitated_by_management	0.040518
t_extent_of_positive_relationships	0.098453

t_extent_of_student_involvement_during_class	0.023562
t_extent_of_teaching_methods_variety	0.108129
t_extent_of_opinion_on_school	0.166899
t_extent_of_class_behaviour	0.025933
t_extent_of_resource_variety	0.014967
t_extent_of_good_work_by_non_teachers	0.054066
t_number_of_individual_training_topics	0.090980
t_extent_of_results_satisfaction	0.133450
level_MAT	0.072164
level_LEN	0.062272

↪ t_number_of_teaching_hours_per_week \

t_gender	-0.063430
t_age	0.161169
t_number_of_years_as_teacher	0.112551
t_number_of_years_in_school	-0.077058
t_has_taught_same_group_last_two_years	0.089767
t_number_of_students_in_group	-0.053343
t_number_of_teaching_hours_per_week	1.000000
t_average_explanation_time	0.042068
t_degree_of_individual_training_incidence	-0.067785
t_seat_configuration	0.066526
t_behaviour_problems_solution	-0.083733
t_number_of_meetings_with_families	0.001374
t_is_enrolled_in_school_training_plan	0.000971
t_number_of_training_hours_last_six_years	-0.034319
t_extent_of_evaluation_variety	0.038189
t_extent_of_pfc_incidence	-0.112766
t_extent_of_work_hampered	0.137327
t_extent_of_family_interest	-0.048010
t_extent_of_family_support	-0.030155
t_extent_of_work_facilitated_by_management	-0.061017
t_extent_of_positive_relationships	-0.095523
t_extent_of_student_involvement_during_class	0.022033
t_extent_of_teaching_methods_variety	-0.093943
t_extent_of_opinion_on_school	-0.064910
t_extent_of_class_behaviour	-0.067174
t_extent_of_resource_variety	0.044952
t_extent_of_good_work_by_non_teachers	-0.066695
t_number_of_individual_training_topics	-0.050515
t_extent_of_results_satisfaction	-0.078304
level_MAT	-0.025020
level_LEN	-0.032582

t_gender	t_average_explanation_time \	-0.120649
----------	------------------------------	-----------

t_age	-0.026750
t_number_of_years_as_teacher	-0.013882
t_number_of_years_in_school	0.002995
t_has_taught_same_group_last_two_years	-0.048343
t_number_of_students_in_group	0.040897
t_number_of_teaching_hours_per_week	0.042068
t_average_explanation_time	1.000000
t_degree_of_individual_training_incidence	-0.058005
t_seat_configuration	-0.020963
t_behaviour_problems_solution	0.041441
t_number_of_meetings_with_families	0.019709
t_is_enrolled_in_school_training_plan	0.002856
t_number_of_training_hours_last_six_years	-0.037435
t_extent_of_evaluation_variety	0.051429
t_extent_of_pfc_incidence	0.046514
t_extent_of_work_hampered	-0.018812
t_extent_of_family_interest	0.066496
t_extent_of_family_support	-0.031747
t_extent_of_work_facilitated_by_management	-0.005192
t_extent_of_positive_relationships	-0.036478
t_extent_of_student_involvement_during_class	0.032047
t_extent_of_teaching_methods_variety	-0.075840
t_extent_of_opinion_on_school	0.019898
t_extent_of_class_behaviour	-0.082599
t_extent_of_resource_variety	-0.017647
t_extent_of_good_work_by_non_teachers	0.007743
t_number_of_individual_training_topics	-0.050090
t_extent_of_results_satisfaction	-0.013376
level_MAT	0.021821
level_LEN	-0.007260

t_degree_of_individual_training_incidence	-0.018532
t_gender	-0.070094
t_age	-0.059970
t_number_of_years_as_teacher	0.059977
t_number_of_years_in_school	0.055218
t_has_taught_same_group_last_two_years	0.065211
t_number_of_students_in_group	-0.067785
t_number_of_teaching_hours_per_week	

t_average_explanation_time	-0.
↪058005	
t_degree_of_individual_training_incidence	1.
↪000000	
t_seat_configuration	0.
↪021087	
t_behaviour_problems_solution	0.
↪099269	
t_number_of_meetings_with_families	-0.
↪031389	
t_is_enrolled_in_school_training_plan	0.
↪003682	
t_number_of_training_hours_last_six_years	0.
↪216023	
t_extent_of_evaluation_variety	0.
↪016456	
t_extent_of_pfc_incidence	0.
↪248168	
t_extent_of_work_hampered	-0.
↪121155	
t_extent_of_family_interest	0.
↪129112	
t_extent_of_family_support	0.
↪158743	
t_extent_of_work_facilitated_by_management	0.
↪223523	
t_extent_of_positive_relationships	0.
↪256147	
t_extent_of_student_involvement_during_class	0.
↪125704	
t_extent_of_teaching_methods_variety	0.
↪155654	
t_extent_of_opinion_on_school	0.
↪186371	
t_extent_of_class_behaviour	0.
↪177928	
t_extent_of_resource_variety	0.
↪067379	
t_extent_of_good_work_by_non_teachers	0.
↪203504	
t_number_of_individual_training_topics	0.
↪088132	
t_extent_of_results_satisfaction	0.
↪112130	
level_MAT	0.
↪044813	

level_LEN
↪058198

0.

	t_seat_configuration	...	\
t_gender	-0.084971	...	
t_age	-0.007977	...	
t_number_of_years_as_teacher	-0.027214	...	
t_number_of_years_in_school	-0.040416	...	
t_has_taught_same_group_last_two_years	0.038558	...	
t_number_of_students_in_group	-0.009042	...	
t_number_of_teaching_hours_per_week	0.066526	...	
t_average_explanation_time	-0.020963	...	
t_degree_of_individual_training_incidence	0.021087	...	
t_seat_configuration	1.000000	...	
t_behaviour_problems_solution	-0.019926	...	
t_number_of_meetings_with_families	-0.018988	...	
t_is_enrolled_in_school_training_plan	-0.022430	...	
t_number_of_training_hours_last_six_years	-0.010753	...	
t_extent_of_evaluation_variety	0.009719	...	
t_extent_of_pfc_incidence	0.009078	...	
t_extent_of_work_hampered	0.008878	...	
t_extent_of_family_interest	-0.074942	...	
t_extent_of_family_support	-0.033717	...	
t_extent_of_work_facilitated_by_management	-0.024033	...	
t_extent_of_positive_relationships	-0.042779	...	
t_extent_of_student_involvement_during_class	0.049695	...	
t_extent_of_teaching_methods_variety	0.001414	...	
t_extent_of_opinion_on_school	-0.012346	...	
t_extent_of_class_behaviour	-0.010096	...	
t_extent_of_resource_variety	0.037983	...	
t_extent_of_good_work_by_non_teachers	-0.070059	...	
t_number_of_individual_training_topics	-0.059415	...	
t_extent_of_results_satisfaction	0.027751	...	
level_MAT	-0.036483	...	
level_LEN	-0.014880	...	

↪t_extent_of_student_involvement_during_class ↵ \

t_gender	↪-0.020070	↵
t_age	↪ 0.012214	↵
t_number_of_years_as_teacher	↪ 0.024339	↵
t_number_of_years_in_school	↪ 0.047081	↵

t_has_taught_same_group_last_two_years	↪ 0.049167	⌞
t_number_of_students_in_group	↪ 0.023562	⌞
t_number_of_teaching_hours_per_week	↪ 0.022033	⌞
t_average_explanation_time	↪ 0.032047	⌞
t_degree_of_individual_training_incidence	↪ 0.125704	⌞
t_seat_configuration	↪ 0.049695	⌞
t_behaviour_problems_solution	↪ 0.042778	⌞
t_number_of_meetings_with_families	↪ 0.026070	⌞
t_is_enrolled_in_school_training_plan	↪ -0.029118	⌞
t_number_of_training_hours_last_six_years	↪ -0.002080	⌞
t_extent_of_evaluation_variety	↪ 0.329147	⌞
t_extent_of_pfc_incidence	↪ 0.243255	⌞
t_extent_of_work_hampered	↪ -0.178566	⌞
t_extent_of_family_interest	↪ 0.380032	⌞
t_extent_of_family_support	↪ 0.348986	⌞
t_extent_of_work_facilitated_by_management	↪ 0.239679	⌞
t_extent_of_positive_relationships	↪ 0.372735	⌞
t_extent_of_student_involvement_during_class	↪ 1.000000	⌞
t_extent_of_teaching_methods_variety	↪ 0.397081	⌞
t_extent_of_opinion_on_school	↪ 0.237279	⌞
t_extent_of_class_behaviour	↪ 0.439388	⌞
t_extent_of_resource_variety	↪ 0.332805	⌞
t_extent_of_good_work_by_non_teachers	↪ 0.230650	⌞

t_number_of_individual_training_topics		□
↪ 0.069593		
t_extent_of_results_satisfaction		□
↪ 0.293641		
level_MAT		□
↪ 0.107216		
level_LEN		□
↪ 0.118362		
↪t_extent_of_teaching_methods_variety \		□
t_gender	-0.	
↪009020		
t_age	-0.	
↪125742		
t_number_of_years_as_teacher	-0.	
↪046590		
t_number_of_years_in_school	0.	
↪024319		
t_has_taught_same_group_last_two_years	-0.	
↪050967		
t_number_of_students_in_group	0.	
↪108129		
t_number_of_teaching_hours_per_week	-0.	
↪093943		
t_average_explanation_time	-0.	
↪075840		
t_degree_of_individual_training_incidence	0.	
↪155654		
t_seat_configuration	0.	
↪001414		
t_behaviour_problems_solution	0.	
↪030955		
t_number_of_meetings_with_families	0.	
↪049045		
t_is_enrolled_in_school_training_plan	-0.	
↪026398		
t_number_of_training_hours_last_six_years	0.	
↪106046		
t_extent_of_evaluation_variety	0.	
↪488262		
t_extent_of_pfc_incidence	0.	
↪266336		
t_extent_of_work_hampered	-0.	
↪134169		

t_extent_of_family_interest	0.
↪196369	
t_extent_of_family_support	0.
↪146982	
t_extent_of_work_facilitated_by_management	0.
↪257184	
t_extent_of_positive_relationships	0.
↪296364	
t_extent_of_student_involvement_during_class	0.
↪397081	
t_extent_of_teaching_methods_variety	1.
↪000000	
t_extent_of_opinion_on_school	0.
↪249851	
t_extent_of_class_behaviour	0.
↪203190	
t_extent_of_resource_variety	0.
↪422508	
t_extent_of_good_work_by_non_teachers	0.
↪292029	
t_number_of_individual_training_topics	0.
↪153010	
t_extent_of_results_satisfaction	0.
↪129843	
level_MAT	0.
↪081908	
level_LEN	0.
↪088948	

	t_extent_of_opinion_on_school \
t_gender	0.015698
t_age	-0.135391
t_number_of_years_as_teacher	-0.047873
t_number_of_years_in_school	0.168196
t_has_taught_same_group_last_two_years	-0.037333
t_number_of_students_in_group	0.166899
t_number_of_teaching_hours_per_week	-0.064910
t_average_explanation_time	0.019898
t_degree_of_individual_training_incidence	0.186371
t_seat_configuration	-0.012346
t_behaviour_problems_solution	0.097148
t_number_of_meetings_with_families	-0.022824
t_is_enrolled_in_school_training_plan	-0.093654
t_number_of_training_hours_last_six_years	0.023433
t_extent_of_evaluation_variety	0.196332
t_extent_of_pfc_incidence	0.379672

t_extent_of_work_hampered	-0.227954
t_extent_of_family_interest	0.155561
t_extent_of_family_support	0.208881
t_extent_of_work_facilitated_by_management	0.538857
t_extent_of_positive_relationships	0.455556
t_extent_of_student_involvement_during_class	0.237279
t_extent_of_teaching_methods_variety	0.249851
t_extent_of_opinion_on_school	1.000000
t_extent_of_class_behaviour	0.236334
t_extent_of_resource_variety	0.218786
t_extent_of_good_work_by_non_teachers	0.440960
t_number_of_individual_training_topics	0.081463
t_extent_of_results_satisfaction	0.171718
level_MAT	0.150765
level_LEN	0.140045

	t_extent_of_class_behaviour \	
t_gender		0.062479
t_age		-0.124766
t_number_of_years_as_teacher		-0.063738
t_number_of_years_in_school		0.041879
t_has_taught_same_group_last_two_years		0.085873
t_number_of_students_in_group		0.025933
t_number_of_teaching_hours_per_week		-0.067174
t_average_explanation_time		-0.082599
t_degree_of_individual_training_incidence		0.177928
t_seat_configuration		-0.010096
t_behaviour_problems_solution		-0.010889
t_number_of_meetings_with_families		0.008213
t_is_enrolled_in_school_training_plan		-0.006256
t_number_of_training_hours_last_six_years		0.023666
t_extent_of_evaluation_variety		0.120303
t_extent_of_pfc_incidence		0.198886
t_extent_of_work_hampered		-0.153506
t_extent_of_family_interest		0.188523
t_extent_of_family_support		0.293336
t_extent_of_work_facilitated_by_management		0.194664
t_extent_of_positive_relationships		0.370046
t_extent_of_student_involvement_during_class		0.439388
t_extent_of_teaching_methods_variety		0.203190
t_extent_of_opinion_on_school		0.236334
t_extent_of_class_behaviour		1.000000
t_extent_of_resource_variety		0.157627
t_extent_of_good_work_by_non_teachers		0.155926
t_number_of_individual_training_topics		-0.001298
t_extent_of_results_satisfaction		0.375323
level_MAT		0.147322
level_LEN		0.152896

	t_extent_of_resource_variety \
t_gender	-0.066329
t_age	-0.065868
t_number_of_years_as_teacher	0.007308
t_number_of_years_in_school	0.056558
t_has_taught_same_group_last_two_years	-0.000101
t_number_of_students_in_group	0.014967
t_number_of_teaching_hours_per_week	0.044952
t_average_explanation_time	-0.017647
t_degree_of_individual_training_incidence	0.067379
t_seat_configuration	0.037983
t_behaviour_problems_solution	0.101157
t_number_of_meetings_with_families	0.045830
t_is_enrolled_in_school_training_plan	0.024593
t_number_of_training_hours_last_six_years	0.085218
t_extent_of_evaluation_variety	0.382609
t_extent_of_pfc_incidence	0.318244
t_extent_of_work_hampered	-0.197045
t_extent_of_family_interest	0.209696
t_extent_of_family_support	0.148578
t_extent_of_work_facilitated_by_management	0.238818
t_extent_of_positive_relationships	0.223770
t_extent_of_student_involvement_during_class	0.332805
t_extent_of_teaching_methods_variety	0.422508
t_extent_of_opinion_on_school	0.218786
t_extent_of_class_behaviour	0.157627
t_extent_of_resource_variety	1.000000
t_extent_of_good_work_by_non_teachers	0.278978
t_number_of_individual_training_topics	0.112688
t_extent_of_results_satisfaction	0.060411
level_MAT	0.096082
level_LEN	0.073635

↳ t_extent_of_good_work_by_non_teachers \

t_gender	-0.
↳022120	
t_age	-0.
↳041972	
t_number_of_years_as_teacher	-0.
↳042023	
t_number_of_years_in_school	0.
↳052474	
t_has_taught_same_group_last_two_years	-0.
↳060428	

t_number_of_students_in_group	0.
↪054066	
t_number_of_teaching_hours_per_week	-0.
↪066695	
t_average_explanation_time	0.
↪007743	
t_degree_of_individual_training_incidence	0.
↪203504	
t_seat_configuration	-0.
↪070059	
t_behaviour_problems_solution	0.
↪095364	
t_number_of_meetings_with_families	0.
↪018461	
t_is_enrolled_in_school_training_plan	-0.
↪039312	
t_number_of_training_hours_last_six_years	0.
↪023957	
t_extent_of_evaluation_variety	0.
↪202534	
t_extent_of_pfc_incidence	0.
↪406188	
t_extent_of_work_hampered	-0.
↪319484	
t_extent_of_family_interest	0.
↪208978	
t_extent_of_family_support	0.
↪133561	
t_extent_of_work_facilitated_by_management	0.
↪627799	
t_extent_of_positive_relationships	0.
↪458111	
t_extent_of_student_involvement_during_class	0.
↪230650	
t_extent_of_teaching_methods_variety	0.
↪292029	
t_extent_of_opinion_on_school	0.
↪440960	
t_extent_of_class_behaviour	0.
↪155926	
t_extent_of_resource_variety	0.
↪278978	
t_extent_of_good_work_by_non_teachers	1.
↪000000	
t_number_of_individual_training_topics	0.
↪042748	

t_extent_of_results_satisfaction	0.
↪056610	
level_MAT	0.
↪067266	
level_LEN	0.
↪044953	
↪t_number_of_individual_training_topics \	
t_gender	0.
↪017530	
t_age	-0.
↪112589	
t_number_of_years_as_teacher	-0.
↪098857	
t_number_of_years_in_school	-0.
↪087610	
t_has_taught_same_group_last_two_years	0.
↪014706	
t_number_of_students_in_group	0.
↪090980	
t_number_of_teaching_hours_per_week	-0.
↪050515	
t_average_explanation_time	-0.
↪050090	
t_degree_of_individual_training_incidence	0.
↪088132	
t_seat_configuration	-0.
↪059415	
t_behaviour_problems_solution	-0.
↪028153	
t_number_of_meetings_with_families	-0.
↪020567	
t_is_enrolled_in_school_training_plan	-0.
↪059328	
t_number_of_training_hours_last_six_years	0.
↪309339	
t_extent_of_evaluation_variety	0.
↪100841	
t_extent_of_pfc_incidence	0.
↪110764	
t_extent_of_work_hampered	-0.
↪025433	
t_extent_of_family_interest	0.
↪073678	

t_extent_of_family_support	-0.
↪005552	
t_extent_of_work_facilitated_by_management	0.
↪035817	
t_extent_of_positive_relationships	0.
↪035576	
t_extent_of_student_involvement_during_class	0.
↪069593	
t_extent_of_teaching_methods_variety	0.
↪153010	
t_extent_of_opinion_on_school	0.
↪081463	
t_extent_of_class_behaviour	-0.
↪001298	
t_extent_of_resource_variety	0.
↪112688	
t_extent_of_good_work_by_non_teachers	0.
↪042748	
t_number_of_individual_training_topics	1.
↪000000	
t_extent_of_results_satisfaction	0.
↪026206	
level_MAT	0.
↪002458	
level_LEN	0.
↪013597	

	t_extent_of_results_satisfaction \
t_gender	-0.004949
t_age	-0.096093
t_number_of_years_as_teacher	-0.028828
t_number_of_years_in_school	0.052541
t_has_taught_same_group_last_two_years	0.029821
t_number_of_students_in_group	0.133450
t_number_of_teaching_hours_per_week	-0.078304
t_average_explanation_time	-0.013376
t_degree_of_individual_training_incidence	0.112130
t_seat_configuration	0.027751
t_behaviour_problems_solution	0.071713
t_number_of_meetings_with_families	-0.054693
t_is_enrolled_in_school_training_plan	-0.033996
t_number_of_training_hours_last_six_years	-0.064414
t_extent_of_evaluation_variety	0.057017
t_extent_of_pfc_incidence	0.163975
t_extent_of_work_hampered	-0.125497
t_extent_of_family_interest	0.107760
t_extent_of_family_support	0.237994

t_extent_of_work_facilitated_by_management	0.107434
t_extent_of_positive_relationships	0.232811
t_extent_of_student_involvement_during_class	0.293641
t_extent_of_teaching_methods_variety	0.129843
t_extent_of_opinion_on_school	0.171718
t_extent_of_class_behaviour	0.375323
t_extent_of_resource_variety	0.060411
t_extent_of_good_work_by_non_teachers	0.056610
t_number_of_individual_training_topics	0.026206
t_extent_of_results_satisfaction	1.000000
level_MAT	0.113681
level_LEN	0.133119

	level_MAT	level_LEN
t_gender	-0.000143	-0.005552
t_age	-0.008295	-0.024233
t_number_of_years_as_teacher	0.075011	0.050907
t_number_of_years_in_school	0.151830	0.108267
t_has_taught_same_group_last_two_years	0.007877	-0.012844
t_number_of_students_in_group	0.072164	0.062272
t_number_of_teaching_hours_per_week	-0.025020	-0.032582
t_average_explanation_time	0.021821	-0.007260
t_degree_of_individual_training_incidence	0.044813	0.058198
t_seat_configuration	-0.036483	-0.014880
t_behaviour_problems_solution	0.049166	0.021260
t_number_of_meetings_with_families	-0.001645	-0.012352
t_is_enrolled_in_school_training_plan	-0.013105	-0.010456
t_number_of_training_hours_last_six_years	-0.007779	-0.006782
t_extent_of_evaluation_variety	0.049057	0.051692
t_extent_of_pfc_incidence	0.094366	0.097278
t_extent_of_work_hampered	-0.112958	-0.110672
t_extent_of_family_interest	0.046225	0.050040
t_extent_of_family_support	0.109480	0.115443
t_extent_of_work_facilitated_by_management	0.044843	0.065517
t_extent_of_positive_relationships	0.083215	0.079217
t_extent_of_student_involvement_during_class	0.107216	0.118362
t_extent_of_teaching_methods_variety	0.081908	0.088948
t_extent_of_opinion_on_school	0.150765	0.140045
t_extent_of_class_behaviour	0.147322	0.152896
t_extent_of_resource_variety	0.096082	0.073635
t_extent_of_good_work_by_non_teachers	0.067266	0.044953
t_number_of_individual_training_topics	0.002458	0.013597
t_extent_of_results_satisfaction	0.113681	0.133119
level_MAT	1.000000	0.562532
level_LEN	0.562532	1.000000

[31 rows x 31 columns]

5.7 Principals Questionnaire

```
[85]: df_sel_principals, correlation_matrix_sel_principals =  
      ↪compute_exploratory_analysis(df=df_sel, column_pattern="p_",  
      ↪missing_threshold=0.1, annot=False)
```

GOOD FEATURES: 35

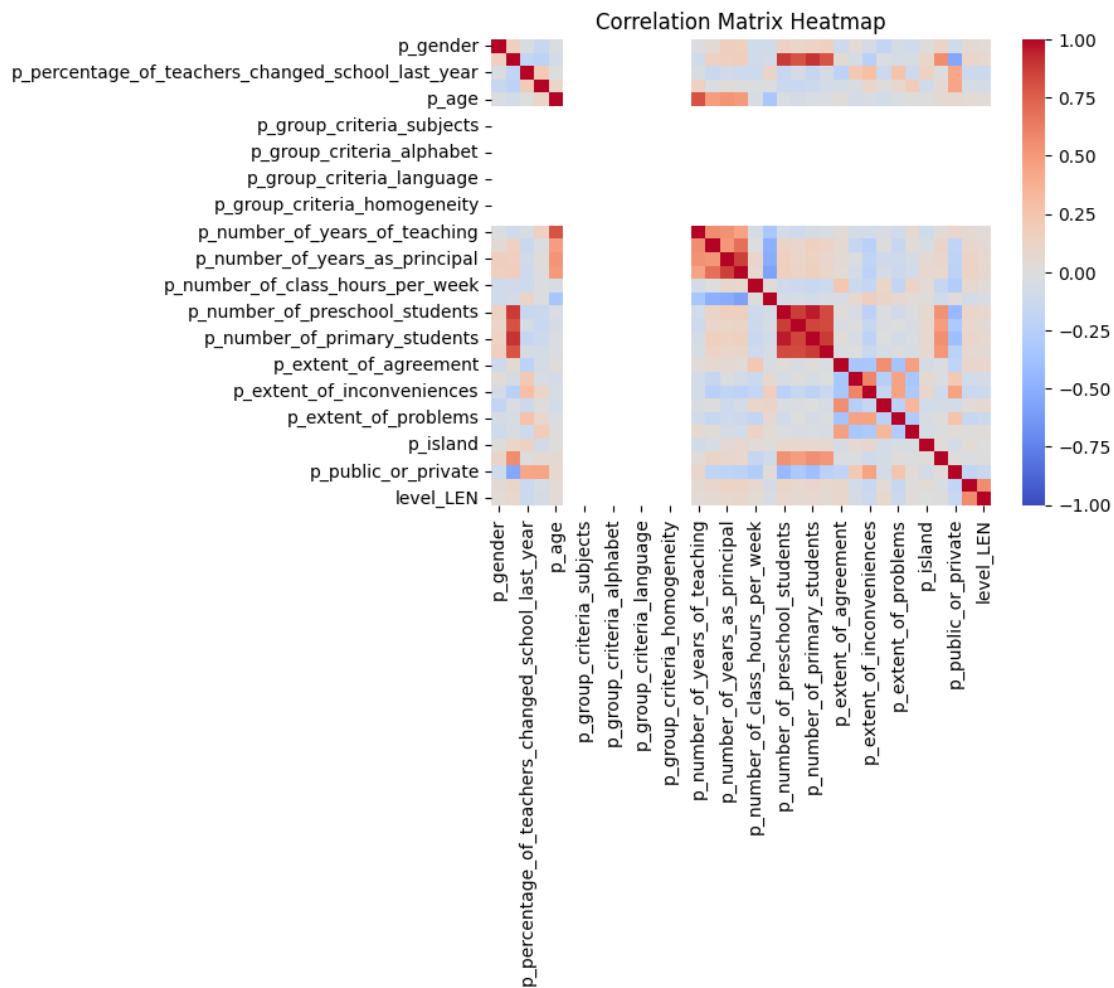
p_gender	0.048916
p_number_of_teachers_in_school	0.055318
p_percentage_of_teachers_changed_school_last_year	0.050264
p_attitude_teacher_training_courses	0.056105
p_age	0.051724
p_group_criteria_random	0.000000
p_group_criteria_subjects	0.000000
p_group_criteria_other	0.000000
p_group_criteria_alphabet	0.000000
p_group_criteria_gender	0.000000
p_group_criteria_language	0.000000
p_group_criteria_performance	0.000000
p_group_criteria_homogeneity	0.000000
p_group_criteria_heterogeneity	0.000000
p_number_of_years_of_teaching	0.058520
p_number_of_years_in_school	0.053690
p_number_of_years_as_principal	0.061889
p_number_of_years_as_principal_in_school	0.054420
p_number_of_class_hours_per_week	0.098001
p_number_of_principals_10_years	0.057958
p_number_of_preschool_students	0.067056
p_number_of_preschool_groups	0.075199
p_number_of_primary_students	0.054757
p_number_of_primary_groups	0.070650
p_extent_of_agreement	0.047568
p_extent_of_factors_limiting_effectiveness	0.048916
p_extent_of_inconveniences	0.047568
p_extent_of_management	0.047568
p_extent_of_problems	0.047568
p_extent_of_satisfaction	0.047568
p_island	0.000000
p_number_of_school_tech_resources	0.000000
p_public_or_private	0.000000
level_MAT	0.000000
level_LEN	0.009772
dtype: float64	

BAD FEATURES

p_number_of_abroad_students_country_no_spanish	0.117657
p_number_of_abroad_students_country_spanish	0.133101
p_number_of_teachers_1_education	1.000000

p_number_of_mandatory_teachers_2_education	1.000000
p_number_of_teachers_1_education_more_5_years	1.000000
p_number_of_mandatory_teachers_2_education_more_5_years	1.000000
p_extent_of_satisfaction_students_results	1.000000
p_extent_of_expected_students_results	1.000000
p_extent_of_students_improvement_expectation	1.000000
p_school_teacher_training_plan	1.000000
p_school_teacher_training_plan_main_theme	1.000000
p_number_of_students	0.249017
p_number_of_groups	0.229754
p_number_of_groups_evaluated_grade	1.000000
p_rate_4_grade_mandatory_2_education_students_different_nationality	1.000000
p_rate_3_grade_mandatory_2_education_students_different_nationality	1.000000
p_rate_6_grade_mandatory_2_education_students_different_nationality	1.000000
level_ING	1.000000

dtype: float64



id_questionnaire	p_gender	p_number_of_teachers_in_school	\
1	MALE	19.0	
2	FEMALE	26.0	
3	FEMALE	42.0	
4	MALE	15.0	
5	MALE	22.0	
...	
20426	FEMALE	24.0	
20427	FEMALE	15.0	
20428	MALE	27.0	
20429	FEMALE	23.0	
20430	MALE	15.0	

id_questionnaire	p_percentage_of_teachers_changed_school_last_year	\
1	3.0	
2	2.0	
3	2.0	
4	1.0	
5	1.0	
...	...	
20426	4.0	
20427	2.0	
20428	1.0	
20429	1.0	
20430	3.0	

id_questionnaire	p_attitude_teacher_training_courses	p_age	\
1	3.0	50.0	
2	3.0	58.0	
3	2.0	45.0	
4	3.0	46.0	
5	2.0	56.0	
...	
20426	3.0	46.0	
20427	3.0	49.0	
20428	3.0	59.0	
20429	4.0	64.0	
20430	2.0	51.0	

id_questionnaire	p_group_criteria_random	p_group_criteria_subjects	\
1	False	False	
2	False	False	
3	False	False	
4	False	False	

5	False	False
...
20426	False	False
20427	False	False
20428	False	False
20429	False	False
20430	False	False

	p_group_criteria_other	p_group_criteria_alphabet	\
id_questionnaire			
1	False	False	
2	False	False	
3	False	False	
4	False	False	
5	False	False	
...	
20426	False	False	
20427	False	False	
20428	False	False	
20429	False	False	
20430	False	False	

	p_group_criteria_gender	...	\
id_questionnaire			
1	False	...	
2	False	...	
3	False	...	
4	False	...	
5	False	...	
...	
20426	False	...	
20427	False	...	
20428	False	...	
20429	False	...	
20430	False	...	

	p_extent_of_factors_limiting_effectiveness	\
id_questionnaire		
1	0.666667	
2	0.375000	
3	0.500000	
4	0.625000	
5	0.458333	
...	...	
20426	0.583333	
20427	0.333333	
20428	0.458333	
20429	0.083333	

20430

0.583333

id_questionnaire	p_extent_of_inconveniences	p_extent_of_management \
1	0.750000	0.666667
2	0.305556	0.857143
3	0.606061	0.666667
4	0.666667	0.809524
5	0.333333	0.714286
...
20426	0.555556	0.952381
20427	0.416667	0.761905
20428	0.388889	0.809524
20429	0.000000	1.000000
20430	0.777778	0.666667

id_questionnaire	p_extent_of_problems	p_extent_of_satisfaction \
1	0.314815	0.611111
2	0.074074	0.888889
3	0.314815	0.666667
4	0.129630	0.555556
5	0.055556	0.722222
...
20426	0.129630	0.666667
20427	0.055556	0.777778
20428	0.156863	0.888889
20429	0.025641	1.000000
20430	0.129630	0.555556

id_questionnaire	p_island	p_number_of_school_tech_resources \
1	TENERIFE_CAPITAL	196.0
2	GRAN_CANARIA_CAPITAL	157.0
3	GRAN_CANARIA_CAPITAL	62.0
4	TENERIFE_CAPITAL	131.0
5	TENERIFE_CAPITAL	118.0
...
20426	GRAN_CANARIA_PROVINCE	264.0
20427	TENERIFE_CAPITAL	178.0
20428	GRAN_CANARIA_CAPITAL	311.0
20429	TENERIFE_CAPITAL	30.0
20430	TENERIFE_CAPITAL	54.0

id_questionnaire	p_public_or_private	level_MAT	level_LEN
1	PUBLIC	3.0	3.0
2	PUBLIC	1.0	1.0

3	PRIVATE	1.0	3.0
4	PUBLIC	2.0	2.0
5	PRIVATE	4.0	3.0
...
20426	PUBLIC	1.0	1.0
20427	PUBLIC	2.0	3.0
20428	PUBLIC	2.0	2.0
20429	PRIVATE	3.0	3.0
20430	PUBLIC	2.0	3.0

[17806 rows x 35 columns]

	p_gender \
p_gender	1.000000
p_number_of_teachers_in_school	0.159623
p_percentage_of_teachers_changed_school_last_year	-0.045782
p_attitude_teacher_training_courses	-0.155428
p_age	-0.024211
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.020059
p_number_of_years_in_school	0.047504
p_number_of_years_as_principal	0.167834
p_number_of_years_as_principal_in_school	0.155515
p_number_of_class_hours_per_week	-0.096391
p_number_of_principals_10_years	-0.091932
p_number_of_preschool_students	0.146669
p_number_of_preschool_groups	0.107065
p_number_of_primary_students	0.128383
p_number_of_primary_groups	0.160362
p_extent_of_agreement	-0.108918
p_extent_of_factors_limiting_effectiveness	0.038589
p_extent_of_inconveniences	-0.067401
p_extent_of_management	-0.176006
p_extent_of_problems	-0.056263
p_extent_of_satisfaction	-0.095692
p_island	0.009655
p_number_of_school_tech_resources	0.111593
p_public_or_private	-0.099218
level_MAT	0.053646
level_LEN	0.044194

↪ p_number_of_teachers_in_school \

p_gender	0.159623
p_number_of_teachers_in_school	1.000000
p_percentage_of_teachers_changed_school_last_year	-0.185265
p_attitude_teacher_training_courses	-0.215315
p_age	-0.071810
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.094354
p_number_of_years_in_school	0.142224
p_number_of_years_as_principal	0.161235
p_number_of_years_as_principal_in_school	0.165651
p_number_of_class_hours_per_week	-0.091841
p_number_of_principals_10_years	-0.066192
p_number_of_preschool_students	0.880144
p_number_of_preschool_groups	0.802674
p_number_of_primary_students	0.898519
p_number_of_primary_groups	0.794199
p_extent_of_agreement	0.053070
p_extent_of_factors_limiting_effectiveness	-0.061530
p_extent_of_inconveniences	-0.240613
p_extent_of_management	-0.027679
p_extent_of_problems	-0.054402
p_extent_of_satisfaction	-0.041268
p_island	0.107208
p_number_of_school_tech_resources	0.556685
p_public_or_private	-0.547485
level_MAT	0.102478
level_LEN	0.085440

↪ p_percentage_of_teachers_changed_school_last_year \

p_gender	
↪	-0.045782
p_number_of_teachers_in_school	
↪	-0.185265
p_percentage_of_teachers_changed_school_last_year	
↪	1.000000

p_attitude_teacher_training_courses	0.201196	□
p_age	-0.004783	□
p_group_criteria_random	NaN	□
p_group_criteria_subjects	NaN	□
p_group_criteria_other	NaN	□
p_group_criteria_alphabet	NaN	□
p_group_criteria_gender	NaN	□
p_group_criteria_language	NaN	□
p_group_criteria_performance	NaN	□
p_group_criteria_homogeneity	NaN	□
p_group_criteria_heterogeneity	NaN	□
p_number_of_years_of_teaching	-0.044489	□
p_number_of_years_in_school	-0.134029	□
p_number_of_years_as_principal	-0.091029	□
p_number_of_years_as_principal_in_school	-0.114515	□
p_number_of_class_hours_per_week	-0.114304	□
p_number_of_principals_10_years	0.110425	□
p_number_of_preschool_students	-0.151578	□
p_number_of_preschool_groups	-0.060950	□
p_number_of_primary_students	-0.162638	□
p_number_of_primary_groups	-0.061627	□
p_extent_of_agreement	-0.221886	□
p_extent_of_factors_limiting_effectiveness	0.212428	□

p_extent_of_inconveniences		□
↪	0.281746	
p_extent_of_management		□
↪	-0.120551	
p_extent_of_problems		□
↪	0.262679	
p_extent_of_satisfaction		□
↪	-0.113346	
p_island		□
↪	0.135192	
p_number_of_school_tech_resources		□
↪	0.004551	
p_public_or_private		□
↪	0.434857	
level_MAT		□
↪	-0.132012	
level_LEN		□
↪	-0.124095	
↪p_attitude_teacher_training_courses \		□
p_gender		-0.
↪155428		
p_number_of_teachers_in_school		-0.
↪215315		
p_percentage_of_teachers_changed_school_last_year		0.
↪201196		
p_attitude_teacher_training_courses		1.
↪000000		
p_age		0.
↪106640		
p_group_criteria_random		□
↪ NaN		
p_group_criteria_subjects		□
↪ NaN		
p_group_criteria_other		□
↪ NaN		
p_group_criteria_alphabet		□
↪ NaN		
p_group_criteria_gender		□
↪ NaN		
p_group_criteria_language		□
↪ NaN		
p_group_criteria_performance		□
↪ NaN		

p_group_criteria_homogeneity	0.
↪ NaN	
p_group_criteria_heterogeneity	0.
↪ NaN	
p_number_of_years_of_teaching	0.
↪ 148419	
p_number_of_years_in_school	-0.
↪ 017325	
p_number_of_years_as_principal	0.
↪ 016209	
p_number_of_years_as_principal_in_school	0.
↪ 008065	
p_number_of_class_hours_per_week	-0.
↪ 009184	
p_number_of_principals_10_years	-0.
↪ 014383	
p_number_of_preschool_students	-0.
↪ 136417	
p_number_of_preschool_groups	-0.
↪ 144948	
p_number_of_primary_students	-0.
↪ 121403	
p_number_of_primary_groups	-0.
↪ 076768	
p_extent_of_agreement	-0.
↪ 059628	
p_extent_of_factors_limiting_effectiveness	-0.
↪ 044016	
p_extent_of_inconveniences	0.
↪ 101608	
p_extent_of_management	0.
↪ 010726	
p_extent_of_problems	0.
↪ 083005	
p_extent_of_satisfaction	0.
↪ 173843	
p_island	-0.
↪ 087793	
p_number_of_school_tech_resources	0.
↪ 059893	
p_public_or_private	0.
↪ 444140	
level_MAT	-0.
↪ 053197	
level_LEN	-0.
↪ 061591	

	p_age \
p_gender	-0.024211
p_number_of_teachers_in_school	-0.071810
p_percentage_of_teachers_changed_school_last_year	-0.004783
p_attitude_teacher_training_courses	0.106640
p_age	1.000000
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.797119
p_number_of_years_in_school	0.479378
p_number_of_years_as_principal	0.529904
p_number_of_years_as_principal_in_school	0.496292
p_number_of_class_hours_per_week	-0.046085
p_number_of_principals_10_years	-0.336140
p_number_of_preschool_students	-0.024041
p_number_of_preschool_groups	-0.061992
p_number_of_primary_students	-0.023715
p_number_of_primary_groups	-0.027358
p_extent_of_agreement	0.016133
p_extent_of_factors_limiting_effectiveness	-0.066245
p_extent_of_inconveniences	-0.045046
p_extent_of_management	-0.085886
p_extent_of_problems	-0.030693
p_extent_of_satisfaction	0.014912
p_island	0.003910
p_number_of_school_tech_resources	0.061936
p_public_or_private	0.072827
level_MAT	0.044703
level_LEN	0.037460

	p_group_criteria_random \
p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN

p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_class_hours_per_week	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_subjects \	
p_gender		NaN
p_number_of_teachers_in_school		NaN
p_percentage_of_teachers_changed_school_last_year		NaN
p_attitude_teacher_training_courses		NaN
p_age		NaN
p_group_criteria_random		NaN
p_group_criteria_subjects		NaN
p_group_criteria_other		NaN
p_group_criteria_alphabet		NaN
p_group_criteria_gender		NaN
p_group_criteria_language		NaN
p_group_criteria_performance		NaN
p_group_criteria_homogeneity		NaN
p_group_criteria_heterogeneity		NaN
p_number_of_years_of_teaching		NaN
p_number_of_years_in_school		NaN
p_number_of_years_as_principal		NaN
p_number_of_years_as_principal_in_school		NaN
p_number_of_class_hours_per_week		NaN
p_number_of_principals_10_years		NaN

p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_other \
p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_class_hours_per_week	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN

p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_alphabet	\
p_gender		NaN
p_number_of_teachers_in_school		NaN
p_percentage_of_teachers_changed_school_last_year		NaN
p_attitude_teacher_training_courses		NaN
p_age		NaN
p_group_criteria_random		NaN
p_group_criteria_subjects		NaN
p_group_criteria_other		NaN
p_group_criteria_alphabet		NaN
p_group_criteria_gender		NaN
p_group_criteria_language		NaN
p_group_criteria_performance		NaN
p_group_criteria_homogeneity		NaN
p_group_criteria_heterogeneity		NaN
p_number_of_years_of_teaching		NaN
p_number_of_years_in_school		NaN
p_number_of_years_as_principal		NaN
p_number_of_years_as_principal_in_school		NaN
p_number_of_class_hours_per_week		NaN
p_number_of_principals_10_years		NaN
p_number_of_preschool_students		NaN
p_number_of_preschool_groups		NaN
p_number_of_primary_students		NaN
p_number_of_primary_groups		NaN
p_extent_of_agreement		NaN
p_extent_of_factors_limiting_effectiveness		NaN
p_extent_of_inconveniences		NaN
p_extent_of_management		NaN
p_extent_of_problems		NaN
p_extent_of_satisfaction		NaN
p_island		NaN
p_number_of_school_tech_resources		NaN
p_public_or_private		NaN
level_MAT		NaN
level_LEN		NaN

	p_group_criteria_gender	\
p_gender		NaN
p_number_of_teachers_in_school		NaN
p_percentage_of_teachers_changed_school_last_year		NaN
p_attitude_teacher_training_courses		NaN
p_age		NaN

p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_class_hours_per_week	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	...	\
p_gender	...	
p_number_of_teachers_in_school	...	
p_percentage_of_teachers_changed_school_last_year	...	
p_attitude_teacher_training_courses	...	
p_age	...	
p_group_criteria_random	...	
p_group_criteria_subjects	...	
p_group_criteria_other	...	
p_group_criteria_alphabet	...	
p_group_criteria_gender	...	
p_group_criteria_language	...	
p_group_criteria_performance	...	
p_group_criteria_homogeneity	...	
p_group_criteria_heterogeneity	...	
p_number_of_years_of_teaching	...	
p_number_of_years_in_school	...	

p_number_of_years_as_principal	...	
p_number_of_years_as_principal_in_school	...	
p_number_of_class_hours_per_week	...	
p_number_of_principals_10_years	...	
p_number_of_preschool_students	...	
p_number_of_preschool_groups	...	
p_number_of_primary_students	...	
p_number_of_primary_groups	...	
p_extent_of_agreement	...	
p_extent_of_factors_limiting_effectiveness	...	
p_extent_of_inconveniences	...	
p_extent_of_management	...	
p_extent_of_problems	...	
p_extent_of_satisfaction	...	
p_island	...	
p_number_of_school_tech_resources	...	
p_public_or_private	...	
level_MAT	...	
level_LEN	...	
↪ p_extent_of_factors_limiting_effectiveness \		⌋
p_gender		⌋
↪ 0.038589		
p_number_of_teachers_in_school		⌋
↪ -0.061530		
p_percentage_of_teachers_changed_school_last_year		⌋
↪ 0.212428		
p_attitude_teacher_training_courses		⌋
↪ -0.044016		
p_age		⌋
↪ -0.066245		
p_group_criteria_random		⌋
↪ NaN		
p_group_criteria_subjects		⌋
↪ NaN		
p_group_criteria_other		⌋
↪ NaN		
p_group_criteria_alphabet		⌋
↪ NaN		
p_group_criteria_gender		⌋
↪ NaN		
p_group_criteria_language		⌋
↪ NaN		
p_group_criteria_performance		⌋
↪ NaN		

p_group_criteria_homogeneity	NaN	␣
p_group_criteria_heterogeneity	NaN	␣
p_number_of_years_of_teaching	-0.073058	␣
p_number_of_years_in_school	-0.160859	␣
p_number_of_years_as_principal	-0.022741	␣
p_number_of_years_as_principal_in_school	-0.075019	␣
p_number_of_class_hours_per_week	-0.121794	␣
p_number_of_principals_10_years	0.067971	␣
p_number_of_preschool_students	-0.051689	␣
p_number_of_preschool_groups	0.000143	␣
p_number_of_primary_students	-0.053022	␣
p_number_of_primary_groups	-0.004780	␣
p_extent_of_agreement	-0.330075	␣
p_extent_of_factors_limiting_effectiveness	1.000000	␣
p_extent_of_inconveniences	0.599354	␣
p_extent_of_management	-0.263730	␣
p_extent_of_problems	0.448839	␣
p_extent_of_satisfaction	-0.342985	␣
p_island	0.055340	␣
p_number_of_school_tech_resources	-0.020285	␣
p_public_or_private	0.215288	␣
level_MAT	-0.080427	␣
level_LEN	-0.062271	␣

	p_extent_of_inconveniences \
p_gender	-0.067401
p_number_of_teachers_in_school	-0.240613
p_percentage_of_teachers_changed_school_last_year	0.281746
p_attitude_teacher_training_courses	0.101608
p_age	-0.045046
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.073439
p_number_of_years_in_school	-0.244957
p_number_of_years_as_principal	-0.201010
p_number_of_years_as_principal_in_school	-0.226483
p_number_of_class_hours_per_week	-0.165187
p_number_of_principals_10_years	0.168138
p_number_of_preschool_students	-0.230243
p_number_of_preschool_groups	-0.150370
p_number_of_primary_students	-0.215991
p_number_of_primary_groups	-0.138480
p_extent_of_agreement	-0.340746
p_extent_of_factors_limiting_effectiveness	0.599354
p_extent_of_inconveniences	1.000000
p_extent_of_management	-0.162628
p_extent_of_problems	0.459965
p_extent_of_satisfaction	-0.311237
p_island	0.093214
p_number_of_school_tech_resources	-0.097926
p_public_or_private	0.454093
level_MAT	-0.142821
level_LEN	-0.121366

	p_extent_of_management \
p_gender	-0.176006
p_number_of_teachers_in_school	-0.027679
p_percentage_of_teachers_changed_school_last_year	-0.120551
p_attitude_teacher_training_courses	0.010726
p_age	-0.085886
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN

p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.035008
p_number_of_years_in_school	0.005647
p_number_of_years_as_principal	-0.108370
p_number_of_years_as_principal_in_school	-0.066178
p_number_of_class_hours_per_week	0.108932
p_number_of_principals_10_years	0.116240
p_number_of_preschool_students	-0.028844
p_number_of_preschool_groups	-0.016151
p_number_of_primary_students	-0.003104
p_number_of_primary_groups	-0.052893
p_extent_of_agreement	0.553430
p_extent_of_factors_limiting_effectiveness	-0.263730
p_extent_of_inconveniences	-0.162628
p_extent_of_management	1.000000
p_extent_of_problems	-0.232462
p_extent_of_satisfaction	0.337143
p_island	-0.082008
p_number_of_school_tech_resources	-0.074406
p_public_or_private	-0.079065
level_MAT	0.038399
level_LEN	0.074959

	p_extent_of_problems \
p_gender	-0.056263
p_number_of_teachers_in_school	-0.054402
p_percentage_of_teachers_changed_school_last_year	0.262679
p_attitude_teacher_training_courses	0.083005
p_age	-0.030693
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.089881
p_number_of_years_in_school	-0.154163
p_number_of_years_as_principal	-0.111545
p_number_of_years_as_principal_in_school	-0.128506
p_number_of_class_hours_per_week	-0.175868
p_number_of_principals_10_years	0.064288

p_number_of_preschool_students	-0.086756
p_number_of_preschool_groups	-0.002107
p_number_of_primary_students	-0.065957
p_number_of_primary_groups	-0.008929
p_extent_of_agreement	-0.365932
p_extent_of_factors_limiting_effectiveness	0.448839
p_extent_of_inconveniences	0.459965
p_extent_of_management	-0.232462
p_extent_of_problems	1.000000
p_extent_of_satisfaction	-0.295867
p_island	-0.043104
p_number_of_school_tech_resources	0.012170
p_public_or_private	0.253034
level_MAT	-0.152743
level_LEN	-0.117485

	p_extent_of_satisfaction \
p_gender	-0.095692
p_number_of_teachers_in_school	-0.041268
p_percentage_of_teachers_changed_school_last_year	-0.113346
p_attitude_teacher_training_courses	0.173843
p_age	0.014912
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.064672
p_number_of_years_in_school	0.020068
p_number_of_years_as_principal	-0.059122
p_number_of_years_as_principal_in_school	-0.085462
p_number_of_class_hours_per_week	0.121301
p_number_of_principals_10_years	0.041289
p_number_of_preschool_students	-0.018821
p_number_of_preschool_groups	-0.032589
p_number_of_primary_students	-0.028711
p_number_of_primary_groups	-0.041391
p_extent_of_agreement	0.461188
p_extent_of_factors_limiting_effectiveness	-0.342985
p_extent_of_inconveniences	-0.311237
p_extent_of_management	0.337143
p_extent_of_problems	-0.295867
p_extent_of_satisfaction	1.000000
p_island	-0.049515

p_number_of_school_tech_resources	-0.009083
p_public_or_private	0.033937
level_MAT	0.014379
level_LEN	0.030505

	p_island \
p_gender	0.009655
p_number_of_teachers_in_school	0.107208
p_percentage_of_teachers_changed_school_last_year	0.135192
p_attitude_teacher_training_courses	-0.087793
p_age	0.003910
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.099812
p_number_of_years_in_school	-0.041988
p_number_of_years_as_principal	0.059209
p_number_of_years_as_principal_in_school	0.081011
p_number_of_class_hours_per_week	0.033677
p_number_of_principals_10_years	0.018335
p_number_of_preschool_students	0.072474
p_number_of_preschool_groups	0.085912
p_number_of_primary_students	0.066444
p_number_of_primary_groups	0.063813
p_extent_of_agreement	-0.057438
p_extent_of_factors_limiting_effectiveness	0.055340
p_extent_of_inconveniences	0.093214
p_extent_of_management	-0.082008
p_extent_of_problems	-0.043104
p_extent_of_satisfaction	-0.049515
p_island	1.000000
p_number_of_school_tech_resources	0.027829
p_public_or_private	-0.074768
level_MAT	0.044030
level_LEN	-0.001744

↪ p_number_of_school_tech_resources \

p_gender

↪ 111593

0.

p_number_of_teachers_in_school	0.
↪556685	
p_percentage_of_teachers_changed_school_last_year	0.
↪004551	
p_attitude_teacher_training_courses	0.
↪059893	
p_age	0.
↪061936	
p_group_criteria_random	␣
↪NaN	
p_group_criteria_subjects	␣
↪NaN	
p_group_criteria_other	␣
↪NaN	
p_group_criteria_alphabet	␣
↪NaN	
p_group_criteria_gender	␣
↪NaN	
p_group_criteria_language	␣
↪NaN	
p_group_criteria_performance	␣
↪NaN	
p_group_criteria_homogeneity	␣
↪NaN	
p_group_criteria_heterogeneity	␣
↪NaN	
p_number_of_years_of_teaching	0.
↪032275	
p_number_of_years_in_school	0.
↪113806	
p_number_of_years_as_principal	0.
↪093311	
p_number_of_years_as_principal_in_school	0.
↪084736	
p_number_of_class_hours_per_week	-0.
↪161634	
p_number_of_principals_10_years	-0.
↪095084	
p_number_of_preschool_students	0.
↪535117	
p_number_of_preschool_groups	0.
↪477388	
p_number_of_primary_students	0.
↪550508	
p_number_of_primary_groups	0.
↪512754	

p_extent_of_agreement	-0.
↪029804	
p_extent_of_factors_limiting_effectiveness	-0.
↪020285	
p_extent_of_inconveniences	-0.
↪097926	
p_extent_of_management	-0.
↪074406	
p_extent_of_problems	0.
↪012170	
p_extent_of_satisfaction	-0.
↪009083	
p_island	0.
↪027829	
p_number_of_school_tech_resources	1.
↪000000	
p_public_or_private	0.
↪025741	
level_MAT	-0.
↪028566	
level_LEN	0.
↪005489	

	p_public_or_private \
p_gender	-0.099218
p_number_of_teachers_in_school	-0.547485
p_percentage_of_teachers_changed_school_last_year	0.434857
p_attitude_teacher_training_courses	0.444140
p_age	0.072827
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.098537
p_number_of_years_in_school	-0.199116
p_number_of_years_as_principal	-0.209854
p_number_of_years_as_principal_in_school	-0.223973
p_number_of_class_hours_per_week	-0.274363
p_number_of_principals_10_years	0.028208
p_number_of_preschool_students	-0.433325
p_number_of_preschool_groups	-0.319348
p_number_of_primary_students	-0.391661

p_number_of_primary_groups	-0.228892
p_extent_of_agreement	-0.198010
p_extent_of_factors_limiting_effectiveness	0.215288
p_extent_of_inconveniences	0.454093
p_extent_of_management	-0.079065
p_extent_of_problems	0.253034
p_extent_of_satisfaction	0.033937
p_island	-0.074768
p_number_of_school_tech_resources	0.025741
p_public_or_private	1.000000
level_MAT	-0.182115
level_LEN	-0.141777

	level_MAT	level_LEN
p_gender	0.053646	0.044194
p_number_of_teachers_in_school	0.102478	0.085440
p_percentage_of_teachers_changed_school_last_year	-0.132012	-0.124095
p_attitude_teacher_training_courses	-0.053197	-0.061591
p_age	0.044703	0.037460
p_group_criteria_random	NaN	NaN
p_group_criteria_subjects	NaN	NaN
p_group_criteria_other	NaN	NaN
p_group_criteria_alphabet	NaN	NaN
p_group_criteria_gender	NaN	NaN
p_group_criteria_language	NaN	NaN
p_group_criteria_performance	NaN	NaN
p_group_criteria_homogeneity	NaN	NaN
p_group_criteria_heterogeneity	NaN	NaN
p_number_of_years_of_teaching	0.059156	0.035318
p_number_of_years_in_school	0.089493	0.060077
p_number_of_years_as_principal	0.118076	0.085325
p_number_of_years_as_principal_in_school	0.120255	0.089276
p_number_of_class_hours_per_week	0.075927	0.041353
p_number_of_principals_10_years	-0.063751	-0.033567
p_number_of_preschool_students	0.106308	0.085683
p_number_of_preschool_groups	0.084685	0.059827
p_number_of_primary_students	0.097782	0.083027
p_number_of_primary_groups	0.074565	0.048145
p_extent_of_agreement	0.089626	0.096205
p_extent_of_factors_limiting_effectiveness	-0.080427	-0.062271
p_extent_of_inconveniences	-0.142821	-0.121366
p_extent_of_management	0.038399	0.074959
p_extent_of_problems	-0.152743	-0.117485
p_extent_of_satisfaction	0.014379	0.030505
p_island	0.044030	-0.001744
p_number_of_school_tech_resources	-0.028566	0.005489
p_public_or_private	-0.182115	-0.141777
level_MAT	1.000000	0.562532

level_LEN 0.562532 1.000000

[35 rows x 35 columns]

```
[86]: df_sel_principals, _ = compute_exploratory_analysis(df=df_sel,
    ↪column_pattern="p_", missing_threshold=0.09, annot=False)
df_sel_principals
```

GOOD FEATURES: 34

p_gender	0.048916
p_number_of_teachers_in_school	0.055318
p_percentage_of_teachers_changed_school_last_year	0.050264
p_attitude_teacher_training_courses	0.056105
p_age	0.051724
p_group_criteria_random	0.000000
p_group_criteria_subjects	0.000000
p_group_criteria_other	0.000000
p_group_criteria_alphabet	0.000000
p_group_criteria_gender	0.000000
p_group_criteria_language	0.000000
p_group_criteria_performance	0.000000
p_group_criteria_homogeneity	0.000000
p_group_criteria_heterogeneity	0.000000
p_number_of_years_of_teaching	0.058520
p_number_of_years_in_school	0.053690
p_number_of_years_as_principal	0.061889
p_number_of_years_as_principal_in_school	0.054420
p_number_of_principals_10_years	0.057958
p_number_of_preschool_students	0.067056
p_number_of_preschool_groups	0.075199
p_number_of_primary_students	0.054757
p_number_of_primary_groups	0.070650
p_extent_of_agreement	0.047568
p_extent_of_factors_limiting_effectiveness	0.048916
p_extent_of_inconveniences	0.047568
p_extent_of_management	0.047568
p_extent_of_problems	0.047568
p_extent_of_satisfaction	0.047568
p_island	0.000000
p_number_of_school_tech_resources	0.000000
p_public_or_private	0.000000
level_MAT	0.000000
level_LEN	0.009772

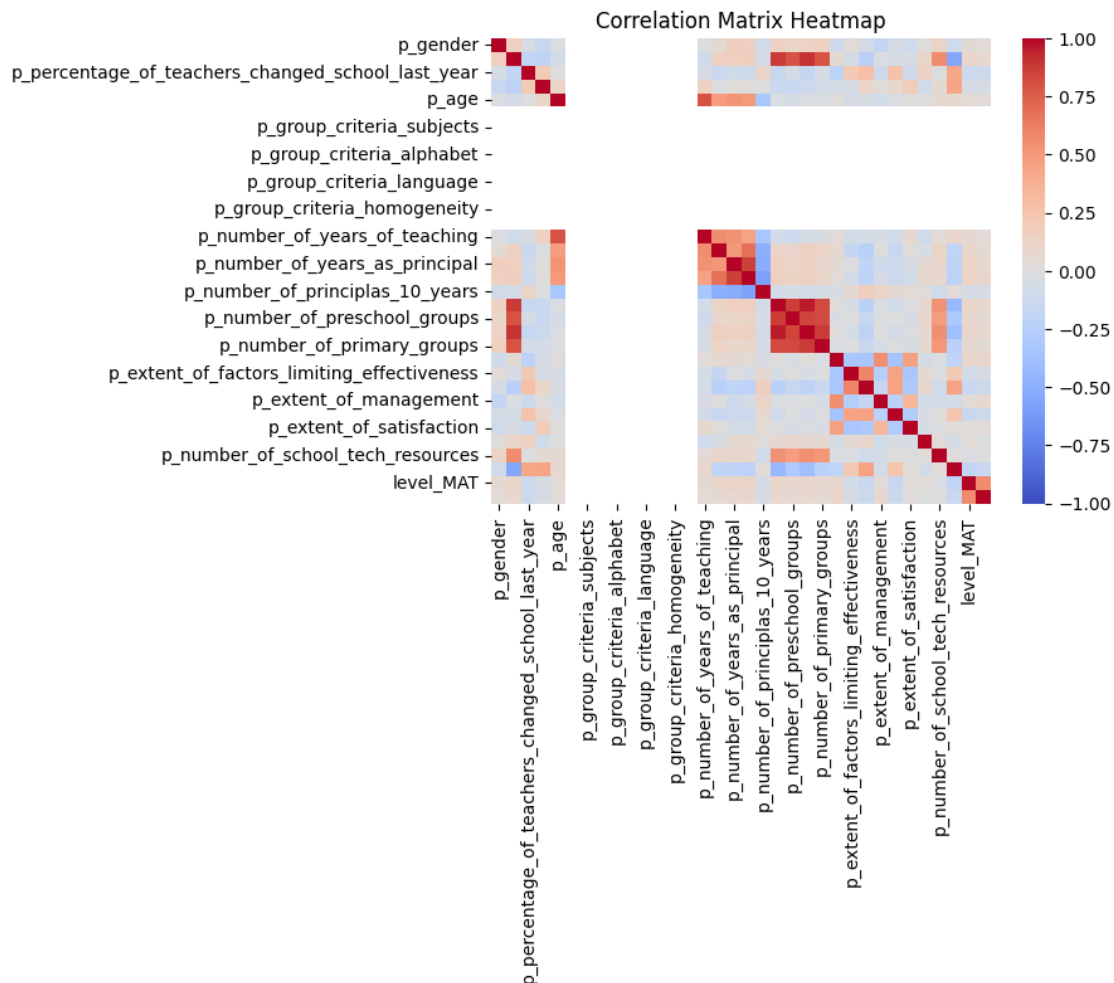
dtype: float64

BAD FEATURES

p_number_of_abroad_students_country_no_spanish	0.117657
p_number_of_abroad_students_country_spanish	0.133101

p_number_of_teachers_1_education	1.000000
p_number_of_mandatory_teachers_2_education	1.000000
p_number_of_teachers_1_education_more_5_years	1.000000
p_number_of_mandatory_teachers_2_education_more_5_years	1.000000
p_extent_of_satisfaction_students_results	1.000000
p_extent_of_expected_students_results	1.000000
p_extent_of_students_improvement_expectation	1.000000
p_school_teacher_training_plan	1.000000
p_school_teacher_training_plan_main_theme	1.000000
p_number_of_class_hours_per_week	0.098001
p_number_of_students	0.249017
p_number_of_groups	0.229754
p_number_of_groups_evaluated_grade	1.000000
p_rate_4_grade_mandatory_2_education_students_different_nationality	1.000000
p_rate_3_grade_mandatory_2_education_students_different_nationality	1.000000
p_rate_6_grade_mandatory_2_education_students_different_nationality	1.000000
level_ING	1.000000

dtype: float64



id_questionnaire	p_gender	p_number_of_teachers_in_school	\
1	MALE	19.0	
2	FEMALE	26.0	
3	FEMALE	42.0	
4	MALE	15.0	
5	MALE	22.0	
...	
20426	FEMALE	24.0	
20427	FEMALE	15.0	
20428	MALE	27.0	
20429	FEMALE	23.0	
20430	MALE	15.0	

id_questionnaire	p_percentage_of_teachers_changed_school_last_year	\
1	3.0	
2	2.0	
3	2.0	
4	1.0	
5	1.0	
...	...	
20426	4.0	
20427	2.0	
20428	1.0	
20429	1.0	
20430	3.0	

id_questionnaire	p_attitude_teacher_training_courses	p_age	\
1	3.0	50.0	
2	3.0	58.0	
3	2.0	45.0	
4	3.0	46.0	
5	2.0	56.0	
...	
20426	3.0	46.0	
20427	3.0	49.0	
20428	3.0	59.0	
20429	4.0	64.0	
20430	2.0	51.0	

id_questionnaire	p_group_criteria_random	p_group_criteria_subjects	\
1	False	False	

2	False	False
3	False	False
4	False	False
5	False	False
...
20426	False	False
20427	False	False
20428	False	False
20429	False	False
20430	False	False

	p_group_criteria_other	p_group_criteria_alphabet \
id_questionnaire		
1	False	False
2	False	False
3	False	False
4	False	False
5	False	False
...
20426	False	False
20427	False	False
20428	False	False
20429	False	False
20430	False	False

	p_group_criteria_gender	...	\
id_questionnaire			
1	False	...	
2	False	...	
3	False	...	
4	False	...	
5	False	...	
...	
20426	False	...	
20427	False	...	
20428	False	...	
20429	False	...	
20430	False	...	

	p_extent_of_factors_limiting_effectiveness \
id_questionnaire	
1	0.666667
2	0.375000
3	0.500000
4	0.625000
5	0.458333
...	...
20426	0.583333

20427	0.333333
20428	0.458333
20429	0.083333
20430	0.583333

id_questionnaire	p_extent_of_inconveniences	p_extent_of_management \
1	0.750000	0.666667
2	0.305556	0.857143
3	0.606061	0.666667
4	0.666667	0.809524
5	0.333333	0.714286
...
20426	0.555556	0.952381
20427	0.416667	0.761905
20428	0.388889	0.809524
20429	0.000000	1.000000
20430	0.777778	0.666667

id_questionnaire	p_extent_of_problems	p_extent_of_satisfaction \
1	0.314815	0.611111
2	0.074074	0.888889
3	0.314815	0.666667
4	0.129630	0.555556
5	0.055556	0.722222
...
20426	0.129630	0.666667
20427	0.055556	0.777778
20428	0.156863	0.888889
20429	0.025641	1.000000
20430	0.129630	0.555556

id_questionnaire	p_island	p_number_of_school_tech_resources \
1	TENERIFE_CAPITAL	196.0
2	GRAN_CANARIA_CAPITAL	157.0
3	GRAN_CANARIA_CAPITAL	62.0
4	TENERIFE_CAPITAL	131.0
5	TENERIFE_CAPITAL	118.0
...
20426	GRAN_CANARIA_PROVINCE	264.0
20427	TENERIFE_CAPITAL	178.0
20428	GRAN_CANARIA_CAPITAL	311.0
20429	TENERIFE_CAPITAL	30.0
20430	TENERIFE_CAPITAL	54.0

p_public_or_private	level_MAT	level_LEN
---------------------	-----------	-----------

id_questionnaire			
1	PUBLIC	3.0	3.0
2	PUBLIC	1.0	1.0
3	PRIVATE	1.0	3.0
4	PUBLIC	2.0	2.0
5	PRIVATE	4.0	3.0
...
20426	PUBLIC	1.0	1.0
20427	PUBLIC	2.0	3.0
20428	PUBLIC	2.0	2.0
20429	PRIVATE	3.0	3.0
20430	PUBLIC	2.0	3.0

[17806 rows x 34 columns]

	p_gender \
p_gender	1.000000
p_number_of_teachers_in_school	0.159623
p_percentage_of_teachers_changed_school_last_year	-0.045782
p_attitude_teacher_training_courses	-0.155428
p_age	-0.024211
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.020059
p_number_of_years_in_school	0.047504
p_number_of_years_as_principal	0.167834
p_number_of_years_as_principal_in_school	0.155515
p_number_of_principals_10_years	-0.091932
p_number_of_preschool_students	0.146669
p_number_of_preschool_groups	0.107065
p_number_of_primary_students	0.128383
p_number_of_primary_groups	0.160362
p_extent_of_agreement	-0.108918
p_extent_of_factors_limiting_effectiveness	0.038589
p_extent_of_inconveniences	-0.067401
p_extent_of_management	-0.176006
p_extent_of_problems	-0.056263
p_extent_of_satisfaction	-0.095692
p_island	0.009655
p_number_of_school_tech_resources	0.111593
p_public_or_private	-0.099218

level_MAT	0.053646
level_LEN	0.044194

p_number_of_teachers_in_school \	
p_gender	0.159623
p_number_of_teachers_in_school	1.000000
p_percentage_of_teachers_changed_school_last_year	-0.185265
p_attitude_teacher_training_courses	-0.215315
p_age	-0.071810
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.094354
p_number_of_years_in_school	0.142224
p_number_of_years_as_principal	0.161235
p_number_of_years_as_principal_in_school	0.165651
p_number_of_principals_10_years	-0.066192
p_number_of_preschool_students	0.880144
p_number_of_preschool_groups	0.802674
p_number_of_primary_students	0.898519
p_number_of_primary_groups	0.794199
p_extent_of_agreement	0.053070
p_extent_of_factors_limiting_effectiveness	-0.061530
p_extent_of_inconveniences	-0.240613
p_extent_of_management	-0.027679
p_extent_of_problems	-0.054402
p_extent_of_satisfaction	-0.041268
p_island	0.107208
p_number_of_school_tech_resources	0.556685
p_public_or_private	-0.547485
level_MAT	0.102478
level_LEN	0.085440

p_percentage_of_teachers_changed_school_last_year \	
p_gender	
-0.045782	
p_number_of_teachers_in_school	
-0.185265	

p_percentage_of_teachers_changed_school_last_year	1.000000	␣
p_attitude_teacher_training_courses	0.201196	␣
p_age	-0.004783	␣
p_group_criteria_random	NaN	␣
p_group_criteria_subjects	NaN	␣
p_group_criteria_other	NaN	␣
p_group_criteria_alphabet	NaN	␣
p_group_criteria_gender	NaN	␣
p_group_criteria_language	NaN	␣
p_group_criteria_performance	NaN	␣
p_group_criteria_homogeneity	NaN	␣
p_group_criteria_heterogeneity	NaN	␣
p_number_of_years_of_teaching	-0.044489	␣
p_number_of_years_in_school	-0.134029	␣
p_number_of_years_as_principal	-0.091029	␣
p_number_of_years_as_principal_in_school	-0.114515	␣
p_number_of_principals_10_years	0.110425	␣
p_number_of_preschool_students	-0.151578	␣
p_number_of_preschool_groups	-0.060950	␣
p_number_of_primary_students	-0.162638	␣
p_number_of_primary_groups	-0.061627	␣
p_extent_of_agreement	-0.221886	␣
p_extent_of_factors_limiting_effectiveness	0.212428	␣

p_extent_of_inconveniences		□
↪	0.281746	
p_extent_of_management		□
↪	-0.120551	
p_extent_of_problems		□
↪	0.262679	
p_extent_of_satisfaction		□
↪	-0.113346	
p_island		□
↪	0.135192	
p_number_of_school_tech_resources		□
↪	0.004551	
p_public_or_private		□
↪	0.434857	
level_MAT		□
↪	-0.132012	
level_LEN		□
↪	-0.124095	
↪p_attitude_teacher_training_courses \		□
p_gender		-0.
↪155428		
p_number_of_teachers_in_school		-0.
↪215315		
p_percentage_of_teachers_changed_school_last_year		0.
↪201196		
p_attitude_teacher_training_courses		1.
↪000000		
p_age		0.
↪106640		
p_group_criteria_random		□
↪ NaN		
p_group_criteria_subjects		□
↪ NaN		
p_group_criteria_other		□
↪ NaN		
p_group_criteria_alphabet		□
↪ NaN		
p_group_criteria_gender		□
↪ NaN		
p_group_criteria_language		□
↪ NaN		
p_group_criteria_performance		□
↪ NaN		

p_group_criteria_homogeneity	␣	NaN	␣
p_group_criteria_heterogeneity	␣	NaN	␣
p_number_of_years_of_teaching	␣	148419	0.
p_number_of_years_in_school	␣	017325	-0.
p_number_of_years_as_principal	␣	016209	0.
p_number_of_years_as_principal_in_school	␣	008065	0.
p_number_of_principals_10_years	␣	014383	-0.
p_number_of_preschool_students	␣	136417	-0.
p_number_of_preschool_groups	␣	144948	-0.
p_number_of_primary_students	␣	121403	-0.
p_number_of_primary_groups	␣	076768	-0.
p_extent_of_agreement	␣	059628	-0.
p_extent_of_factors_limiting_effectiveness	␣	044016	-0.
p_extent_of_inconveniences	␣	101608	0.
p_extent_of_management	␣	010726	0.
p_extent_of_problems	␣	083005	0.
p_extent_of_satisfaction	␣	173843	0.
p_island	␣	087793	-0.
p_number_of_school_tech_resources	␣	059893	0.
p_public_or_private	␣	444140	0.
level_MAT	␣	053197	-0.
level_LEN	␣	061591	-0.

p_age \

p_gender	-0.024211
p_number_of_teachers_in_school	-0.071810
p_percentage_of_teachers_changed_school_last_year	-0.004783
p_attitude_teacher_training_courses	0.106640
p_age	1.000000
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.797119
p_number_of_years_in_school	0.479378
p_number_of_years_as_principal	0.529904
p_number_of_years_as_principal_in_school	0.496292
p_number_of_principals_10_years	-0.336140
p_number_of_preschool_students	-0.024041
p_number_of_preschool_groups	-0.061992
p_number_of_primary_students	-0.023715
p_number_of_primary_groups	-0.027358
p_extent_of_agreement	0.016133
p_extent_of_factors_limiting_effectiveness	-0.066245
p_extent_of_inconveniences	-0.045046
p_extent_of_management	-0.085886
p_extent_of_problems	-0.030693
p_extent_of_satisfaction	0.014912
p_island	0.003910
p_number_of_school_tech_resources	0.061936
p_public_or_private	0.072827
level_MAT	0.044703
level_LEN	0.037460

	p_group_criteria_random \
p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN

p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_subjects \
p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN

p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_other \
p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

p_group_criteria_alphabet \

p_gender	NaN
p_number_of_teachers_in_school	NaN
p_percentage_of_teachers_changed_school_last_year	NaN
p_attitude_teacher_training_courses	NaN
p_age	NaN
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	p_group_criteria_gender \	
p_gender		NaN
p_number_of_teachers_in_school		NaN
p_percentage_of_teachers_changed_school_last_year		NaN
p_attitude_teacher_training_courses		NaN
p_age		NaN
p_group_criteria_random		NaN
p_group_criteria_subjects		NaN
p_group_criteria_other		NaN
p_group_criteria_alphabet		NaN
p_group_criteria_gender		NaN
p_group_criteria_language		NaN
p_group_criteria_performance		NaN

p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	NaN
p_number_of_years_in_school	NaN
p_number_of_years_as_principal	NaN
p_number_of_years_as_principal_in_school	NaN
p_number_of_principals_10_years	NaN
p_number_of_preschool_students	NaN
p_number_of_preschool_groups	NaN
p_number_of_primary_students	NaN
p_number_of_primary_groups	NaN
p_extent_of_agreement	NaN
p_extent_of_factors_limiting_effectiveness	NaN
p_extent_of_inconveniences	NaN
p_extent_of_management	NaN
p_extent_of_problems	NaN
p_extent_of_satisfaction	NaN
p_island	NaN
p_number_of_school_tech_resources	NaN
p_public_or_private	NaN
level_MAT	NaN
level_LEN	NaN

	...	\
p_gender	...	
p_number_of_teachers_in_school	...	
p_percentage_of_teachers_changed_school_last_year	...	
p_attitude_teacher_training_courses	...	
p_age	...	
p_group_criteria_random	...	
p_group_criteria_subjects	...	
p_group_criteria_other	...	
p_group_criteria_alphabet	...	
p_group_criteria_gender	...	
p_group_criteria_language	...	
p_group_criteria_performance	...	
p_group_criteria_homogeneity	...	
p_group_criteria_heterogeneity	...	
p_number_of_years_of_teaching	...	
p_number_of_years_in_school	...	
p_number_of_years_as_principal	...	
p_number_of_years_as_principal_in_school	...	
p_number_of_principals_10_years	...	
p_number_of_preschool_students	...	
p_number_of_preschool_groups	...	
p_number_of_primary_students	...	
p_number_of_primary_groups	...	
p_extent_of_agreement	...	

p_extent_of_factors_limiting_effectiveness	...	
p_extent_of_inconveniences	...	
p_extent_of_management	...	
p_extent_of_problems	...	
p_extent_of_satisfaction	...	
p_island	...	
p_number_of_school_tech_resources	...	
p_public_or_private	...	
level_MAT	...	
level_LEN	...	
↪ p_extent_of_factors_limiting_effectiveness \		⌋
p_gender		⌋
↪ 0.038589		
p_number_of_teachers_in_school		⌋
↪ -0.061530		
p_percentage_of_teachers_changed_school_last_year		⌋
↪ 0.212428		
p_attitude_teacher_training_courses		⌋
↪ -0.044016		
p_age		⌋
↪ -0.066245		
p_group_criteria_random		⌋
↪ NaN		
p_group_criteria_subjects		⌋
↪ NaN		
p_group_criteria_other		⌋
↪ NaN		
p_group_criteria_alphabet		⌋
↪ NaN		
p_group_criteria_gender		⌋
↪ NaN		
p_group_criteria_language		⌋
↪ NaN		
p_group_criteria_performance		⌋
↪ NaN		
p_group_criteria_homogeneity		⌋
↪ NaN		
p_group_criteria_heterogeneity		⌋
↪ NaN		
p_number_of_years_of_teaching		⌋
↪ -0.073058		
p_number_of_years_in_school		⌋
↪ -0.160859		

p_number_of_years_as_principal		□
↪ -0.022741		
p_number_of_years_as_principal_in_school		□
↪ -0.075019		
p_number_of_principals_10_years		□
↪ 0.067971		
p_number_of_preschool_students		□
↪ -0.051689		
p_number_of_preschool_groups		□
↪ 0.000143		
p_number_of_primary_students		□
↪ -0.053022		
p_number_of_primary_groups		□
↪ -0.004780		
p_extent_of_agreement		□
↪ -0.330075		
p_extent_of_factors_limiting_effectiveness		□
↪ 1.000000		
p_extent_of_inconveniences		□
↪ 0.599354		
p_extent_of_management		□
↪ -0.263730		
p_extent_of_problems		□
↪ 0.448839		
p_extent_of_satisfaction		□
↪ -0.342985		
p_island		□
↪ 0.055340		
p_number_of_school_tech_resources		□
↪ -0.020285		
p_public_or_private		□
↪ 0.215288		
level_MAT		□
↪ -0.080427		
level_LEN		□
↪ -0.062271		

	p_extent_of_inconveniences \
p_gender	-0.067401
p_number_of_teachers_in_school	-0.240613
p_percentage_of_teachers_changed_school_last_year	0.281746
p_attitude_teacher_training_courses	0.101608
p_age	-0.045046
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN

p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.073439
p_number_of_years_in_school	-0.244957
p_number_of_years_as_principal	-0.201010
p_number_of_years_as_principal_in_school	-0.226483
p_number_of_principals_10_years	0.168138
p_number_of_preschool_students	-0.230243
p_number_of_preschool_groups	-0.150370
p_number_of_primary_students	-0.215991
p_number_of_primary_groups	-0.138480
p_extent_of_agreement	-0.340746
p_extent_of_factors_limiting_effectiveness	0.599354
p_extent_of_inconveniences	1.000000
p_extent_of_management	-0.162628
p_extent_of_problems	0.459965
p_extent_of_satisfaction	-0.311237
p_island	0.093214
p_number_of_school_tech_resources	-0.097926
p_public_or_private	0.454093
level_MAT	-0.142821
level_LEN	-0.121366

	p_extent_of_management \
p_gender	-0.176006
p_number_of_teachers_in_school	-0.027679
p_percentage_of_teachers_changed_school_last_year	-0.120551
p_attitude_teacher_training_courses	0.010726
p_age	-0.085886
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.035008
p_number_of_years_in_school	0.005647
p_number_of_years_as_principal	-0.108370
p_number_of_years_as_principal_in_school	-0.066178
p_number_of_principals_10_years	0.116240
p_number_of_preschool_students	-0.028844

p_number_of_preschool_groups	-0.016151
p_number_of_primary_students	-0.003104
p_number_of_primary_groups	-0.052893
p_extent_of_agreement	0.553430
p_extent_of_factors_limiting_effectiveness	-0.263730
p_extent_of_inconveniences	-0.162628
p_extent_of_management	1.000000
p_extent_of_problems	-0.232462
p_extent_of_satisfaction	0.337143
p_island	-0.082008
p_number_of_school_tech_resources	-0.074406
p_public_or_private	-0.079065
level_MAT	0.038399
level_LEN	0.074959

	p_extent_of_problems \
p_gender	-0.056263
p_number_of_teachers_in_school	-0.054402
p_percentage_of_teachers_changed_school_last_year	0.262679
p_attitude_teacher_training_courses	0.083005
p_age	-0.030693
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.089881
p_number_of_years_in_school	-0.154163
p_number_of_years_as_principal	-0.111545
p_number_of_years_as_principal_in_school	-0.128506
p_number_of_principals_10_years	0.064288
p_number_of_preschool_students	-0.086756
p_number_of_preschool_groups	-0.002107
p_number_of_primary_students	-0.065957
p_number_of_primary_groups	-0.008929
p_extent_of_agreement	-0.365932
p_extent_of_factors_limiting_effectiveness	0.448839
p_extent_of_inconveniences	0.459965
p_extent_of_management	-0.232462
p_extent_of_problems	1.000000
p_extent_of_satisfaction	-0.295867
p_island	-0.043104
p_number_of_school_tech_resources	0.012170
p_public_or_private	0.253034

level_MAT	-0.152743
level_LEN	-0.117485
	p_extent_of_satisfaction \
p_gender	-0.095692
p_number_of_teachers_in_school	-0.041268
p_percentage_of_teachers_changed_school_last_year	-0.113346
p_attitude_teacher_training_courses	0.173843
p_age	0.014912
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.064672
p_number_of_years_in_school	0.020068
p_number_of_years_as_principal	-0.059122
p_number_of_years_as_principal_in_school	-0.085462
p_number_of_principals_10_years	0.041289
p_number_of_preschool_students	-0.018821
p_number_of_preschool_groups	-0.032589
p_number_of_primary_students	-0.028711
p_number_of_primary_groups	-0.041391
p_extent_of_agreement	0.461188
p_extent_of_factors_limiting_effectiveness	-0.342985
p_extent_of_inconveniences	-0.311237
p_extent_of_management	0.337143
p_extent_of_problems	-0.295867
p_extent_of_satisfaction	1.000000
p_island	-0.049515
p_number_of_school_tech_resources	-0.009083
p_public_or_private	0.033937
level_MAT	0.014379
level_LEN	0.030505
	p_island \
p_gender	0.009655
p_number_of_teachers_in_school	0.107208
p_percentage_of_teachers_changed_school_last_year	0.135192
p_attitude_teacher_training_courses	-0.087793
p_age	0.003910
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN

p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	-0.099812
p_number_of_years_in_school	-0.041988
p_number_of_years_as_principal	0.059209
p_number_of_years_as_principal_in_school	0.081011
p_number_of_principals_10_years	0.018335
p_number_of_preschool_students	0.072474
p_number_of_preschool_groups	0.085912
p_number_of_primary_students	0.066444
p_number_of_primary_groups	0.063813
p_extent_of_agreement	-0.057438
p_extent_of_factors_limiting_effectiveness	0.055340
p_extent_of_inconveniences	0.093214
p_extent_of_management	-0.082008
p_extent_of_problems	-0.043104
p_extent_of_satisfaction	-0.049515
p_island	1.000000
p_number_of_school_tech_resources	0.027829
p_public_or_private	-0.074768
level_MAT	0.044030
level_LEN	-0.001744

p_number_of_school_tech_resources \	
p_gender	0.
111593	
p_number_of_teachers_in_school	0.
556685	
p_percentage_of_teachers_changed_school_last_year	0.
004551	
p_attitude_teacher_training_courses	0.
059893	
p_age	0.
061936	
p_group_criteria_random	
NaN	
p_group_criteria_subjects	
NaN	
p_group_criteria_other	
NaN	
p_group_criteria_alphabet	
NaN	

p_group_criteria_gender	↪NaN	↪
p_group_criteria_language	↪NaN	↪
p_group_criteria_performance	↪NaN	↪
p_group_criteria_homogeneity	↪NaN	↪
p_group_criteria_heterogeneity	↪NaN	↪
p_number_of_years_of_teaching	↪032275	0.
p_number_of_years_in_school	↪113806	0.
p_number_of_years_as_principal	↪093311	0.
p_number_of_years_as_principal_in_school	↪084736	0.
p_number_of_principals_10_years	↪095084	-0.
p_number_of_preschool_students	↪535117	0.
p_number_of_preschool_groups	↪477388	0.
p_number_of_primary_students	↪550508	0.
p_number_of_primary_groups	↪512754	0.
p_extent_of_agreement	↪029804	-0.
p_extent_of_factors_limiting_effectiveness	↪020285	-0.
p_extent_of_inconveniences	↪097926	-0.
p_extent_of_management	↪074406	-0.
p_extent_of_problems	↪012170	0.
p_extent_of_satisfaction	↪009083	-0.
p_island	↪027829	0.
p_number_of_school_tech_resources	↪000000	1.
p_public_or_private	↪025741	0.

level_MAT	-0.
↪028566	
level_LEN	0.
↪005489	

	p_public_or_private \
p_gender	-0.099218
p_number_of_teachers_in_school	-0.547485
p_percentage_of_teachers_changed_school_last_year	0.434857
p_attitude_teacher_training_courses	0.444140
p_age	0.072827
p_group_criteria_random	NaN
p_group_criteria_subjects	NaN
p_group_criteria_other	NaN
p_group_criteria_alphabet	NaN
p_group_criteria_gender	NaN
p_group_criteria_language	NaN
p_group_criteria_performance	NaN
p_group_criteria_homogeneity	NaN
p_group_criteria_heterogeneity	NaN
p_number_of_years_of_teaching	0.098537
p_number_of_years_in_school	-0.199116
p_number_of_years_as_principal	-0.209854
p_number_of_years_as_principal_in_school	-0.223973
p_number_of_principals_10_years	0.028208
p_number_of_preschool_students	-0.433325
p_number_of_preschool_groups	-0.319348
p_number_of_primary_students	-0.391661
p_number_of_primary_groups	-0.228892
p_extent_of_agreement	-0.198010
p_extent_of_factors_limiting_effectiveness	0.215288
p_extent_of_inconveniences	0.454093
p_extent_of_management	-0.079065
p_extent_of_problems	0.253034
p_extent_of_satisfaction	0.033937
p_island	-0.074768
p_number_of_school_tech_resources	0.025741
p_public_or_private	1.000000
level_MAT	-0.182115
level_LEN	-0.141777

	level_MAT	level_LEN
p_gender	0.053646	0.044194
p_number_of_teachers_in_school	0.102478	0.085440
p_percentage_of_teachers_changed_school_last_year	-0.132012	-0.124095
p_attitude_teacher_training_courses	-0.053197	-0.061591
p_age	0.044703	0.037460

p_group_criteria_random	NaN	NaN
p_group_criteria_subjects	NaN	NaN
p_group_criteria_other	NaN	NaN
p_group_criteria_alphabet	NaN	NaN
p_group_criteria_gender	NaN	NaN
p_group_criteria_language	NaN	NaN
p_group_criteria_performance	NaN	NaN
p_group_criteria_homogeneity	NaN	NaN
p_group_criteria_heterogeneity	NaN	NaN
p_number_of_years_of_teaching	0.059156	0.035318
p_number_of_years_in_school	0.089493	0.060077
p_number_of_years_as_principal	0.118076	0.085325
p_number_of_years_as_principal_in_school	0.120255	0.089276
p_number_of_principals_10_years	-0.063751	-0.033567
p_number_of_preschool_students	0.106308	0.085683
p_number_of_preschool_groups	0.084685	0.059827
p_number_of_primary_students	0.097782	0.083027
p_number_of_primary_groups	0.074565	0.048145
p_extent_of_agreement	0.089626	0.096205
p_extent_of_factors_limiting_effectiveness	-0.080427	-0.062271
p_extent_of_inconveniences	-0.142821	-0.121366
p_extent_of_management	0.038399	0.074959
p_extent_of_problems	-0.152743	-0.117485
p_extent_of_satisfaction	0.014379	0.030505
p_island	0.044030	-0.001744
p_number_of_school_tech_resources	-0.028566	0.005489
p_public_or_private	-0.182115	-0.141777
level_MAT	1.000000	0.562532
level_LEN	0.562532	1.000000

[34 rows x 34 columns]

```
[86]:      p_gender  p_number_of_teachers_in_school  \
id_questionnaire
1      MALE      19.0
2      FEMALE     26.0
3      FEMALE     42.0
4      MALE      15.0
5      MALE      22.0
...      ...      ...
20426   FEMALE     24.0
20427   FEMALE     15.0
20428   MALE      27.0
20429   FEMALE     23.0
20430   MALE      15.0

      p_percentage_of_teachers_changed_school_last_year  \
```

id_questionnaire	
1	3.0
2	2.0
3	2.0
4	1.0
5	1.0
...	...
20426	4.0
20427	2.0
20428	1.0
20429	1.0
20430	3.0

	p_attitude_teacher_training_courses	p_age \
id_questionnaire		
1	3.0	50.0
2	3.0	58.0
3	2.0	45.0
4	3.0	46.0
5	2.0	56.0
...
20426	3.0	46.0
20427	3.0	49.0
20428	3.0	59.0
20429	4.0	64.0
20430	2.0	51.0

	p_group_criteria_random	p_group_criteria_subjects \
id_questionnaire		
1	False	False
2	False	False
3	False	False
4	False	False
5	False	False
...
20426	False	False
20427	False	False
20428	False	False
20429	False	False
20430	False	False

	p_group_criteria_other	p_group_criteria_alphabet \
id_questionnaire		
1	False	False
2	False	False
3	False	False
4	False	False

5	False	False
...
20426	False	False
20427	False	False
20428	False	False
20429	False	False
20430	False	False

	p_group_criteria_gender	\
id_questionnaire		
1	False	...
2	False	...
3	False	...
4	False	...
5	False	...
...
20426	False	...
20427	False	...
20428	False	...
20429	False	...
20430	False	...

	p_extent_of_factors_limiting_effectiveness	\
id_questionnaire		
1	0.666667	
2	0.375000	
3	0.500000	
4	0.625000	
5	0.458333	
...	...	
20426	0.583333	
20427	0.333333	
20428	0.458333	
20429	0.083333	
20430	0.583333	

	p_extent_of_inconveniences	p_extent_of_management	\
id_questionnaire			
1	0.750000	0.666667	
2	0.305556	0.857143	
3	0.606061	0.666667	
4	0.666667	0.809524	
5	0.333333	0.714286	
...	
20426	0.555556	0.952381	
20427	0.416667	0.761905	
20428	0.388889	0.809524	

20429	0.000000	1.000000
20430	0.777778	0.666667

	p_extent_of_problems	p_extent_of_satisfaction \
id_questionnaire		
1	0.314815	0.611111
2	0.074074	0.888889
3	0.314815	0.666667
4	0.129630	0.555556
5	0.055556	0.722222
...
20426	0.129630	0.666667
20427	0.055556	0.777778
20428	0.156863	0.888889
20429	0.025641	1.000000
20430	0.129630	0.555556

	p_island	p_number_of_school_tech_resources \
id_questionnaire		
1	TENERIFE_CAPITAL	196.0
2	GRAN_CANARIA_CAPITAL	157.0
3	GRAN_CANARIA_CAPITAL	62.0
4	TENERIFE_CAPITAL	131.0
5	TENERIFE_CAPITAL	118.0
...
20426	GRAN_CANARIA_PROVINCE	264.0
20427	TENERIFE_CAPITAL	178.0
20428	GRAN_CANARIA_CAPITAL	311.0
20429	TENERIFE_CAPITAL	30.0
20430	TENERIFE_CAPITAL	54.0

	p_public_or_private	level_MAT	level_LEN
id_questionnaire			
1	PUBLIC	3.0	3.0
2	PUBLIC	1.0	1.0
3	PRIVATE	1.0	3.0
4	PUBLIC	2.0	2.0
5	PRIVATE	4.0	3.0
...
20426	PUBLIC	1.0	1.0
20427	PUBLIC	2.0	3.0
20428	PUBLIC	2.0	2.0
20429	PRIVATE	3.0	3.0
20430	PUBLIC	2.0	3.0

[17806 rows x 34 columns]

6 Feature Importance

```
[87]: target_feature = "level_MAT"
```

6.1 Utils

```
[88]: get_features = lambda curr_df: [col for col in curr_df.columns if not col.
    ↳startswith("level_")]
```

6.2 Core

```
[89]: get_features = lambda curr_df: [col for col in curr_df.columns if not col.
    ↳startswith("level_")]
concatenated_df = pd.concat(
    [
        df_sel_gen[["id_class_group", "id_school", target_feature]],
        df_sel_students[get_features(df_sel_students)],
        df_sel_families[get_features(df_sel_families)],
        df_sel_teachers[get_features(df_sel_teachers)],
        df_sel_principals[get_features(df_sel_principals)]
    ],
    axis=1
)
X, y = concatenated_df[get_features(concatenated_df)],
    ↳concatenated_df[target_feature]
X
```

```
[89]:
```

	id_class_group	id_school	s_gender	s_birth_year	\
id_questionnaire					
1	A	2415.0	FEMALE	2007.0	
2	A	1842.0	MALE	2007.0	
3	NaN	1432.0	FEMALE	2007.0	
4	A	2280.0	MALE	2007.0	
5	A	2040.0	FEMALE	2007.0	
...	
20426	NaN	1863.0	FEMALE	2001.0	
20427	A	2203.0	FEMALE	2007.0	
20428	B	1368.0	MALE	2001.0	
20429	A	2373.0	MALE	2007.0	
20430	NaN	1907.0	FEMALE	2007.0	

	s_is_living_with_mother	s_has_repeated	s_frequency_of_skips	\
id_questionnaire				
1	True	False	1.000000	
2	True	True	1.000000	
3	True	False	1.000000	
4	True	False	1.000000	

5	True	False	1.000000
...
20426	True	NaN	0.000000
20427	True	False	1.000000
20428	True	NaN	0.333333
20429	True	False	1.000000
20430	True	False	1.000000

	s_homeworks_day_per_week	s_extent_of_teachers_affinity	\
id_questionnaire			
1	0.0	1.0	
2	3.0	1.0	
3	2.0	NaN	
4	3.0	1.0	
5	1.0	1.0	
...	
20426	2.0	NaN	
20427	3.0	1.0	
20428	4.0	NaN	
20429	3.0	1.0	
20430	2.0	1.0	

	s_frequency_of_computer_usage	...	\
id_questionnaire		...	
1	0.111111	...	
2	0.333333	...	
3	0.666667	...	
4	0.333333	...	
5	0.000000	...	
...	
20426	NaN	...	
20427	0.000000	...	
20428	0.222222	...	
20429	0.000000	...	
20430	0.111111	...	

	p_number_of_primary_groups	p_extent_of_agreement	\
id_questionnaire			
1	10.0	0.611111	
2	12.0	0.666667	
3	12.0	0.666667	
4	6.0	0.444444	
5	6.0	0.722222	
...	
20426	10.0	0.444444	
20427	6.0	0.611111	
20428	12.0	0.722222	

20429	6.0	0.888889
20430	6.0	0.666667

	p_extent_of_factors_limiting_effectiveness \
id_questionnaire	
1	0.666667
2	0.375000
3	0.500000
4	0.625000
5	0.458333
...	...
20426	0.583333
20427	0.333333
20428	0.458333
20429	0.083333
20430	0.583333

	p_extent_of_inconveniences	p_extent_of_management \
id_questionnaire		
1	0.750000	0.666667
2	0.305556	0.857143
3	0.606061	0.666667
4	0.666667	0.809524
5	0.333333	0.714286
...
20426	0.555556	0.952381
20427	0.416667	0.761905
20428	0.388889	0.809524
20429	0.000000	1.000000
20430	0.777778	0.666667

	p_extent_of_problems	p_extent_of_satisfaction \
id_questionnaire		
1	0.314815	0.611111
2	0.074074	0.888889
3	0.314815	0.666667
4	0.129630	0.555556
5	0.055556	0.722222
...
20426	0.129630	0.666667
20427	0.055556	0.777778
20428	0.156863	0.888889
20429	0.025641	1.000000
20430	0.129630	0.555556

	p_island	p_number_of_school_tech_resources \
id_questionnaire		

1	TENERIFE_CAPITAL	196.0
2	GRAN_CANARIA_CAPITAL	157.0
3	GRAN_CANARIA_CAPITAL	62.0
4	TENERIFE_CAPITAL	131.0
5	TENERIFE_CAPITAL	118.0
...
20426	GRAN_CANARIA_PROVINCE	264.0
20427	TENERIFE_CAPITAL	178.0
20428	GRAN_CANARIA_CAPITAL	311.0
20429	TENERIFE_CAPITAL	30.0
20430	TENERIFE_CAPITAL	54.0

	p_public_or_private
id_questionnaire	
1	PUBLIC
2	PUBLIC
3	PRIVATE
4	PUBLIC
5	PRIVATE
...	...
20426	PUBLIC
20427	PUBLIC
20428	PUBLIC
20429	PRIVATE
20430	PUBLIC

[17806 rows x 100 columns]

```
[90]: forest = RandomForestClassifier(random_state=0)
forest.fit(discretize_columns(X), y)
```

```
[90]: RandomForestClassifier(random_state=0)
```

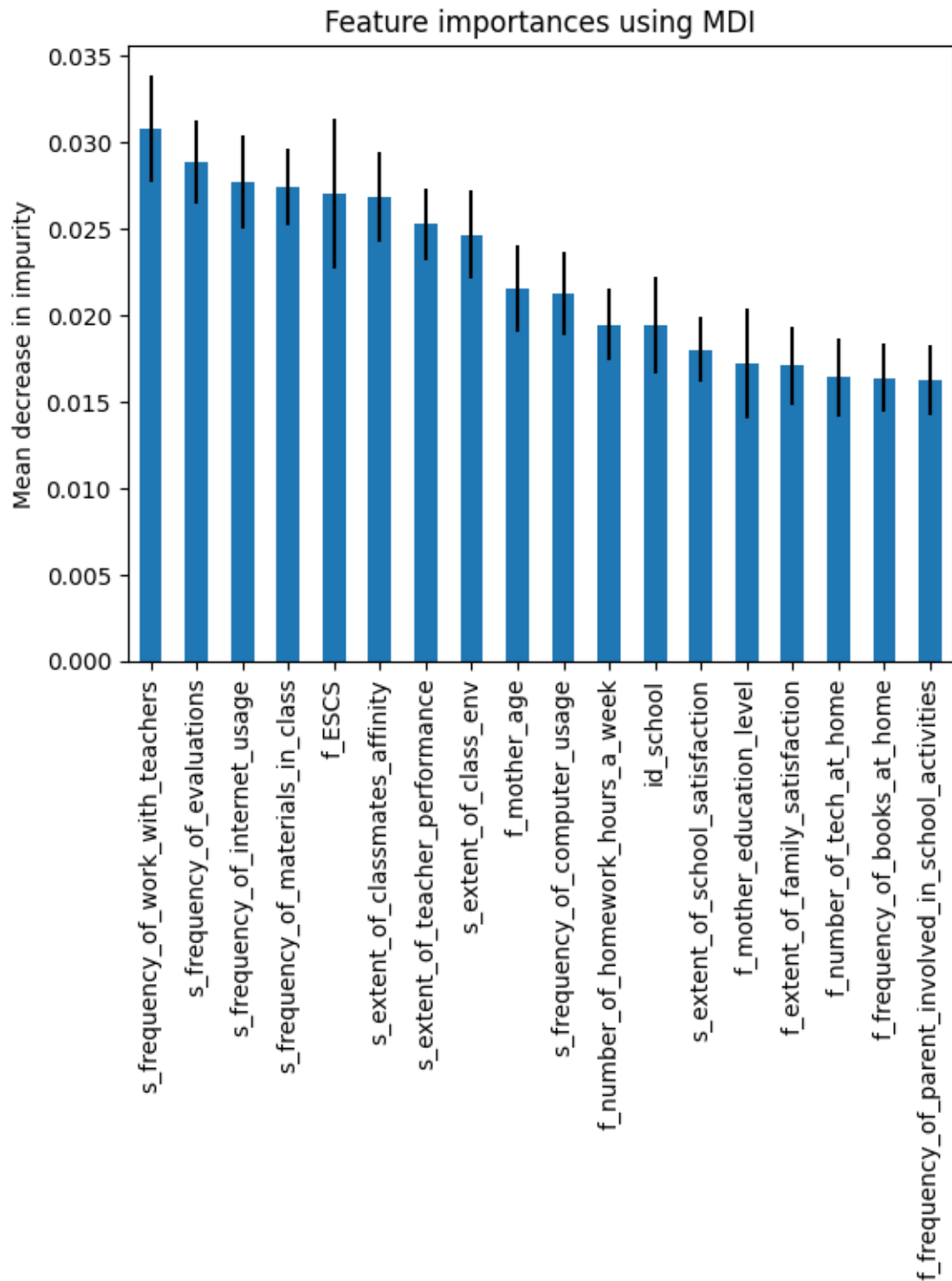
```
[91]: importances = forest.feature_importances_
std = np.std([tree.feature_importances_ for tree in forest.estimators_], axis=0)
```

```
[92]: forest_importances = pd.Series(importances, index=get_features(concatenated_df))
forest_std = pd.Series(std, index=get_features(concatenated_df))
forest_df = pd.concat([forest_importances, forest_std], axis=1)
forest_df.columns = ["importance", "std"]
forest_df = forest_df[forest_df["importance"] > 0.015]
forest_df = forest_df.sort_values("importance", ascending=False)
forest_importances = forest_df["importance"]
std = forest_df["std"].to_numpy()

fig, ax = plt.subplots()
forest_importances.plot.bar(yerr=std, ax=ax)
```

```
ax.set_title("Feature importances using MDI")
ax.set_ylabel("Mean decrease in impurity")
# fig.tight_layout()
```

```
[92]: Text(0, 0.5, 'Mean decrease in impurity')
```



```
[93]: final_features = [col for col in list(forest_importances.index) if col != "id_school"] + ["s_gender"]
final_features
```

```
[93]: ['s_frequency_of_work_with_teachers',
's_frequency_of_evaluations',
's_frequency_of_internet_usage',
's_frequency_of_materials_in_class',
'f_ESCS',
's_extent_of_classmates_affinity',
's_extent_of_teacher_performance',
's_extent_of_class_env',
'f_mother_age',
's_frequency_of_computer_usage',
'f_number_of_homework_hours_a_week',
's_extent_of_school_satisfaction',
'f_mother_education_level',
'f_extent_of_family_satisfaction',
'f_number_of_tech_at_home',
'f_frequency_of_books_at_home',
'f_frequency_of_parent_involved_in_school_activities',
's_gender']
```

7 Bias detection

```
[94]: final_df = concatenated_df[final_features + [target_feature]]
final_df
```

```
[94]:
```

id_questionnaire	s_frequency_of_work_with_teachers \	s_frequency_of_evaluations	s_frequency_of_internet_usage \
1	0.733333		
2	0.533333		
3	0.833333		
4	0.400000		
5	0.566667		
...	...		
20426	0.333333		
20427	0.666667		
20428	0.518519		
20429	0.633333		
20430	0.629630		

1	0.740741	0.066667
2	0.888889	0.600000
3	0.750000	0.666667
4	0.541667	0.266667
5	0.888889	0.133333
...
20426	0.333333	0.333333
20427	0.740741	0.133333
20428	0.740741	0.933333
20429	0.740741	0.533333
20430	0.481481	0.400000

	s_frequency_of_materials_in_class	f_ESCS \
id_questionnaire		
1	0.380952	0.235340
2	0.714286	NaN
3	1.000000	0.261451
4	NaN	0.787122
5	0.666667	3.151773
...
20426	NaN	-1.627731
20427	0.476190	1.372627
20428	0.777778	NaN
20429	0.142857	NaN
20430	0.476190	NaN

	s_extent_of_classmates_affinity \
id_questionnaire	
1	0.714286
2	0.714286
3	0.833333
4	0.857143
5	0.952381
...	...
20426	0.333333
20427	0.857143
20428	0.666667
20429	0.714286
20430	0.619048

	s_extent_of_teacher_performance	s_extent_of_class_env \
id_questionnaire		
1	0.833333	0.791667
2	0.800000	0.791667
3	0.800000	0.833333
4	0.800000	0.750000
5	1.000000	1.000000

...
20426	0.333333	0.333333
20427	0.900000	0.916667
20428	0.566667	0.958333
20429	1.000000	0.875000
20430	0.733333	0.708333

	f_mother_age	s_frequency_of_computer_usage \
id_questionnaire		
1	43.0	0.111111
2	NaN	0.333333
3	45.0	0.666667
4	39.0	0.333333
5	39.0	0.000000
...
20426	31.0	NaN
20427	37.0	0.000000
20428	NaN	0.222222
20429	NaN	0.000000
20430	NaN	0.111111

	f_number_of_homework_hours_a_week \
id_questionnaire	
1	14.0
2	NaN
3	9.0
4	9.0
5	6.0
...	...
20426	4.0
20427	1.0
20428	NaN
20429	NaN
20430	NaN

	s_extent_of_school_satisfaction	f_mother_education_level \
id_questionnaire		
1	0.944444	7.0
2	1.000000	NaN
3	0.888889	3.0
4	0.722222	5.0
5	1.000000	9.0
...
20426	1.000000	3.0
20427	1.000000	5.0
20428	0.833333	NaN
20429	1.000000	NaN

20430	0.722222	NaN
-------	----------	-----

	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
id_questionnaire		
1	0.700000	3.0
2	NaN	NaN
3	0.291667	4.0
4	1.000000	3.0
5	0.833333	7.0
...
20426	0.791667	4.0
20427	0.291667	5.0
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

	f_frequency_of_books_at_home \
id_questionnaire	
1	1.000000
2	NaN
3	0.444444
4	0.333333
5	0.555556
...	...
20426	0.333333
20427	0.444444
20428	NaN
20429	NaN
20430	NaN

	f_frequency_of_parent_involved_in_school_activities \
id_questionnaire	
1	0.333333
2	NaN
3	0.166667
4	0.250000
5	0.250000
...	...
20426	0.000000
20427	0.000000
20428	NaN
20429	NaN
20430	NaN

	s_gender	level_MAT
id_questionnaire		
1	FEMALE	3.0

2	MALE	1.0
3	FEMALE	1.0
4	MALE	2.0
5	FEMALE	4.0
...
20426	FEMALE	1.0
20427	FEMALE	2.0
20428	MALE	2.0
20429	MALE	3.0
20430	FEMALE	2.0

[17806 rows x 19 columns]

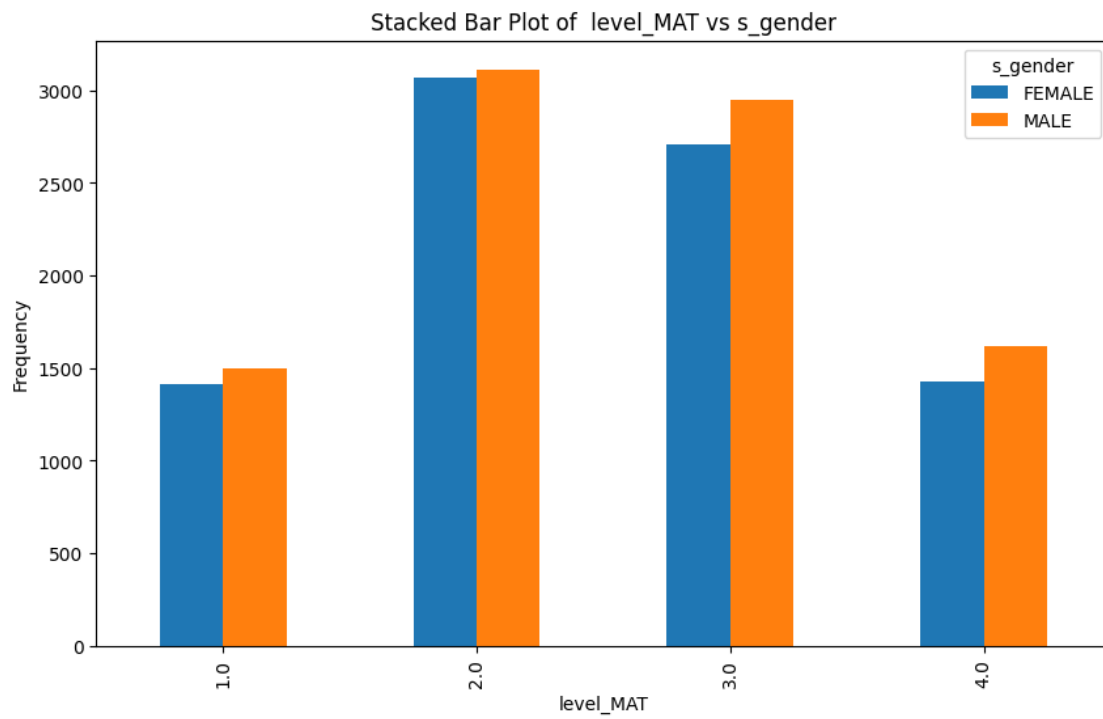
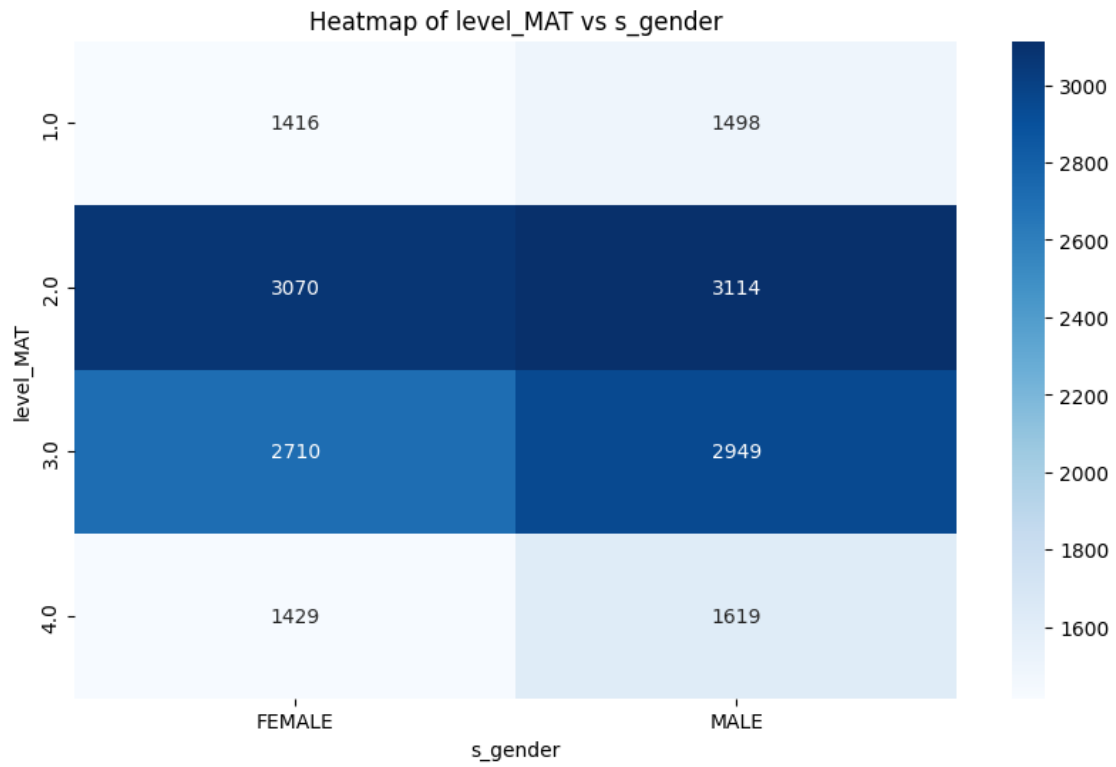
```
[187]: sensitive_feature = "s_gender"

trial_df = final_df[[target_feature, sensitive_feature]]
trial_df = trial_df.dropna()
col_a = trial_df[target_feature]
col_b = trial_df[sensitive_feature]

crosstab = pd.crosstab(col_a, col_b)

# Heatmap of Frequency
plt.figure(figsize=(10, 6))
sns.heatmap(crosstab, annot=True, fmt='d', cmap='Blues')
plt.title(f'Heatmap of {target_feature} vs {sensitive_feature}')
plt.xlabel(sensitive_feature)
plt.ylabel(target_feature)
plt.show()

# Clustered Bar Plot
crosstab.plot(kind='bar', stacked=False, figsize=(10, 6))
plt.title(f'Stacked Bar Plot of {target_feature} vs {sensitive_feature}')
plt.xlabel(target_feature)
plt.ylabel('Frequency')
plt.legend(title=sensitive_feature)
plt.show()
```



```

[101]: sensitive_feature = "f_ESCS"
        favorable_sensitive_label = "ADVANTAGED"
        unfavorable_sensitive_label = "DISADVANTAGED"
        favorable_class_label = "PASSED"
        unfavorable_class_label = "NOT PASSED"

        trial_df = final_df[[target_feature, sensitive_feature]]
        trial_df = trial_df.dropna()
        trial_df[sensitive_feature] = trial_df[sensitive_feature].apply(lambda elem:
            ↪ favorable_sensitive_label if elem > 1.5 else unfavorable_sensitive_label)
        trial_df[target_feature] = trial_df[target_feature].apply(lambda elem:
            ↪ favorable_class_label if elem >= 3.5 else unfavorable_class_label)

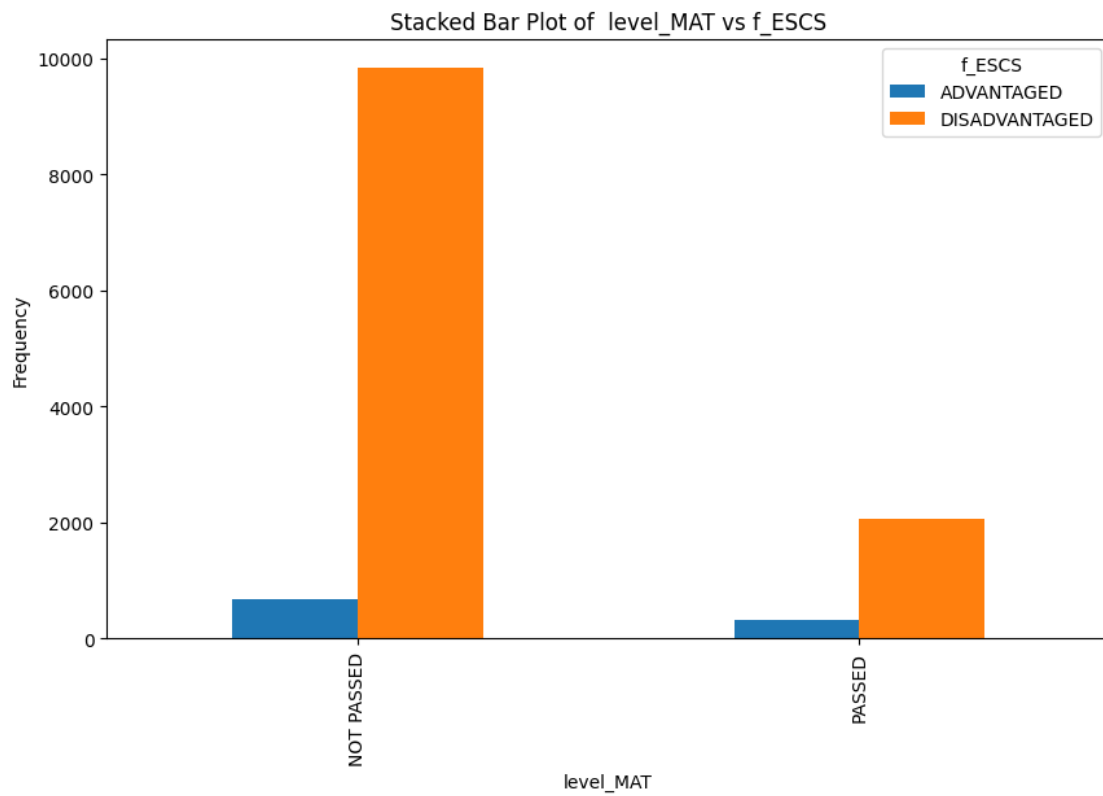
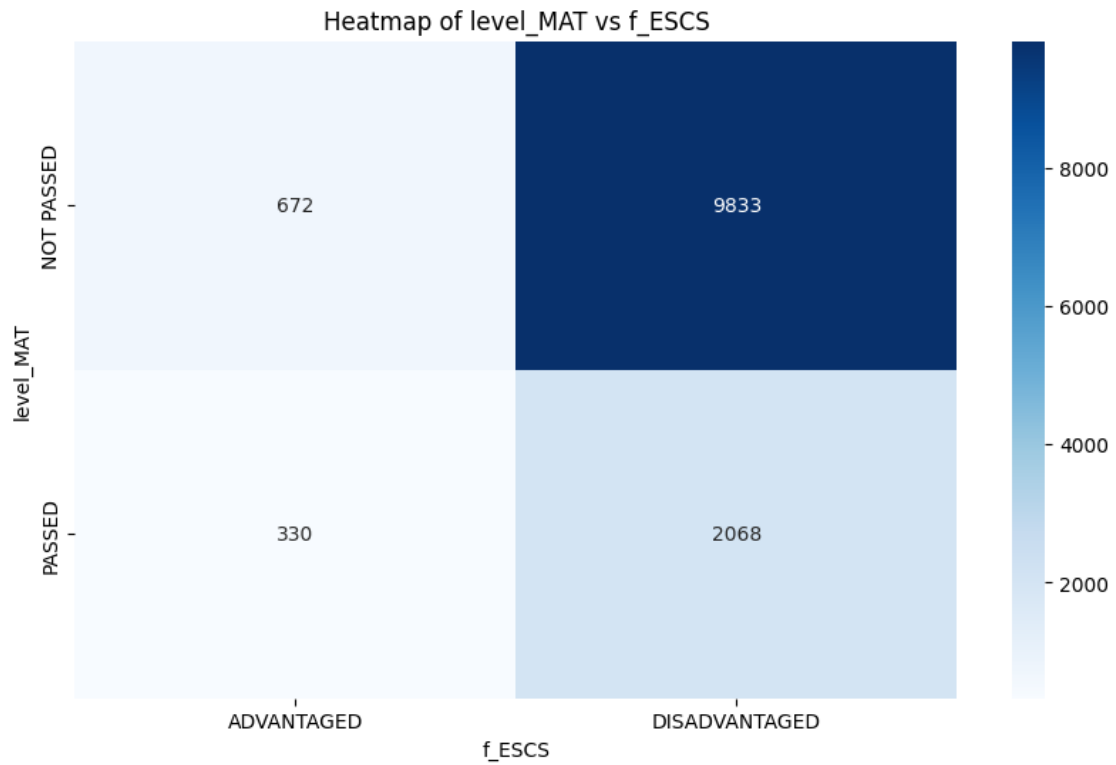
        col_a = trial_df[target_feature]
        col_b = trial_df[sensitive_feature]

        crosstab = pd.crosstab(col_a, col_b)

        # Heatmap of Frequency
        plt.figure(figsize=(10, 6))
        sns.heatmap(crosstab, annot=True, fmt='d', cmap='Blues')
        plt.title(f'Heatmap of {target_feature} vs {sensitive_feature}')
        plt.xlabel(sensitive_feature)
        plt.ylabel(target_feature)
        plt.show()

        # Clustered Bar Plot
        crosstab.plot(kind='bar', stacked=False, figsize=(10, 6))
        plt.title(f'Stacked Bar Plot of {target_feature} vs {sensitive_feature}')
        plt.xlabel(target_feature)
        plt.ylabel('Frequency')
        plt.legend(title=sensitive_feature)
        plt.show()

```

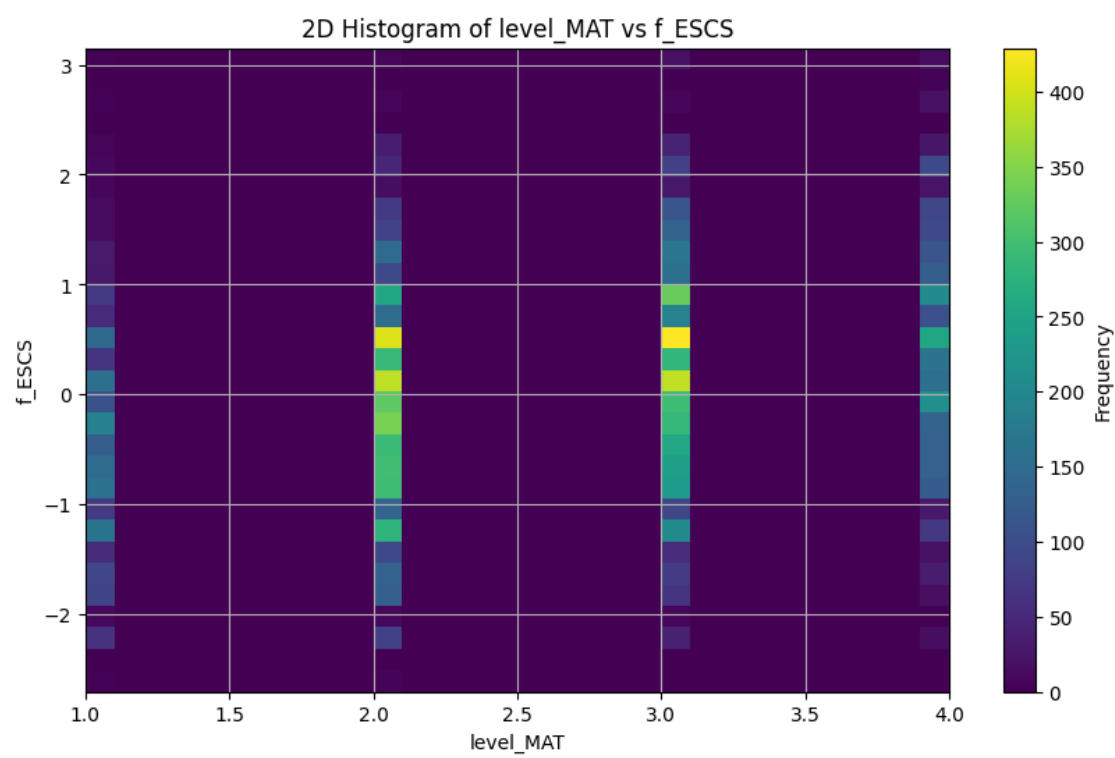
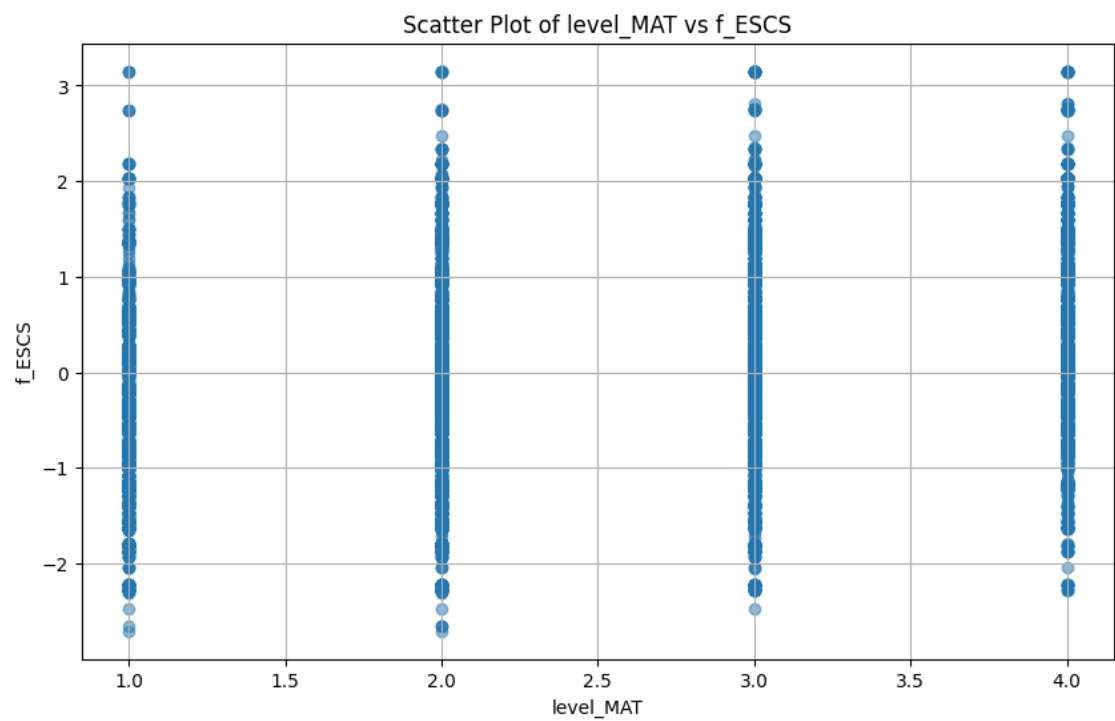


```
[97]: sensitive_feature = "f_ESCS"

trial_df = final_df[[target_feature, sensitive_feature]]
trial_df = trial_df.dropna()
col_a = trial_df[target_feature]
col_b = trial_df[sensitive_feature]

# Scatter Plot
plt.figure(figsize=(10, 6))
plt.scatter(col_a, col_b, alpha=0.5)
plt.title(f"Scatter Plot of {target_feature} vs {sensitive_feature}")
plt.xlabel(target_feature)
plt.ylabel(sensitive_feature)
plt.grid()
plt.show()

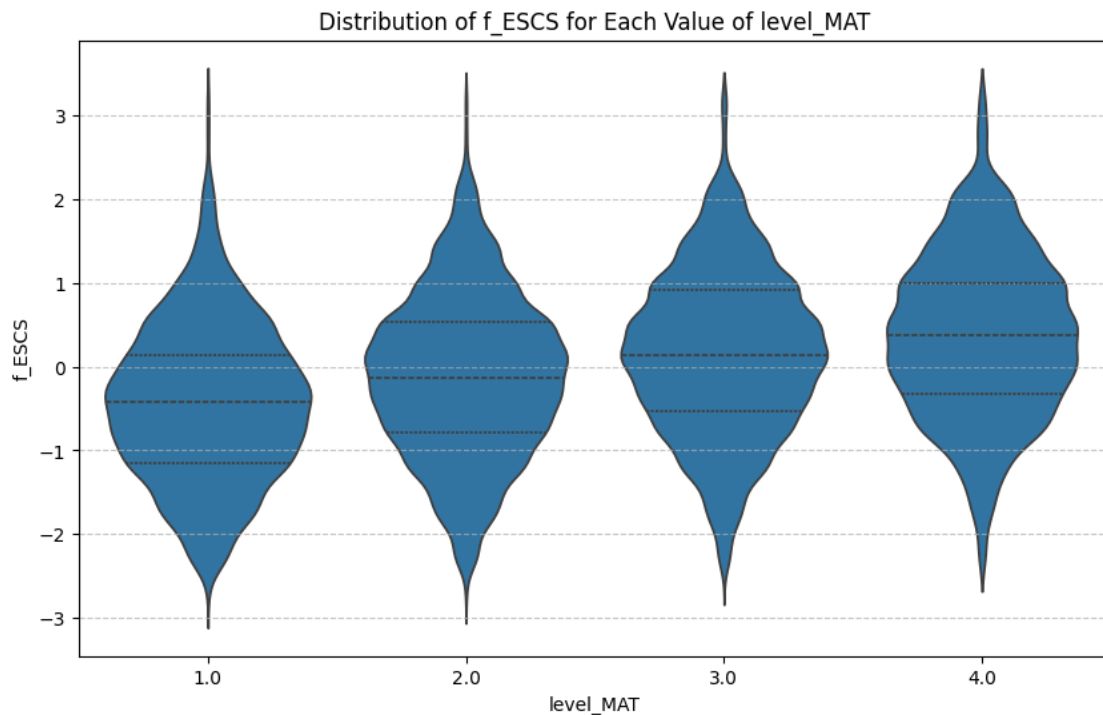
# 2D Histogram (Heatmap)
plt.figure(figsize=(10, 6))
plt.hist2d(col_a, col_b, bins=(30, 30), cmap=plt.cm.viridis)
plt.colorbar(label='Frequency')
plt.title(f"2D Histogram of {target_feature} vs {sensitive_feature}")
plt.xlabel(target_feature)
plt.ylabel(sensitive_feature)
plt.grid()
plt.show()
```

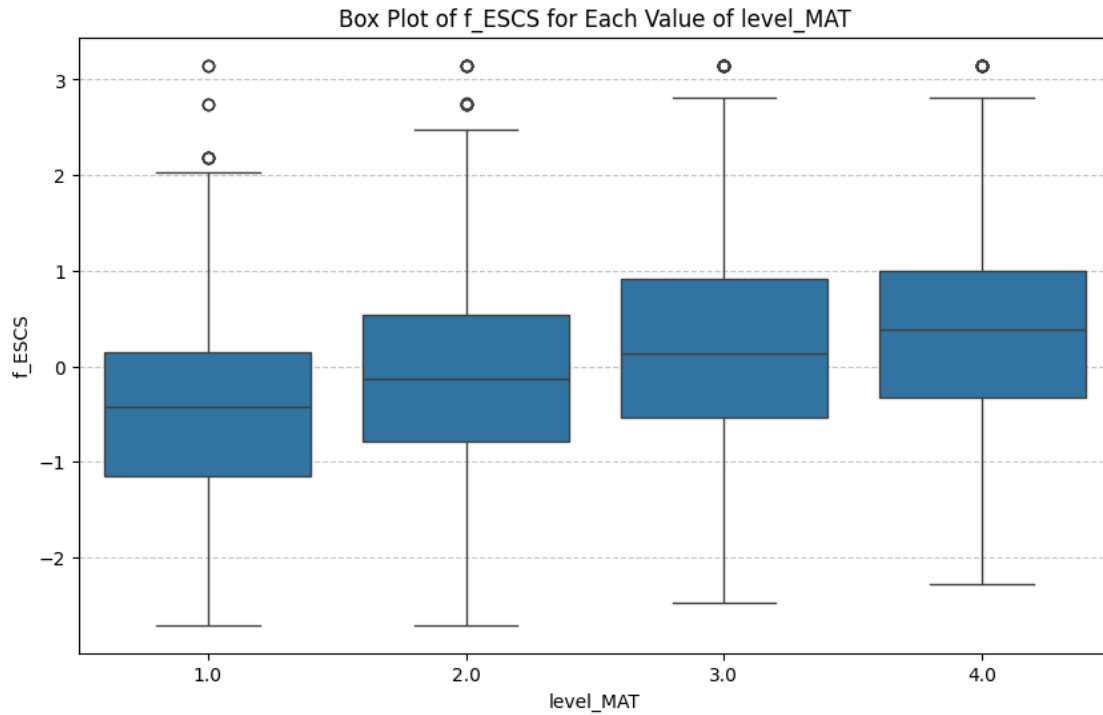


```
[189]: sensitive_feature = "f_ESCS"

plt.figure(figsize=(10, 6))
sns.violinplot(x=target_feature, y=sensitive_feature, data=trial_df,
               inner="quartile")
plt.title(f'Distribution of {sensitive_feature} for Each Value of {target_feature}')
plt.xlabel(target_feature)
plt.ylabel(sensitive_feature)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()

# Box Plot
plt.figure(figsize=(10, 6))
sns.boxplot(x=target_feature, y=sensitive_feature, data=trial_df)
plt.title(f'Box Plot of {sensitive_feature} for Each Value of {target_feature}')
plt.xlabel(target_feature)
plt.ylabel(sensitive_feature)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```





8 Processing

```
[283]: sensitive_feature = "s_gender"
        favorable_sensitive_label = "MALE"
        unfavorable_sensitive_label = "FEMALE"

        class_feature = target_feature
        favorable_class_label = "PASSED"
        unfavorable_class_label = "NOT PASSED"

        metric = "balanced_accuracy"
        fair_metric = "demographic_parity_ratio"
```

8.1 Utils

```
[340]: def encode_feature(df, feature, favorable_label):

        df_encoded = df.copy()

        # Encode class
        if feature in df_encoded.columns:
            df_encoded[feature] = df_encoded[
                feature
```



```

        ].apply(lambda x: 1 if x == favorable_label else 0)
    return df_encoded

```

```

[339]: def decode_feature(df, feature, favorable_label, unfavorable_label):

    df_decoded = df.copy()

    # Encode class
    if feature in df_decoded.columns:
        df_decoded[feature] = df_decoded[
            feature
        ].apply(lambda x: favorable_label if x == 1 else unfavorable_label)
    return df_decoded

```

8.1.1 Estimator

```

[341]: from sklearn.base import BaseEstimator, ClassifierMixin
from sklearn.preprocessing import OneHotEncoder, OrdinalEncoder
from aif360.datasets import BinaryLabelDataset
from aif360.metrics import BinaryLabelDatasetMetric
import pandas as pd

class AIF360SklearnWrapper(BaseEstimator, ClassifierMixin):
    def __init__(
        self,
        aif360_model,
        sensitive_feature,
        favorable_class_label,
        unfavorable_class_label,
        favorable_sens_group,
    ):
        self.aif360_model = aif360_model

        self.sensitive_feature = sensitive_feature
        self.favorable_sens_group = favorable_sens_group

        self.class_feature = "label"
        self.favorable_class_label = favorable_class_label
        self.unfavorable_class_label = unfavorable_class_label

        self.ordinal_encoder = None

    def _encode_features(self, df, fit):

        df_encoded = df.copy()

        # Encode class

```

```

        if self.class_feature in df_encoded.columns:
            df_encoded[self.class_feature] = df_encoded[
                self.class_feature
            ].apply(lambda x: 1 if x == self.favorable_class_label else 0)
        else:
            df_encoded[self.class_feature] = 0 # or any constant value

        # Encode sensitive feature
        if self.sensitive_feature in df_encoded.
↪select_dtypes(include=['object', 'category']).columns:
            df_encoded[self.sensitive_feature] = df_encoded[
                self.sensitive_feature
            ].apply(lambda x: 1 if x == self.favorable_sens_group else 0)

        # # Encode other columns
        # categorical_cols = df_encoded.select_dtypes(include=['object',
↪'category']).columns
        # numeric_cols = df_encoded.select_dtypes(exclude=['object',
↪'category']).columns

        # if fit:
        #     self.ordinal_encoder = OrdinalEncoder()
        #     self.ordinal_encoder.fit(df_encoded[categorical_cols])

        # if len(categorical_cols) > 0:
        #     df_encoded[categorical_cols] = self.ordinal_encoder.
↪transform(df_encoded[categorical_cols])
        # if len(numeric_cols) > 0:
        #     df_encoded[numeric_cols] = df_encoded[numeric_cols]

        # # Check missing values
        # print(df_encoded.isna().sum())
        # df_encoded = df_encoded.set_index(self.sensitive_feature)
        # print(df_encoded.isna().sum())

        return df_encoded

    def fit(self, X, y, **kwargs):
        if not self.sensitive_feature:
            raise ValueError("Protected attribute name must be specified.")

        df_encoded = self._encode_features(df=X.assign(label=y), fit=True)
        # print(df_encoded)
        dataset = BinaryLabelDataset(
            df=df_encoded,
            label_names=[self.class_feature],

```

```

        # sensitive_features=[self.sensitive_feature],
        protected_attribute_names=[self.sensitive_feature],
        favorable_label=1,
        unfavorable_label=0
    )
    self.aif360_model = self.aif360_model.fit(dataset, **kwargs)
    return self

def predict(self, X, decode=False):
    X_encoded = self._encode_features(df=X, fit=False)
    dataset = BinaryLabelDataset(
        df=X_encoded,
        label_names=[self.class_feature],
        # sensitive_features=[self.sensitive_feature],
        protected_attribute_names=[self.sensitive_feature],
        favorable_label=1,
        unfavorable_label=0
    )
    predictions = self.aif360_model.predict(dataset).
    ↪convert_to_dataframe()[0][self.class_feature]
    predictions = predictions.apply(lambda x: self.favorable_class_label if
    ↪x == 1 else self.unfavorable_class_label) if decode else predictions
    return predictions

def predict_proba(self, X):
    if hasattr(self.aif360_model, "predict_proba"):
        X_encoded = self._encode_features(df=X, fit=False)
        dataset = BinaryLabelDataset(
            df=X_encoded,
            label_names=[self.class_feature],
            # sensitive_features=[self.sensitive_feature],
            protected_attribute_names=[self.sensitive_feature],
            favorable_label=1,
            unfavorable_label=0
        )
        probabilities = self.aif360_model.predict_proba(dataset)
        return probabilities
    else:
        raise NotImplementedError("The wrapped model does not support
    ↪probability prediction.")

def score(self, X, y):
    y_pred = self.predict(X)
    return (y_pred == y).mean()

```

8.1.2 Transformers

```
[ ]: import pandas as pd

from sklearn.base import BaseEstimator, TransformerMixin
from sklearn.preprocessing import OrdinalEncoder

from aif360.sklearn.preprocessing import LearnedFairRepresentations

class LFR_wrapper(BaseEstimator, TransformerMixin):
    """
    Wrapper class for LearnedFairRepresentations that allows it to be used with
    ↪ numpy arrays.
    """

    feature_names = []
    enc = None

    def __init__(self, prot_attr, feature_names, **kwargs):
        """
        Initialize the LFR_wrapper.

        Args:
            X: The numpy array of input data.
            y: The numpy array of target labels.
            feature_names: The list of feature names.
            **kwargs: Additional keyword arguments to pass to the
            ↪ LearnedFairRepresentations constructor.
        """
        self.prot_attr = prot_attr
        self.feature_names = feature_names
        self.lfr = LearnedFairRepresentations(
            prot_attr=prot_attr if len(prot_attr) == 1 else "mixin", **kwargs
        )

    def _prepare_dataset(self, X, y=None):
        df = pd.DataFrame(X, columns=self.feature_names)

        if len(self.prot_attr) == 0:
            raise Exception("No sensitive features to mitigate")

        if len(self.prot_attr) > 1:
            df["mixin"] = df[self.prot_attr[0]].astype(str) + df[
                self.prot_attr[1]
            ].astype(str)
            df = df.drop(columns=self.prot_attr)
```

```

        self.enc = OrdinalEncoder()
        df["mixin"] = self.enc.fit_transform(df["mixin"].to_numpy().
↳reshape(-1, 1))
        df = df.set_index("mixin")
    else:
        df = df.set_index(self.prot_attr)

    if y is not None:
        df["target"] = y
        df_X, df_y = df.drop(columns=["target"]), df["target"]
    else:
        df_X, df_y = df, None

    return df_X, df_y

def fit(self, X, y):
    df_X, df_y = self._prepare_dataset(X, y)
    self.lfr.fit(df_X, df_y)
    return self

def transform(self, X):
    df_X, _ = self._prepare_dataset(X, None)
    transformed_df = self.lfr.transform(df_X)
    return transformed_df.values

def fit_transform(self, X, y=None):
    df_X, df_y = self._prepare_dataset(X, y)
    transformed_df = self.lfr.fit_transform(df_X, df_y)
    return transformed_df.values

```

8.1.3 Metrics

Platform

```

[ ]: detection_results = {
    "disparate_impact": [],
    "statistical_parity_difference": []
}

for sens_col, priv_group in sensitive_groups:
    # print(f"{sens_col}: {priv_group}")

    current_df = copy.deepcopy(df)
    current_df[sens_col] = current_df[sens_col].apply(lambda x: 1 if x ==
↳priv_group else 0)

```

```

for out_col, out_group in outcome_groups:
    # print(f"{out_col}: {out_group}")

    current_df[out_col] = current_df[out_col].apply(lambda x: 1 if x ==
↪out_group else 0)

    current_dataset = BinaryLabelDataset(
        favorable_label=1,
        unfavorable_label=0,
        df=current_df[[sens_col, out_col]],
        label_names=[out_col],
        sensitive_features=[sens_col]
    )

    # Step 2: Compute metrics
    metric = BinaryLabelDatasetMetric(
        current_dataset,
        privileged_groups=[{sens_col: 1}],
        unprivileged_groups=[{sens_col: 0}],
    )

    # Disparate Impact
    disparate_impact = metric.disparate_impact()
    # print(f"Disparate Impact: {disparate_impact}")
    detection_results["disparate_impact"] += [
        {
            "when": {
                sens_col: priv_group,
                out_col: out_group
            },
            "value": "NaN" if np.isnan(disparate_impact) else
↪disparate_impact
        }
    ]

    # Statistical Parity Difference
    statistical_parity_difference = metric.statistical_parity_difference()
    # print(f"Statistical Parity Difference:
↪{statistical_parity_difference}")
    detection_results["statistical_parity_difference"] += [
        {
            "when": {
                sens_col: priv_group,
                out_col: out_group
            },
            "value": "NaN" if np.isnan(statistical_parity_difference) else
↪statistical_parity_difference

```

```

    }
]

# print()
detection_results

```

```

[342]: def compute_fair_metric(
        fair_metric_name, sensitive_feature, X, y_true, y_pred
    ):

        # metrics_module = __import__("metrics")
        metrics_module = globals()["metrics"]
        fair_metric = getattr(metrics_module, fair_metric_name)

        X_sensitive = encode_feature(X, sensitive_feature,
        ↪favorable_label=favorable_sensitive_label)
        X_sensitive = X_sensitive[sensitive_feature]

        return fair_metric(
            y_true=y_true,
            y_pred=y_pred,
            sensitive_features=X_sensitive,
        )

```

8.2 Analysis

```

[357]: final_df

```

```

[357]:          s_frequency_of_work_with_teachers  \
id_questionnaire
1                0.733333
3                0.833333
5                0.566667
6                0.533333
8                0.633333
...
20421            0.566667
20422            0.700000
20423            0.400000
20424            0.500000
20427            0.666667

          s_frequency_of_evaluations  s_frequency_of_internet_usage  \
id_questionnaire
1                0.740741                0.066667
3                0.750000                0.666667
5                0.888889                0.133333

```

6	0.851852	0.133333
8	0.703704	0.266667
...
20421	0.777778	0.600000
20422	0.518519	0.266667
20423	0.555556	0.400000
20424	0.555556	0.200000
20427	0.740741	0.133333

	s_frequency_of_materials_in_class	f_ESCS \
id_questionnaire		
1	0.380952	0.235340
3	1.000000	0.261451
5	0.666667	3.151773
6	0.380952	-1.627731
8	0.666667	-0.358498
...
20421	0.476190	-0.764891
20422	0.380952	-0.048524
20423	0.571429	0.072922
20424	0.523810	2.755881
20427	0.476190	1.372627

	s_extent_of_classmates_affinity \
id_questionnaire	
1	0.714286
3	0.833333
5	0.952381
6	0.761905
8	0.857143
...	...
20421	0.904762
20422	0.952381
20423	0.809524
20424	0.904762
20427	0.857143

	s_extent_of_teacher_performance	s_extent_of_class_env \
id_questionnaire		
1	0.833333	0.791667
3	0.800000	0.833333
5	1.000000	1.000000
6	0.866667	1.000000
8	0.966667	0.833333
...
20421	1.000000	1.000000
20422	0.888889	0.916667

20423	1.000000	0.666667
20424	0.766667	0.791667
20427	0.900000	0.916667

	f_mother_age	s_frequency_of_computer_usage \
id_questionnaire		
1	43.0	0.111111
3	45.0	0.666667
5	39.0	0.000000
6	25.0	0.222222
8	37.0	0.333333
...
20421	41.0	0.111111
20422	36.0	0.222222
20423	46.0	0.111111
20424	41.0	0.111111
20427	37.0	0.000000

	f_number_of_homework_hours_a_week \
id_questionnaire	
1	14.0
3	9.0
5	6.0
6	5.0
8	10.0
...	...
20421	14.0
20422	2.0
20423	8.0
20424	2.0
20427	1.0

	s_extent_of_school_satisfaction	f_mother_education_level \
id_questionnaire		
1	0.944444	7.0
3	0.888889	3.0
5	1.000000	9.0
6	1.000000	3.0
8	0.833333	7.0
...
20421	1.000000	5.0
20422	0.944444	4.0
20423	1.000000	3.0
20424	0.833333	9.0
20427	1.000000	5.0

	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
--	---------------------------------	----------------------------

id_questionnaire		
1	0.700000	3.0
3	0.291667	4.0
5	0.833333	7.0
6	0.833333	3.0
8	1.000000	3.0
...
20421	1.000000	4.0
20422	0.708333	5.0
20423	0.833333	10.0
20424	0.666667	10.0
20427	0.291667	5.0

f_frequency_of_books_at_home \	
id_questionnaire	
1	1.000000
3	0.444444
5	0.555556
6	0.333333
8	0.555556
...	...
20421	0.222222
20422	0.555556
20423	0.444444
20424	0.555556
20427	0.444444

f_frequency_of_parent_involved_in_school_activities \	
id_questionnaire	
1	0.333333
3	0.166667
5	0.250000
6	0.083333
8	0.333333
...	...
20421	0.333333
20422	0.000000
20423	0.000000
20424	0.416667
20427	0.000000

id_questionnaire	s_gender	level_MAT
1	FEMALE	3.0
3	FEMALE	1.0
5	FEMALE	4.0
6	FEMALE	1.0

8	MALE	2.0
...
20421	MALE	2.0
20422	FEMALE	3.0
20423	MALE	2.0
20424	MALE	3.0
20427	FEMALE	2.0

[10735 rows x 19 columns]

```
[199]: compute_exploratory_analysis(df=final_df, column_pattern=None,
    ↪missing_threshold=0.3, annot=False)
```

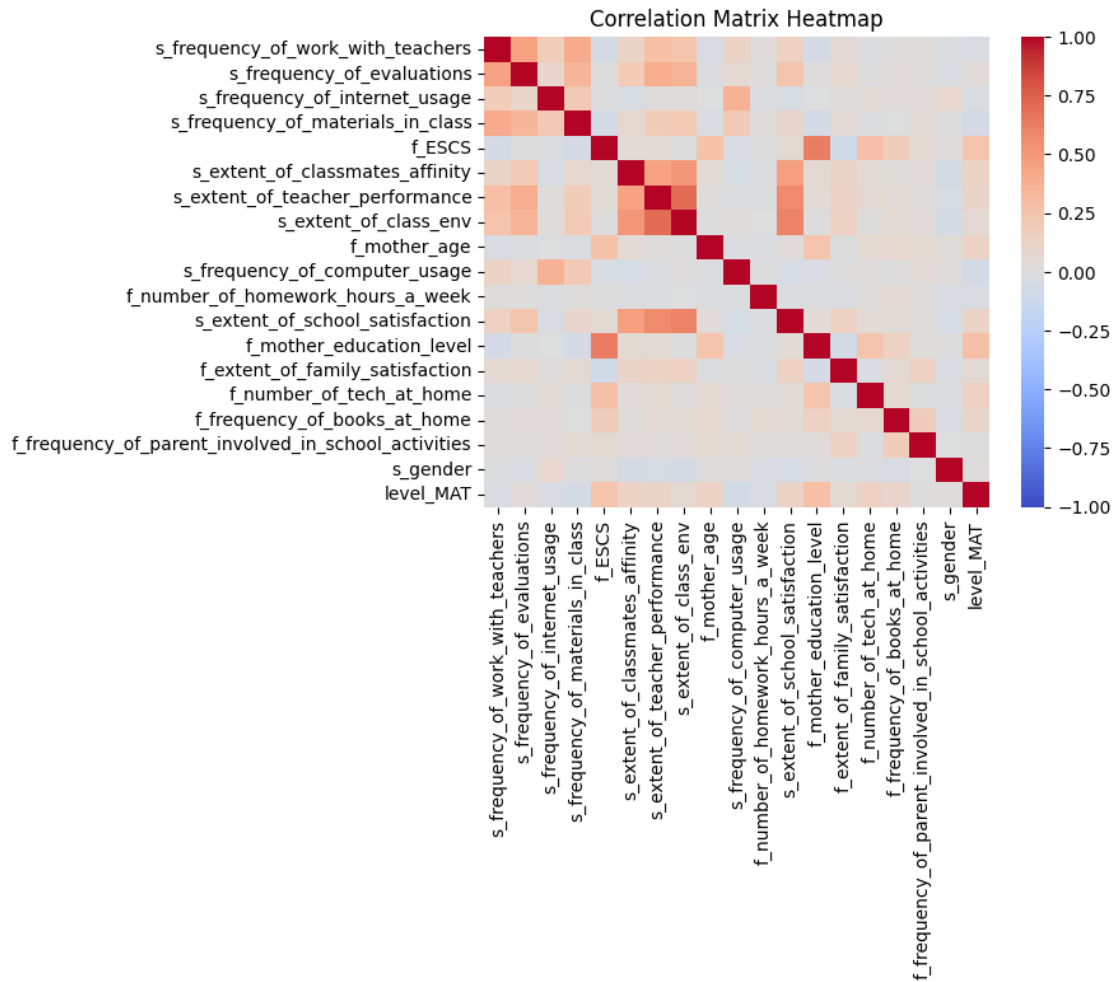
GOOD FEATURES: 19

s_frequency_of_work_with_teachers	0.056329
s_frequency_of_evaluations	0.058632
s_frequency_of_internet_usage	0.055824
s_frequency_of_materials_in_class	0.059811
f_ESCS	0.275357
s_extent_of_classmates_affinity	0.060036
s_extent_of_teacher_performance	0.059025
s_extent_of_class_env	0.059250
f_mother_age	0.281534
s_frequency_of_computer_usage	0.056049
f_number_of_homework_hours_a_week	0.278445
s_extent_of_school_satisfaction	0.057396
f_mother_education_level	0.291082
f_extent_of_family_satisfaction	0.277547
f_number_of_tech_at_home	0.281984
f_frequency_of_books_at_home	0.290295
f_frequency_of_parent_involved_in_school_activities	0.278502
s_gender	0.000056
level_MAT	0.000000

dtype: float64

BAD FEATURES

Series([], dtype: float64)



id_questionnaire	s_frequency_of_work_with_teachers \
1	0.733333
2	0.533333
3	0.833333
4	0.400000
5	0.566667
...	...
20426	0.333333
20427	0.666667
20428	0.518519
20429	0.633333
20430	0.629630

id_questionnaire	s_frequency_of_evaluations	s_frequency_of_internet_usage \
1	0.740741	0.066667

2	0.888889	0.600000
3	0.750000	0.666667
4	0.541667	0.266667
5	0.888889	0.133333
...
20426	0.333333	0.333333
20427	0.740741	0.133333
20428	0.740741	0.933333
20429	0.740741	0.533333
20430	0.481481	0.400000

	s_frequency_of_materials_in_class	f_ESCS \
id_questionnaire		
1	0.380952	0.235340
2	0.714286	NaN
3	1.000000	0.261451
4	NaN	0.787122
5	0.666667	3.151773
...
20426	NaN	-1.627731
20427	0.476190	1.372627
20428	0.777778	NaN
20429	0.142857	NaN
20430	0.476190	NaN

	s_extent_of_classmates_affinity \
id_questionnaire	
1	0.714286
2	0.714286
3	0.833333
4	0.857143
5	0.952381
...	...
20426	0.333333
20427	0.857143
20428	0.666667
20429	0.714286
20430	0.619048

	s_extent_of_teacher_performance	s_extent_of_class_env \
id_questionnaire		
1	0.833333	0.791667
2	0.800000	0.791667
3	0.800000	0.833333
4	0.800000	0.750000
5	1.000000	1.000000
...
20426	0.333333	0.333333

20427	0.900000	0.916667
20428	0.566667	0.958333
20429	1.000000	0.875000
20430	0.733333	0.708333

	f_mother_age	s_frequency_of_computer_usage \
id_questionnaire		
1	43.0	0.111111
2	NaN	0.333333
3	45.0	0.666667
4	39.0	0.333333
5	39.0	0.000000
...
20426	31.0	NaN
20427	37.0	0.000000
20428	NaN	0.222222
20429	NaN	0.000000
20430	NaN	0.111111

	f_number_of_homework_hours_a_week \
id_questionnaire	
1	14.0
2	NaN
3	9.0
4	9.0
5	6.0
...	...
20426	4.0
20427	1.0
20428	NaN
20429	NaN
20430	NaN

	s_extent_of_school_satisfaction	f_mother_education_level \
id_questionnaire		
1	0.944444	7.0
2	1.000000	NaN
3	0.888889	3.0
4	0.722222	5.0
5	1.000000	9.0
...
20426	1.000000	3.0
20427	1.000000	5.0
20428	0.833333	NaN
20429	1.000000	NaN
20430	0.722222	NaN

	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
--	---------------------------------	----------------------------

id_questionnaire		
1	0.700000	3.0
2	NaN	NaN
3	0.291667	4.0
4	1.000000	3.0
5	0.833333	7.0
...
20426	0.791667	4.0
20427	0.291667	5.0
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

id_questionnaire	f_frequency_of_books_at_home \
1	1.000000
2	NaN
3	0.444444
4	0.333333
5	0.555556
...	...
20426	0.333333
20427	0.444444
20428	NaN
20429	NaN
20430	NaN

id_questionnaire	f_frequency_of_parent_involved_in_school_activities \
1	0.333333
2	NaN
3	0.166667
4	0.250000
5	0.250000
...	...
20426	0.000000
20427	0.000000
20428	NaN
20429	NaN
20430	NaN

id_questionnaire	s_gender	level_MAT
1	FEMALE	3.0
2	MALE	1.0
3	FEMALE	1.0
4	MALE	2.0
5	FEMALE	4.0

...
20426	FEMALE	1.0
20427	FEMALE	2.0
20428	MALE	2.0
20429	MALE	3.0
20430	FEMALE	2.0

[17806 rows x 19 columns]

```

↪s_frequency_of_work_with_teachers \
s_frequency_of_work_with_teachers 1.
↪000000
s_frequency_of_evaluations 0.
↪459210
s_frequency_of_internet_usage 0.
↪184811
s_frequency_of_materials_in_class 0.
↪407241
f_ESCS -0.
↪056927
s_extent_of_classmates_affinity 0.
↪115582
s_extent_of_teacher_performance 0.
↪284243
s_extent_of_class_env 0.
↪253826
f_mother_age -0.
↪033653
s_frequency_of_computer_usage 0.
↪120296
f_number_of_homework_hours_a_week 0.
↪023143
s_extent_of_school_satisfaction 0.
↪140080
f_mother_education_level -0.
↪066000
f_extent_of_family_satisfaction 0.
↪057850
f_number_of_tech_at_home -0.
↪016357
f_frequency_of_books_at_home 0.
↪017840
f_frequency_of_parent_involved_in_school_activi... 0.
↪016874
s_gender -0.
↪013262

```


level_MAT	-0.
↪038320	

	s_frequency_of_evaluations \
s_frequency_of_work_with_teachers	0.459210
s_frequency_of_evaluations	1.000000
s_frequency_of_internet_usage	0.097839
s_frequency_of_materials_in_class	0.346366
f_ESCS	0.000752
s_extent_of_classmates_affinity	0.208706
s_extent_of_teacher_performance	0.405473
s_extent_of_class_env	0.353524
f_mother_age	-0.010146
s_frequency_of_computer_usage	0.064127
f_number_of_homework_hours_a_week	0.013561
s_extent_of_school_satisfaction	0.246491
f_mother_education_level	0.002241
f_extent_of_family_satisfaction	0.076432
f_number_of_tech_at_home	0.004816
f_frequency_of_books_at_home	0.024959
f_frequency_of_parent_involved_in_school_activi...	0.012819
s_gender	-0.028531
level_MAT	0.033004

↪s_frequency_of_internet_usage \	
s_frequency_of_work_with_teachers	0.184811
s_frequency_of_evaluations	0.097839
s_frequency_of_internet_usage	1.000000
s_frequency_of_materials_in_class	0.215844
f_ESCS	-0.015618
s_extent_of_classmates_affinity	-0.043809
s_extent_of_teacher_performance	0.020481
s_extent_of_class_env	0.022687
f_mother_age	-0.006085
s_frequency_of_computer_usage	0.371735
f_number_of_homework_hours_a_week	-0.020469
s_extent_of_school_satisfaction	-0.039423
f_mother_education_level	-0.004143
f_extent_of_family_satisfaction	0.008416
f_number_of_tech_at_home	0.040167
f_frequency_of_books_at_home	0.023850
f_frequency_of_parent_involved_in_school_activi...	0.025514
s_gender	0.084017
level_MAT	-0.036665

↪s_frequency_of_materials_in_class \

s_frequency_of_work_with_teachers	0.
↪407241	
s_frequency_of_evaluations	0.
↪346366	
s_frequency_of_internet_usage	0.
↪215844	
s_frequency_of_materials_in_class	1.
↪000000	
f_ESCS	-0.
↪068230	
s_extent_of_classmates_affinity	0.
↪075477	
s_extent_of_teacher_performance	0.
↪192643	
s_extent_of_class_env	0.
↪198315	
f_mother_age	-0.
↪030786	
s_frequency_of_computer_usage	0.
↪215274	
f_number_of_homework_hours_a_week	0.
↪013483	
s_extent_of_school_satisfaction	0.
↪088556	
f_mother_education_level	-0.
↪068052	
f_extent_of_family_satisfaction	0.
↪050423	
f_number_of_tech_at_home	-0.
↪014722	
f_frequency_of_books_at_home	-0.
↪004838	
f_frequency_of_parent_involved_in_school_activi...	0.
↪039642	
s_gender	0.
↪000793	
level_MAT	-0.
↪066594	

	f_ESCS \
s_frequency_of_work_with_teachers	-0.056927
s_frequency_of_evaluations	0.000752
s_frequency_of_internet_usage	-0.015618
s_frequency_of_materials_in_class	-0.068230

f_ESCS	1.000000
s_extent_of_classmates_affinity	0.040619
s_extent_of_teacher_performance	0.039800
s_extent_of_class_env	0.005268
f_mother_age	0.261448
s_frequency_of_computer_usage	-0.044139
f_number_of_homework_hours_a_week	0.007309
s_extent_of_school_satisfaction	0.049781
f_mother_education_level	0.627503
f_extent_of_family_satisfaction	-0.087052
f_number_of_tech_at_home	0.278867
f_frequency_of_books_at_home	0.175323
f_frequency_of_parent_involved_in_school_activi...	0.047424
s_gender	0.009554
level_MAT	0.250165

□

↪s_extent_of_classmates_affinity \	
s_frequency_of_work_with_teachers	0.
↪115582	
s_frequency_of_evaluations	0.
↪208706	
s_frequency_of_internet_usage	-0.
↪043809	
s_frequency_of_materials_in_class	0.
↪075477	
f_ESCS	0.
↪040619	
s_extent_of_classmates_affinity	1.
↪000000	
s_extent_of_teacher_performance	0.
↪455324	
s_extent_of_class_env	0.
↪512753	
f_mother_age	0.
↪040068	
s_frequency_of_computer_usage	-0.
↪052212	
f_number_of_homework_hours_a_week	-0.
↪008666	
s_extent_of_school_satisfaction	0.
↪471544	
f_mother_education_level	0.
↪048646	
f_extent_of_family_satisfaction	0.
↪115887	

f_number_of_tech_at_home	0.
↪037754	
f_frequency_of_books_at_home	0.
↪021444	
f_frequency_of_parent_involved_in_school_activi...	0.030362
s_gender	-0.
↪055421	
level_MAT	0.
↪132252	
↪s_extent_of_teacher_performance \	
s_frequency_of_work_with_teachers	0.
↪284243	
s_frequency_of_evaluations	0.
↪405473	
s_frequency_of_internet_usage	0.
↪020481	
s_frequency_of_materials_in_class	0.
↪192643	
f_ESCS	0.
↪039800	
s_extent_of_classmates_affinity	0.
↪455324	
s_extent_of_teacher_performance	1.
↪000000	
s_extent_of_class_env	0.
↪699774	
f_mother_age	0.
↪011708	
s_frequency_of_computer_usage	-0.
↪011863	
f_number_of_homework_hours_a_week	-0.
↪008315	
s_extent_of_school_satisfaction	0.
↪572676	
f_mother_education_level	0.
↪039756	
f_extent_of_family_satisfaction	0.
↪132071	
f_number_of_tech_at_home	0.
↪047904	
f_frequency_of_books_at_home	0.
↪033481	
f_frequency_of_parent_involved_in_school_activi...	0.015670

s_gender	-0.
↪044767	
level_MAT	0.
↪110412	

	s_extent_of_class_env \
s_frequency_of_work_with_teachers	0.253826
s_frequency_of_evaluations	0.353524
s_frequency_of_internet_usage	0.022687
s_frequency_of_materials_in_class	0.198315
f_ESCS	0.005268
s_extent_of_classmates_affinity	0.512753
s_extent_of_teacher_performance	0.699774
s_extent_of_class_env	1.000000
f_mother_age	0.000189
s_frequency_of_computer_usage	0.014833
f_number_of_homework_hours_a_week	-0.004051
s_extent_of_school_satisfaction	0.603268
f_mother_education_level	0.000608
f_extent_of_family_satisfaction	0.135805
f_number_of_tech_at_home	0.014492
f_frequency_of_books_at_home	0.039548
f_frequency_of_parent_involved_in_school_activi...	0.026473
s_gender	-0.055349
level_MAT	0.059059

	f_mother_age \
s_frequency_of_work_with_teachers	-0.033653
s_frequency_of_evaluations	-0.010146
s_frequency_of_internet_usage	-0.006085
s_frequency_of_materials_in_class	-0.030786
f_ESCS	0.261448
s_extent_of_classmates_affinity	0.040068
s_extent_of_teacher_performance	0.011708
s_extent_of_class_env	0.000189
f_mother_age	1.000000
s_frequency_of_computer_usage	-0.018088
f_number_of_homework_hours_a_week	-0.019828
s_extent_of_school_satisfaction	0.024990
f_mother_education_level	0.250886
f_extent_of_family_satisfaction	-0.013724
f_number_of_tech_at_home	0.058864
f_frequency_of_books_at_home	0.072558
f_frequency_of_parent_involved_in_school_activi...	0.060719
s_gender	0.021320
level_MAT	0.126018

↪s_frequency_of_computer_usage \

s_frequency_of_work_with_teachers	0.120296
s_frequency_of_evaluations	0.064127
s_frequency_of_internet_usage	0.371735
s_frequency_of_materials_in_class	0.215274
f_ESCS	-0.044139
s_extent_of_classmates_affinity	-0.052212
s_extent_of_teacher_performance	-0.011863
s_extent_of_class_env	0.014833
f_mother_age	-0.018088
s_frequency_of_computer_usage	1.000000
f_number_of_homework_hours_a_week	-0.011977
s_extent_of_school_satisfaction	-0.051609
f_mother_education_level	-0.035688
f_extent_of_family_satisfaction	0.007575
f_number_of_tech_at_home	0.012460
f_frequency_of_books_at_home	0.004609
f_frequency_of_parent_involved_in_school_activi...	0.040685
s_gender	0.019260
level_MAT	-0.072532

↪f_number_of_homework_hours_a_week \

s_frequency_of_work_with_teachers	0.
↪023143	
s_frequency_of_evaluations	0.
↪013561	
s_frequency_of_internet_usage	-0.
↪020469	
s_frequency_of_materials_in_class	0.
↪013483	
f_ESCS	0.
↪007309	
s_extent_of_classmates_affinity	-0.
↪008666	
s_extent_of_teacher_performance	-0.
↪008315	
s_extent_of_class_env	-0.
↪004051	
f_mother_age	-0.
↪019828	
s_frequency_of_computer_usage	-0.
↪011977	
f_number_of_homework_hours_a_week	1.
↪000000	

s_extent_of_school_satisfaction	-0.
↪009182	
f_mother_education_level	-0.
↪008756	
f_extent_of_family_satisfaction	-0.
↪019991	
f_number_of_tech_at_home	0.
↪020601	
f_frequency_of_books_at_home	0.
↪057714	
f_frequency_of_parent_involved_in_school_activi...	0.
↪015347	
s_gender	-0.
↪033709	
level_MAT	-0.
↪035409	
□	
↪s_extent_of_school_satisfaction \	
s_frequency_of_work_with_teachers	0.
↪140080	
s_frequency_of_evaluations	0.
↪246491	
s_frequency_of_internet_usage	-0.
↪039423	
s_frequency_of_materials_in_class	0.
↪088556	
f_ESCS	0.
↪049781	
s_extent_of_classmates_affinity	0.
↪471544	
s_extent_of_teacher_performance	0.
↪572676	
s_extent_of_class_env	0.
↪603268	
f_mother_age	0.
↪024990	
s_frequency_of_computer_usage	-0.
↪051609	
f_number_of_homework_hours_a_week	-0.
↪009182	
s_extent_of_school_satisfaction	1.
↪000000	
f_mother_education_level	0.
↪042832	

f_extent_of_family_satisfaction	0.
↪136325	
f_number_of_tech_at_home	0.
↪039812	
f_frequency_of_books_at_home	0.
↪036158	
f_frequency_of_parent_involved_in_school_activi...	0.016555
s_gender	-0.
↪054659	
level_MAT	0.
↪123259	

	f_mother_education_level \
s_frequency_of_work_with_teachers	-0.066000
s_frequency_of_evaluations	0.002241
s_frequency_of_internet_usage	-0.004143
s_frequency_of_materials_in_class	-0.068052
f_ESCS	0.627503
s_extent_of_classmates_affinity	0.048646
s_extent_of_teacher_performance	0.039756
s_extent_of_class_env	0.000608
f_mother_age	0.250886
s_frequency_of_computer_usage	-0.035688
f_number_of_homework_hours_a_week	-0.008756
s_extent_of_school_satisfaction	0.042832
f_mother_education_level	1.000000
f_extent_of_family_satisfaction	-0.065329
f_number_of_tech_at_home	0.254568
f_frequency_of_books_at_home	0.146603
f_frequency_of_parent_involved_in_school_activi...	0.020266
s_gender	0.006079
level_MAT	0.284138

↪f_extent_of_family_satisfaction \

s_frequency_of_work_with_teachers	0.
↪057850	
s_frequency_of_evaluations	0.
↪076432	
s_frequency_of_internet_usage	0.
↪008416	
s_frequency_of_materials_in_class	0.
↪050423	
f_ESCS	-0.
↪087052	
s_extent_of_classmates_affinity	0.
↪115887	

s_extent_of_teacher_performance	0.
↪132071	
s_extent_of_class_env	0.
↪135805	
f_mother_age	-0.
↪013724	
s_frequency_of_computer_usage	0.
↪007575	
f_number_of_homework_hours_a_week	-0.
↪019991	
s_extent_of_school_satisfaction	0.
↪136325	
f_mother_education_level	-0.
↪065329	
f_extent_of_family_satisfaction	1.
↪000000	
f_number_of_tech_at_home	-0.
↪030832	
f_frequency_of_books_at_home	0.
↪057873	
f_frequency_of_parent_involved_in_school_activi...	0.136940
s_gender	-0.
↪024137	
level_MAT	0.
↪050280	

	f_number_of_tech_at_home \
s_frequency_of_work_with_teachers	-0.016357
s_frequency_of_evaluations	0.004816
s_frequency_of_internet_usage	0.040167
s_frequency_of_materials_in_class	-0.014722
f_ESCS	0.278867
s_extent_of_classmates_affinity	0.037754
s_extent_of_teacher_performance	0.047904
s_extent_of_class_env	0.014492
f_mother_age	0.058864
s_frequency_of_computer_usage	0.012460
f_number_of_homework_hours_a_week	0.020601
s_extent_of_school_satisfaction	0.039812
f_mother_education_level	0.254568
f_extent_of_family_satisfaction	-0.030832
f_number_of_tech_at_home	1.000000
f_frequency_of_books_at_home	0.063726
f_frequency_of_parent_involved_in_school_activi...	-0.011387
s_gender	-0.009508
level_MAT	0.143473

	f_frequency_of_books_at_home	
↪ \		
s_frequency_of_work_with_teachers		0.017840
s_frequency_of_evaluations		0.024959
s_frequency_of_internet_usage		0.023850
s_frequency_of_materials_in_class		-0.004838
f_ESCS		0.175323
s_extent_of_classmates_affinity		0.021444
s_extent_of_teacher_performance		0.033481
s_extent_of_class_env		0.039548
f_mother_age		0.072558
s_frequency_of_computer_usage		0.004609
f_number_of_homework_hours_a_week		0.057714
s_extent_of_school_satisfaction		0.036158
f_mother_education_level		0.146603
f_extent_of_family_satisfaction		0.057873
f_number_of_tech_at_home		0.063726
f_frequency_of_books_at_home		1.000000
f_frequency_of_parent_involved_in_school_activi...		0.179154
s_gender		-0.037040
level_MAT		0.097342

	f_frequency_of_parent_involved_in_school_activities	\
s_frequency_of_work_with_teachers		
↪	0.016874	
s_frequency_of_evaluations		
↪	0.012819	
s_frequency_of_internet_usage		
↪	0.025514	
s_frequency_of_materials_in_class		
↪	0.039642	
f_ESCS		
↪	0.047424	
s_extent_of_classmates_affinity		
↪	0.030362	
s_extent_of_teacher_performance		
↪	0.015670	
s_extent_of_class_env		
↪	0.026473	
f_mother_age		
↪	0.060719	
s_frequency_of_computer_usage		
↪	0.040685	
f_number_of_homework_hours_a_week		
↪	0.015347	

s_extent_of_school_satisfaction	0.016555	␣
f_mother_education_level	0.020266	␣
f_extent_of_family_satisfaction	0.136940	␣
f_number_of_tech_at_home	-0.011387	␣
f_frequency_of_books_at_home	0.179154	␣
f_frequency_of_parent_involved_in_school_activi...	1.000000	␣
s_gender	-0.007350	␣
level_MAT	0.002740	␣

	s_gender	level_MAT
s_frequency_of_work_with_teachers	-0.013262	-0.038320
s_frequency_of_evaluations	-0.028531	0.033004
s_frequency_of_internet_usage	0.084017	-0.036665
s_frequency_of_materials_in_class	0.000793	-0.066594
f_ESCS	0.009554	0.250165
s_extent_of_classmates_affinity	-0.055421	0.132252
s_extent_of_teacher_performance	-0.044767	0.110412
s_extent_of_class_env	-0.055349	0.059059
f_mother_age	0.021320	0.126018
s_frequency_of_computer_usage	0.019260	-0.072532
f_number_of_homework_hours_a_week	-0.033709	-0.035409
s_extent_of_school_satisfaction	-0.054659	0.123259
f_mother_education_level	0.006079	0.284138
f_extent_of_family_satisfaction	-0.024137	0.050280
f_number_of_tech_at_home	-0.009508	0.143473
f_frequency_of_books_at_home	-0.037040	0.097342
f_frequency_of_parent_involved_in_school_activi...	-0.007350	0.002740
s_gender	1.000000	0.015317
level_MAT	0.015317	1.000000

```
[199]: (
    id_questionnaire
    1          0.733333
    2          0.533333
    3          0.833333
    4          0.400000
    5          0.566667
    ...
    20426      0.333333
```

20427	0.666667
20428	0.518519
20429	0.633333
20430	0.629630

	s_frequency_of_evaluations	s_frequency_of_internet_usage \
id_questionnaire		
1	0.740741	0.066667
2	0.888889	0.600000
3	0.750000	0.666667
4	0.541667	0.266667
5	0.888889	0.133333
...
20426	0.333333	0.333333
20427	0.740741	0.133333
20428	0.740741	0.933333
20429	0.740741	0.533333
20430	0.481481	0.400000

	s_frequency_of_materials_in_class	f_ESCS \
id_questionnaire		
1	0.380952	0.235340
2	0.714286	NaN
3	1.000000	0.261451
4	NaN	0.787122
5	0.666667	3.151773
...
20426	NaN	-1.627731
20427	0.476190	1.372627
20428	0.777778	NaN
20429	0.142857	NaN
20430	0.476190	NaN

	s_extent_of_classmates_affinity \
id_questionnaire	
1	0.714286
2	0.714286
3	0.833333
4	0.857143
5	0.952381
...	...
20426	0.333333
20427	0.857143
20428	0.666667
20429	0.714286
20430	0.619048

id_questionnaire	s_extent_of_teacher_performance	s_extent_of_class_env \
1	0.833333	0.791667
2	0.800000	0.791667
3	0.800000	0.833333
4	0.800000	0.750000
5	1.000000	1.000000
...
20426	0.333333	0.333333
20427	0.900000	0.916667
20428	0.566667	0.958333
20429	1.000000	0.875000
20430	0.733333	0.708333

id_questionnaire	f_mother_age	s_frequency_of_computer_usage \
1	43.0	0.111111
2	NaN	0.333333
3	45.0	0.666667
4	39.0	0.333333
5	39.0	0.000000
...
20426	31.0	NaN
20427	37.0	0.000000
20428	NaN	0.222222
20429	NaN	0.000000
20430	NaN	0.111111

id_questionnaire	f_number_of_homework_hours_a_week \
1	14.0
2	NaN
3	9.0
4	9.0
5	6.0
...	...
20426	4.0
20427	1.0
20428	NaN
20429	NaN
20430	NaN

id_questionnaire	s_extent_of_school_satisfaction	f_mother_education_level \
1	0.944444	7.0
2	1.000000	NaN
3	0.888889	3.0

4	0.722222	5.0
5	1.000000	9.0
...
20426	1.000000	3.0
20427	1.000000	5.0
20428	0.833333	NaN
20429	1.000000	NaN
20430	0.722222	NaN

id_questionnaire	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
1	0.700000	3.0
2	NaN	NaN
3	0.291667	4.0
4	1.000000	3.0
5	0.833333	7.0
...
20426	0.791667	4.0
20427	0.291667	5.0
20428	NaN	NaN
20429	NaN	NaN
20430	NaN	NaN

id_questionnaire	f_frequency_of_books_at_home \
1	1.000000
2	NaN
3	0.444444
4	0.333333
5	0.555556
...	...
20426	0.333333
20427	0.444444
20428	NaN
20429	NaN
20430	NaN

id_questionnaire	f_frequency_of_parent_involved_in_school_activities \
1	0.333333
2	NaN
3	0.166667
4	0.250000
5	0.250000
...	...
20426	0.000000
20427	0.000000

20428	NaN
20429	NaN
20430	NaN

	s_gender	level_MAT
id_questionnaire		
1	FEMALE	3.0
2	MALE	1.0
3	FEMALE	1.0
4	MALE	2.0
5	FEMALE	4.0
...
20426	FEMALE	1.0
20427	FEMALE	2.0
20428	MALE	2.0
20429	MALE	3.0
20430	FEMALE	2.0

```
[17806 rows x 19 columns],
s_frequency_of_work_with_teachers \
s_frequency_of_work_with_teachers
1.000000
s_frequency_of_evaluations
0.459210
s_frequency_of_internet_usage
0.184811
s_frequency_of_materials_in_class
0.407241
f_ESCS
-0.056927
s_extent_of_classmates_affinity
0.115582
s_extent_of_teacher_performance
0.284243
s_extent_of_class_env
0.253826
f_mother_age
-0.033653
s_frequency_of_computer_usage
0.120296
f_number_of_homework_hours_a_week
0.023143
s_extent_of_school_satisfaction
0.140080
f_mother_education_level
-0.066000
f_extent_of_family_satisfaction
```

0.057850
 f_number_of_tech_at_home
 -0.016357
 f_frequency_of_books_at_home
 0.017840
 f_frequency_of_parent_involved_in_school_activi...
 0.016874
 s_gender
 -0.013262
 level_MAT
 -0.038320

	s_frequency_of_evaluations
\	
s_frequency_of_work_with_teachers	0.459210
s_frequency_of_evaluations	1.000000
s_frequency_of_internet_usage	0.097839
s_frequency_of_materials_in_class	0.346366
f_ESCS	0.000752
s_extent_of_classmates_affinity	0.208706
s_extent_of_teacher_performance	0.405473
s_extent_of_class_env	0.353524
f_mother_age	-0.010146
s_frequency_of_computer_usage	0.064127
f_number_of_homework_hours_a_week	0.013561
s_extent_of_school_satisfaction	0.246491
f_mother_education_level	0.002241
f_extent_of_family_satisfaction	0.076432
f_number_of_tech_at_home	0.004816
f_frequency_of_books_at_home	0.024959
f_frequency_of_parent_involved_in_school_activi...	0.012819
s_gender	-0.028531
level_MAT	0.033004

s_frequency_of_internet_usage \
 s_frequency_of_work_with_teachers
 0.184811
 s_frequency_of_evaluations
 0.097839
 s_frequency_of_internet_usage
 1.000000
 s_frequency_of_materials_in_class
 0.215844
 f_ESCS
 -0.015618
 s_extent_of_classmates_affinity
 -0.043809


```

s_extent_of_teacher_performance
0.020481
s_extent_of_class_env
0.022687
f_mother_age
-0.006085
s_frequency_of_computer_usage
0.371735
f_number_of_homework_hours_a_week
-0.020469
s_extent_of_school_satisfaction
-0.039423
f_mother_education_level
-0.004143
f_extent_of_family_satisfaction
0.008416
f_number_of_tech_at_home
0.040167
f_frequency_of_books_at_home
0.023850
f_frequency_of_parent_involved_in_school_activi...
0.025514
s_gender
0.084017
level_MAT
-0.036665

s_frequency_of_materials_in_class \
s_frequency_of_work_with_teachers
0.407241
s_frequency_of_evaluations
0.346366
s_frequency_of_internet_usage
0.215844
s_frequency_of_materials_in_class
1.000000
f_ESCS
-0.068230
s_extent_of_classmates_affinity
0.075477
s_extent_of_teacher_performance
0.192643
s_extent_of_class_env
0.198315
f_mother_age
-0.030786
s_frequency_of_computer_usage

```

0.215274	
f_number_of_homework_hours_a_week	
0.013483	
s_extent_of_school_satisfaction	
0.088556	
f_mother_education_level	
-0.068052	
f_extent_of_family_satisfaction	
0.050423	
f_number_of_tech_at_home	
-0.014722	
f_frequency_of_books_at_home	
-0.004838	
f_frequency_of_parent_involved_in_school_activi...	
0.039642	
s_gender	
0.000793	
level_MAT	
-0.066594	
	f_ESCS \
s_frequency_of_work_with_teachers	-0.056927
s_frequency_of_evaluations	0.000752
s_frequency_of_internet_usage	-0.015618
s_frequency_of_materials_in_class	-0.068230
f_ESCS	1.000000
s_extent_of_classmates_affinity	0.040619
s_extent_of_teacher_performance	0.039800
s_extent_of_class_env	0.005268
f_mother_age	0.261448
s_frequency_of_computer_usage	-0.044139
f_number_of_homework_hours_a_week	0.007309
s_extent_of_school_satisfaction	0.049781
f_mother_education_level	0.627503
f_extent_of_family_satisfaction	-0.087052
f_number_of_tech_at_home	0.278867
f_frequency_of_books_at_home	0.175323
f_frequency_of_parent_involved_in_school_activi...	0.047424
s_gender	0.009554
level_MAT	0.250165
s_extent_of_classmates_affinity \	
s_frequency_of_work_with_teachers	
0.115582	
s_frequency_of_evaluations	
0.208706	
s_frequency_of_internet_usage	

-0.043809
 s_frequency_of_materials_in_class
 0.075477
 f_ESCS
 0.040619
 s_extent_of_classmates_affinity
 1.000000
 s_extent_of_teacher_performance
 0.455324
 s_extent_of_class_env
 0.512753
 f_mother_age
 0.040068
 s_frequency_of_computer_usage
 -0.052212
 f_number_of_homework_hours_a_week
 -0.008666
 s_extent_of_school_satisfaction
 0.471544
 f_mother_education_level
 0.048646
 f_extent_of_family_satisfaction
 0.115887
 f_number_of_tech_at_home
 0.037754
 f_frequency_of_books_at_home
 0.021444
 f_frequency_of_parent_involved_in_school_activi...
 0.030362
 s_gender
 -0.055421
 level_MAT
 0.132252

 s_extent_of_teacher_performance \\
 s_frequency_of_work_with_teachers
 0.284243
 s_frequency_of_evaluations
 0.405473
 s_frequency_of_internet_usage
 0.020481
 s_frequency_of_materials_in_class
 0.192643
 f_ESCS
 0.039800
 s_extent_of_classmates_affinity
 0.455324

s_extent_of_teacher_performance
1.000000
s_extent_of_class_env
0.699774
f_mother_age
0.011708
s_frequency_of_computer_usage
-0.011863
f_number_of_homework_hours_a_week
-0.008315
s_extent_of_school_satisfaction
0.572676
f_mother_education_level
0.039756
f_extent_of_family_satisfaction
0.132071
f_number_of_tech_at_home
0.047904
f_frequency_of_books_at_home
0.033481
f_frequency_of_parent_involved_in_school_activi...
0.015670
s_gender
-0.044767
level_MAT
0.110412

	s_extent_of_class_env \
s_frequency_of_work_with_teachers	0.253826
s_frequency_of_evaluations	0.353524
s_frequency_of_internet_usage	0.022687
s_frequency_of_materials_in_class	0.198315
f_ESCS	0.005268
s_extent_of_classmates_affinity	0.512753
s_extent_of_teacher_performance	0.699774
s_extent_of_class_env	1.000000
f_mother_age	0.000189
s_frequency_of_computer_usage	0.014833
f_number_of_homework_hours_a_week	-0.004051
s_extent_of_school_satisfaction	0.603268
f_mother_education_level	0.000608
f_extent_of_family_satisfaction	0.135805
f_number_of_tech_at_home	0.014492
f_frequency_of_books_at_home	0.039548
f_frequency_of_parent_involved_in_school_activi...	0.026473
s_gender	-0.055349
level_MAT	0.059059

	f_mother_age \
s_frequency_of_work_with_teachers	-0.033653
s_frequency_of_evaluations	-0.010146
s_frequency_of_internet_usage	-0.006085
s_frequency_of_materials_in_class	-0.030786
f_ESCS	0.261448
s_extent_of_classmates_affinity	0.040068
s_extent_of_teacher_performance	0.011708
s_extent_of_class_env	0.000189
f_mother_age	1.000000
s_frequency_of_computer_usage	-0.018088
f_number_of_homework_hours_a_week	-0.019828
s_extent_of_school_satisfaction	0.024990
f_mother_education_level	0.250886
f_extent_of_family_satisfaction	-0.013724
f_number_of_tech_at_home	0.058864
f_frequency_of_books_at_home	0.072558
f_frequency_of_parent_involved_in_school_activi...	0.060719
s_gender	0.021320
level_MAT	0.126018
s_frequency_of_computer_usage \	
s_frequency_of_work_with_teachers	
0.120296	
s_frequency_of_evaluations	
0.064127	
s_frequency_of_internet_usage	
0.371735	
s_frequency_of_materials_in_class	
0.215274	
f_ESCS	
-0.044139	
s_extent_of_classmates_affinity	
-0.052212	
s_extent_of_teacher_performance	
-0.011863	
s_extent_of_class_env	
0.014833	
f_mother_age	
-0.018088	
s_frequency_of_computer_usage	
1.000000	
f_number_of_homework_hours_a_week	
-0.011977	
s_extent_of_school_satisfaction	
-0.051609	

```

f_mother_education_level
-0.035688
f_extent_of_family_satisfaction
0.007575
f_number_of_tech_at_home
0.012460
f_frequency_of_books_at_home
0.004609
f_frequency_of_parent_involved_in_school_activi...
0.040685
s_gender
0.019260
level_MAT
-0.072532

f_number_of_homework_hours_a_week \
s_frequency_of_work_with_teachers
0.023143
s_frequency_of_evaluations
0.013561
s_frequency_of_internet_usage
-0.020469
s_frequency_of_materials_in_class
0.013483
f_ESCS
0.007309
s_extent_of_classmates_affinity
-0.008666
s_extent_of_teacher_performance
-0.008315
s_extent_of_class_env
-0.004051
f_mother_age
-0.019828
s_frequency_of_computer_usage
-0.011977
f_number_of_homework_hours_a_week
1.000000
s_extent_of_school_satisfaction
-0.009182
f_mother_education_level
-0.008756
f_extent_of_family_satisfaction
-0.019991
f_number_of_tech_at_home
0.020601
f_frequency_of_books_at_home

```

0.057714
 f_frequency_of_parent_involved_in_school_activi...
 0.015347
 s_gender
 -0.033709
 level_MAT
 -0.035409

 s_extent_of_school_satisfaction \\
 s_frequency_of_work_with_teachers
 0.140080
 s_frequency_of_evaluations
 0.246491
 s_frequency_of_internet_usage
 -0.039423
 s_frequency_of_materials_in_class
 0.088556
 f_ESCS
 0.049781
 s_extent_of_classmates_affinity
 0.471544
 s_extent_of_teacher_performance
 0.572676
 s_extent_of_class_env
 0.603268
 f_mother_age
 0.024990
 s_frequency_of_computer_usage
 -0.051609
 f_number_of_homework_hours_a_week
 -0.009182
 s_extent_of_school_satisfaction
 1.000000
 f_mother_education_level
 0.042832
 f_extent_of_family_satisfaction
 0.136325
 f_number_of_tech_at_home
 0.039812
 f_frequency_of_books_at_home
 0.036158
 f_frequency_of_parent_involved_in_school_activi...
 0.016555
 s_gender
 -0.054659
 level_MAT
 0.123259

	f_mother_education_level \
s_frequency_of_work_with_teachers	-0.066000
s_frequency_of_evaluations	0.002241
s_frequency_of_internet_usage	-0.004143
s_frequency_of_materials_in_class	-0.068052
f_ESCS	0.627503
s_extent_of_classmates_affinity	0.048646
s_extent_of_teacher_performance	0.039756
s_extent_of_class_env	0.000608
f_mother_age	0.250886
s_frequency_of_computer_usage	-0.035688
f_number_of_homework_hours_a_week	-0.008756
s_extent_of_school_satisfaction	0.042832
f_mother_education_level	1.000000
f_extent_of_family_satisfaction	-0.065329
f_number_of_tech_at_home	0.254568
f_frequency_of_books_at_home	0.146603
f_frequency_of_parent_involved_in_school_activi...	0.020266
s_gender	0.006079
level_MAT	0.284138
f_extent_of_family_satisfaction \	
s_frequency_of_work_with_teachers	0.057850
s_frequency_of_evaluations	0.076432
s_frequency_of_internet_usage	0.008416
s_frequency_of_materials_in_class	0.050423
f_ESCS	-0.087052
s_extent_of_classmates_affinity	0.115887
s_extent_of_teacher_performance	0.132071
s_extent_of_class_env	0.135805
f_mother_age	-0.013724
s_frequency_of_computer_usage	0.007575
f_number_of_homework_hours_a_week	-0.019991
s_extent_of_school_satisfaction	0.136325

f_mother_education_level
 -0.065329
 f_extent_of_family_satisfaction
 1.000000
 f_number_of_tech_at_home
 -0.030832
 f_frequency_of_books_at_home
 0.057873
 f_frequency_of_parent_involved_in_school_activi...
 0.136940
 s_gender
 -0.024137
 level_MAT
 0.050280

	f_number_of_tech_at_home \
s_frequency_of_work_with_teachers	-0.016357
s_frequency_of_evaluations	0.004816
s_frequency_of_internet_usage	0.040167
s_frequency_of_materials_in_class	-0.014722
f_ESCS	0.278867
s_extent_of_classmates_affinity	0.037754
s_extent_of_teacher_performance	0.047904
s_extent_of_class_env	0.014492
f_mother_age	0.058864
s_frequency_of_computer_usage	0.012460
f_number_of_homework_hours_a_week	0.020601
s_extent_of_school_satisfaction	0.039812
f_mother_education_level	0.254568
f_extent_of_family_satisfaction	-0.030832
f_number_of_tech_at_home	1.000000
f_frequency_of_books_at_home	0.063726
f_frequency_of_parent_involved_in_school_activi...	-0.011387
s_gender	-0.009508
level_MAT	0.143473

f_frequency_of_books_at_home \
 s_frequency_of_work_with_teachers
 0.017840
 s_frequency_of_evaluations
 0.024959
 s_frequency_of_internet_usage
 0.023850
 s_frequency_of_materials_in_class
 -0.004838
 f_ESCS
 0.175323

```

s_extent_of_classmates_affinity
0.021444
s_extent_of_teacher_performance
0.033481
s_extent_of_class_env
0.039548
f_mother_age
0.072558
s_frequency_of_computer_usage
0.004609
f_number_of_homework_hours_a_week
0.057714
s_extent_of_school_satisfaction
0.036158
f_mother_education_level
0.146603
f_extent_of_family_satisfaction
0.057873
f_number_of_tech_at_home
0.063726
f_frequency_of_books_at_home
1.000000
f_frequency_of_parent_involved_in_school_activi...
0.179154
s_gender
-0.037040
level_MAT
0.097342

f_frequency_of_parent_involved_in_school_activities \
s_frequency_of_work_with_teachers
0.016874
s_frequency_of_evaluations
0.012819
s_frequency_of_internet_usage
0.025514
s_frequency_of_materials_in_class
0.039642
f_ESCS
0.047424
s_extent_of_classmates_affinity
0.030362
s_extent_of_teacher_performance
0.015670
s_extent_of_class_env
0.026473
f_mother_age

```

```

0.060719
s_frequency_of_computer_usage
0.040685
f_number_of_homework_hours_a_week
0.015347
s_extent_of_school_satisfaction
0.016555
f_mother_education_level
0.020266
f_extent_of_family_satisfaction
0.136940
f_number_of_tech_at_home
-0.011387
f_frequency_of_books_at_home
0.179154
f_frequency_of_parent_involved_in_school_activi...
1.000000
s_gender
-0.007350
level_MAT
0.002740

```

	s_gender	level_MAT
s_frequency_of_work_with_teachers	-0.013262	-0.038320
s_frequency_of_evaluations	-0.028531	0.033004
s_frequency_of_internet_usage	0.084017	-0.036665
s_frequency_of_materials_in_class	0.000793	-0.066594
f_ESCS	0.009554	0.250165
s_extent_of_classmates_affinity	-0.055421	0.132252
s_extent_of_teacher_performance	-0.044767	0.110412
s_extent_of_class_env	-0.055349	0.059059
f_mother_age	0.021320	0.126018
s_frequency_of_computer_usage	0.019260	-0.072532
f_number_of_homework_hours_a_week	-0.033709	-0.035409
s_extent_of_school_satisfaction	-0.054659	0.123259
f_mother_education_level	0.006079	0.284138
f_extent_of_family_satisfaction	-0.024137	0.050280
f_number_of_tech_at_home	-0.009508	0.143473
f_frequency_of_books_at_home	-0.037040	0.097342
f_frequency_of_parent_involved_in_school_activi...	-0.007350	0.002740
s_gender	1.000000	0.015317
level_MAT	0.015317	1.000000)

```
[203]: final_df = final_df[~final_df.isnull().any(axis=1)]
```

```
[204]: compute_exploratory_analysis(df=final_df, column_pattern=None,
↳missing_threshold=0.3, annot=False)
```

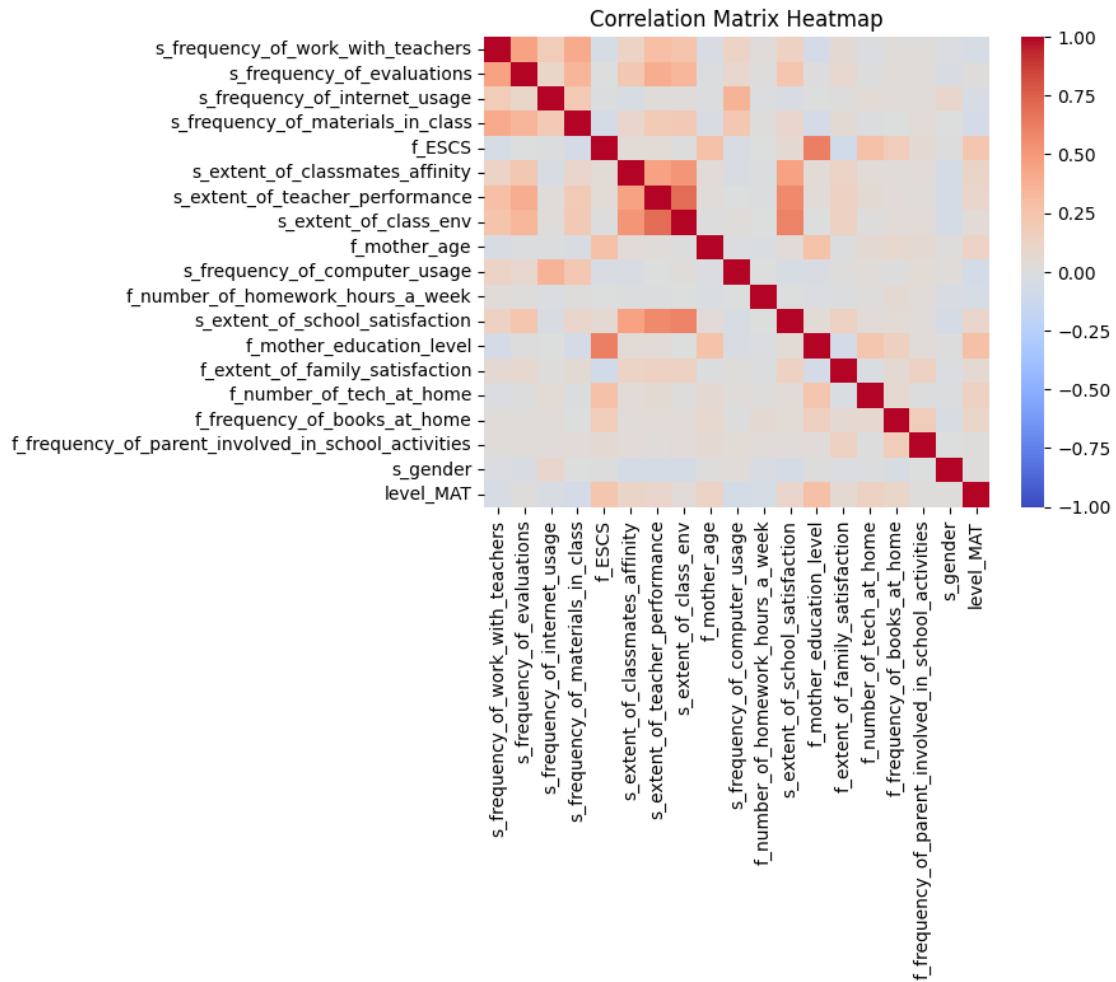
GOOD FEATURES: 19

s_frequency_of_work_with_teachers	0.0
s_frequency_of_evaluations	0.0
s_frequency_of_internet_usage	0.0
s_frequency_of_materials_in_class	0.0
f_ESCS	0.0
s_extent_of_classmates_affinity	0.0
s_extent_of_teacher_performance	0.0
s_extent_of_class_env	0.0
f_mother_age	0.0
s_frequency_of_computer_usage	0.0
f_number_of_homework_hours_a_week	0.0
s_extent_of_school_satisfaction	0.0
f_mother_education_level	0.0
f_extent_of_family_satisfaction	0.0
f_number_of_tech_at_home	0.0
f_frequency_of_books_at_home	0.0
f_frequency_of_parent_involved_in_school_activities	0.0
s_gender	0.0
level_MAT	0.0

dtype: float64

BAD FEATURES

Series([], dtype: float64)



```

s_frequency_of_work_with_teachers \
id_questionnaire
1                                0.733333
3                                0.833333
5                                0.566667
6                                0.533333
8                                0.633333
...                               ...
20421                           0.566667
20422                           0.700000
20423                           0.400000
20424                           0.500000
20427                           0.666667

```

```

s_frequency_of_evaluations  s_frequency_of_internet_usage \
id_questionnaire
1                            0.740741                        0.066667

```

3	0.750000	0.666667
5	0.888889	0.133333
6	0.851852	0.133333
8	0.703704	0.266667
...
20421	0.777778	0.600000
20422	0.518519	0.266667
20423	0.555556	0.400000
20424	0.555556	0.200000
20427	0.740741	0.133333

id_questionnaire	s_frequency_of_materials_in_class	f_ESCS \
1	0.380952	0.235340
3	1.000000	0.261451
5	0.666667	3.151773
6	0.380952	-1.627731
8	0.666667	-0.358498
...
20421	0.476190	-0.764891
20422	0.380952	-0.048524
20423	0.571429	0.072922
20424	0.523810	2.755881
20427	0.476190	1.372627

id_questionnaire	s_extent_of_classmates_affinity \
1	0.714286
3	0.833333
5	0.952381
6	0.761905
8	0.857143
...	...
20421	0.904762
20422	0.952381
20423	0.809524
20424	0.904762
20427	0.857143

id_questionnaire	s_extent_of_teacher_performance	s_extent_of_class_env \
1	0.833333	0.791667
3	0.800000	0.833333
5	1.000000	1.000000
6	0.866667	1.000000
8	0.966667	0.833333
...
20421	1.000000	1.000000

20422	0.888889	0.916667
20423	1.000000	0.666667
20424	0.766667	0.791667
20427	0.900000	0.916667

	f_mother_age	s_frequency_of_computer_usage \
id_questionnaire		
1	43.0	0.111111
3	45.0	0.666667
5	39.0	0.000000
6	25.0	0.222222
8	37.0	0.333333
...
20421	41.0	0.111111
20422	36.0	0.222222
20423	46.0	0.111111
20424	41.0	0.111111
20427	37.0	0.000000

	f_number_of_homework_hours_a_week \
id_questionnaire	
1	14.0
3	9.0
5	6.0
6	5.0
8	10.0
...	...
20421	14.0
20422	2.0
20423	8.0
20424	2.0
20427	1.0

	s_extent_of_school_satisfaction	f_mother_education_level \
id_questionnaire		
1	0.944444	7.0
3	0.888889	3.0
5	1.000000	9.0
6	1.000000	3.0
8	0.833333	7.0
...
20421	1.000000	5.0
20422	0.944444	4.0
20423	1.000000	3.0
20424	0.833333	9.0
20427	1.000000	5.0

	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
--	---------------------------------	----------------------------

id_questionnaire		
1	0.700000	3.0
3	0.291667	4.0
5	0.833333	7.0
6	0.833333	3.0
8	1.000000	3.0
...
20421	1.000000	4.0
20422	0.708333	5.0
20423	0.833333	10.0
20424	0.666667	10.0
20427	0.291667	5.0

f_frequency_of_books_at_home \	
id_questionnaire	
1	1.000000
3	0.444444
5	0.555556
6	0.333333
8	0.555556
...	...
20421	0.222222
20422	0.555556
20423	0.444444
20424	0.555556
20427	0.444444

f_frequency_of_parent_involved_in_school_activities \	
id_questionnaire	
1	0.333333
3	0.166667
5	0.250000
6	0.083333
8	0.333333
...	...
20421	0.333333
20422	0.000000
20423	0.000000
20424	0.416667
20427	0.000000

id_questionnaire	s_gender	level_MAT
1	FEMALE	3.0
3	FEMALE	1.0
5	FEMALE	4.0
6	FEMALE	1.0
8	MALE	2.0


```

...
20421      MALE      2.0
20422      FEMALE    3.0
20423      MALE      2.0
20424      MALE      3.0
20427      FEMALE    2.0

```

[10735 rows x 19 columns]

```

    ↪s_frequency_of_work_with_teachers \
s_frequency_of_work_with_teachers      1.
    ↪000000
s_frequency_of_evaluations              0.
    ↪454881
s_frequency_of_internet_usage            0.
    ↪194741
s_frequency_of_materials_in_class        0.
    ↪409276
f_ESCS                                  -0.
    ↪054629
s_extent_of_classmates_affinity          0.
    ↪125635
s_extent_of_teacher_performance          0.
    ↪284706
s_extent_of_class_env                    0.
    ↪256114
f_mother_age                            -0.
    ↪037594
s_frequency_of_computer_usage            0.
    ↪126924
f_number_of_homework_hours_a_week        0.
    ↪024993
s_extent_of_school_satisfaction          0.
    ↪136178
f_mother_education_level                 -0.
    ↪068564
f_extent_of_family_satisfaction          0.
    ↪056726
f_number_of_tech_at_home                  -0.
    ↪017440
f_frequency_of_books_at_home              0.
    ↪023068
f_frequency_of_parent_involved_in_school_activi... 0.
    ↪017585
s_gender                                -0.
    ↪018127

```

level_MAT	-0.
↪052970	

	s_frequency_of_evaluations \
s_frequency_of_work_with_teachers	0.454881
s_frequency_of_evaluations	1.000000
s_frequency_of_internet_usage	0.106232
s_frequency_of_materials_in_class	0.351169
f_ESCS	-0.000939
s_extent_of_classmates_affinity	0.221110
s_extent_of_teacher_performance	0.397563
s_extent_of_class_env	0.342555
f_mother_age	-0.012102
s_frequency_of_computer_usage	0.080795
f_number_of_homework_hours_a_week	0.014494
s_extent_of_school_satisfaction	0.237172
f_mother_education_level	0.003207
f_extent_of_family_satisfaction	0.077667
f_number_of_tech_at_home	0.000708
f_frequency_of_books_at_home	0.028708
f_frequency_of_parent_involved_in_school_activi...	0.017756
s_gender	-0.035529
level_MAT	0.014032

↪s_frequency_of_internet_usage \	
s_frequency_of_work_with_teachers	0.194741
s_frequency_of_evaluations	0.106232
s_frequency_of_internet_usage	1.000000
s_frequency_of_materials_in_class	0.206295
f_ESCS	-0.012901
s_extent_of_classmates_affinity	-0.043188
s_extent_of_teacher_performance	0.022901
s_extent_of_class_env	0.018326
f_mother_age	-0.008697
s_frequency_of_computer_usage	0.366037
f_number_of_homework_hours_a_week	-0.019355
s_extent_of_school_satisfaction	-0.033456
f_mother_education_level	-0.005847
f_extent_of_family_satisfaction	0.015617
f_number_of_tech_at_home	0.045162
f_frequency_of_books_at_home	0.025437
f_frequency_of_parent_involved_in_school_activi...	0.026814
s_gender	0.086346
level_MAT	-0.039068

↪s_frequency_of_materials_in_class \

s_frequency_of_work_with_teachers	0.
↪409276	
s_frequency_of_evaluations	0.
↪351169	
s_frequency_of_internet_usage	0.
↪206295	
s_frequency_of_materials_in_class	1.
↪000000	
f_ESCS	-0.
↪068555	
s_extent_of_classmates_affinity	0.
↪099953	
s_extent_of_teacher_performance	0.
↪201062	
s_extent_of_class_env	0.
↪211431	
f_mother_age	-0.
↪033654	
s_frequency_of_computer_usage	0.
↪229184	
f_number_of_homework_hours_a_week	0.
↪013167	
s_extent_of_school_satisfaction	0.
↪099923	
f_mother_education_level	-0.
↪069164	
f_extent_of_family_satisfaction	0.
↪051873	
f_number_of_tech_at_home	-0.
↪014363	
f_frequency_of_books_at_home	-0.
↪007351	
f_frequency_of_parent_involved_in_school_activi...	0.
↪038761	
s_gender	-0.
↪002908	
level_MAT	-0.
↪065283	

	f_ESCS \
s_frequency_of_work_with_teachers	-0.054629
s_frequency_of_evaluations	-0.000939
s_frequency_of_internet_usage	-0.012901
s_frequency_of_materials_in_class	-0.068555

f_ESCS	1.000000
s_extent_of_classmates_affinity	0.039126
s_extent_of_teacher_performance	0.042015
s_extent_of_class_env	0.006251
f_mother_age	0.266225
s_frequency_of_computer_usage	-0.037662
f_number_of_homework_hours_a_week	-0.003042
s_extent_of_school_satisfaction	0.056717
f_mother_education_level	0.621918
f_extent_of_family_satisfaction	-0.086076
f_number_of_tech_at_home	0.268389
f_frequency_of_books_at_home	0.177241
f_frequency_of_parent_involved_in_school_activi...	0.058164
s_gender	0.006280
level_MAT	0.241087

□

↪s_extent_of_classmates_affinity \	
s_frequency_of_work_with_teachers	0.
↪125635	
s_frequency_of_evaluations	0.
↪221110	
s_frequency_of_internet_usage	-0.
↪043188	
s_frequency_of_materials_in_class	0.
↪099953	
f_ESCS	0.
↪039126	
s_extent_of_classmates_affinity	1.
↪000000	
s_extent_of_teacher_performance	0.
↪457112	
s_extent_of_class_env	0.
↪518564	
f_mother_age	0.
↪038003	
s_frequency_of_computer_usage	-0.
↪035870	
f_number_of_homework_hours_a_week	-0.
↪010784	
s_extent_of_school_satisfaction	0.
↪459867	
f_mother_education_level	0.
↪045788	
f_extent_of_family_satisfaction	0.
↪123810	

f_number_of_tech_at_home	0.
↪033261	
f_frequency_of_books_at_home	0.
↪024182	
f_frequency_of_parent_involved_in_school_activi...	0.034539
s_gender	-0.
↪056301	
level_MAT	0.
↪113418	
↪s_extent_of_teacher_performance \	
s_frequency_of_work_with_teachers	0.
↪284706	
s_frequency_of_evaluations	0.
↪397563	
s_frequency_of_internet_usage	0.
↪022901	
s_frequency_of_materials_in_class	0.
↪201062	
f_ESCS	0.
↪042015	
s_extent_of_classmates_affinity	0.
↪457112	
s_extent_of_teacher_performance	1.
↪000000	
s_extent_of_class_env	0.
↪700419	
f_mother_age	0.
↪017997	
s_frequency_of_computer_usage	-0.
↪001503	
f_number_of_homework_hours_a_week	-0.
↪010408	
s_extent_of_school_satisfaction	0.
↪572681	
f_mother_education_level	0.
↪036255	
f_extent_of_family_satisfaction	0.
↪136448	
f_number_of_tech_at_home	0.
↪047870	
f_frequency_of_books_at_home	0.
↪036880	
f_frequency_of_parent_involved_in_school_activi...	0.023407

s_gender	-0.
↪054769	
level_MAT	0.
↪094313	

	s_extent_of_class_env \
s_frequency_of_work_with_teachers	0.256114
s_frequency_of_evaluations	0.342555
s_frequency_of_internet_usage	0.018326
s_frequency_of_materials_in_class	0.211431
f_ESCS	0.006251
s_extent_of_classmates_affinity	0.518564
s_extent_of_teacher_performance	0.700419
s_extent_of_class_env	1.000000
f_mother_age	0.002250
s_frequency_of_computer_usage	0.021759
f_number_of_homework_hours_a_week	-0.003077
s_extent_of_school_satisfaction	0.605949
f_mother_education_level	-0.002109
f_extent_of_family_satisfaction	0.140228
f_number_of_tech_at_home	0.014887
f_frequency_of_books_at_home	0.038779
f_frequency_of_parent_involved_in_school_activi...	0.035032
s_gender	-0.059657
level_MAT	0.035256

	f_mother_age \
s_frequency_of_work_with_teachers	-0.037594
s_frequency_of_evaluations	-0.012102
s_frequency_of_internet_usage	-0.008697
s_frequency_of_materials_in_class	-0.033654
f_ESCS	0.266225
s_extent_of_classmates_affinity	0.038003
s_extent_of_teacher_performance	0.017997
s_extent_of_class_env	0.002250
f_mother_age	1.000000
s_frequency_of_computer_usage	-0.021918
f_number_of_homework_hours_a_week	-0.027460
s_extent_of_school_satisfaction	0.031414
f_mother_education_level	0.257911
f_extent_of_family_satisfaction	-0.014695
f_number_of_tech_at_home	0.056860
f_frequency_of_books_at_home	0.073622
f_frequency_of_parent_involved_in_school_activi...	0.060336
s_gender	0.014927
level_MAT	0.127296

↪s_frequency_of_computer_usage \

s_frequency_of_work_with_teachers	0.126924
s_frequency_of_evaluations	0.080795
s_frequency_of_internet_usage	0.366037
s_frequency_of_materials_in_class	0.229184
f_ESCS	-0.037662
s_extent_of_classmates_affinity	-0.035870
s_extent_of_teacher_performance	-0.001503
s_extent_of_class_env	0.021759
f_mother_age	-0.021918
s_frequency_of_computer_usage	1.000000
f_number_of_homework_hours_a_week	-0.007628
s_extent_of_school_satisfaction	-0.040981
f_mother_education_level	-0.034161
f_extent_of_family_satisfaction	0.015375
f_number_of_tech_at_home	0.020655
f_frequency_of_books_at_home	0.001636
f_frequency_of_parent_involved_in_school_activi...	0.043040
s_gender	0.021234
level_MAT	-0.072849

↪f_number_of_homework_hours_a_week \

s_frequency_of_work_with_teachers	0.
↪024993	
s_frequency_of_evaluations	0.
↪014494	
s_frequency_of_internet_usage	-0.
↪019355	
s_frequency_of_materials_in_class	0.
↪013167	
f_ESCS	-0.
↪003042	
s_extent_of_classmates_affinity	-0.
↪010784	
s_extent_of_teacher_performance	-0.
↪010408	
s_extent_of_class_env	-0.
↪003077	
f_mother_age	-0.
↪027460	
s_frequency_of_computer_usage	-0.
↪007628	
f_number_of_homework_hours_a_week	1.
↪000000	

s_extent_of_school_satisfaction	-0.
↪007675	
f_mother_education_level	-0.
↪021123	
f_extent_of_family_satisfaction	-0.
↪014983	
f_number_of_tech_at_home	0.
↪013894	
f_frequency_of_books_at_home	0.
↪060948	
f_frequency_of_parent_involved_in_school_activi...	0.
↪025651	
s_gender	-0.
↪035562	
level_MAT	-0.
↪049336	

↪s_extent_of_school_satisfaction \	
s_frequency_of_work_with_teachers	0.
↪136178	
s_frequency_of_evaluations	0.
↪237172	
s_frequency_of_internet_usage	-0.
↪033456	
s_frequency_of_materials_in_class	0.
↪099923	
f_ESCS	0.
↪056717	
s_extent_of_classmates_affinity	0.
↪459867	
s_extent_of_teacher_performance	0.
↪572681	
s_extent_of_class_env	0.
↪605949	
f_mother_age	0.
↪031414	
s_frequency_of_computer_usage	-0.
↪040981	
f_number_of_homework_hours_a_week	-0.
↪007675	
s_extent_of_school_satisfaction	1.
↪000000	
f_mother_education_level	0.
↪046073	

f_extent_of_family_satisfaction	0.
↪144083	
f_number_of_tech_at_home	0.
↪035594	
f_frequency_of_books_at_home	0.
↪042360	
f_frequency_of_parent_involved_in_school_activi...	0.024147
s_gender	-0.
↪057120	
level_MAT	0.
↪099471	

	f_mother_education_level \
s_frequency_of_work_with_teachers	-0.068564
s_frequency_of_evaluations	0.003207
s_frequency_of_internet_usage	-0.005847
s_frequency_of_materials_in_class	-0.069164
f_ESCS	0.621918
s_extent_of_classmates_affinity	0.045788
s_extent_of_teacher_performance	0.036255
s_extent_of_class_env	-0.002109
f_mother_age	0.257911
s_frequency_of_computer_usage	-0.034161
f_number_of_homework_hours_a_week	-0.021123
s_extent_of_school_satisfaction	0.046073
f_mother_education_level	1.000000
f_extent_of_family_satisfaction	-0.067264
f_number_of_tech_at_home	0.239962
f_frequency_of_books_at_home	0.151003
f_frequency_of_parent_involved_in_school_activi...	0.027669
s_gender	0.002725
level_MAT	0.278425

↪

↪f_extent_of_family_satisfaction \	
s_frequency_of_work_with_teachers	0.
↪056726	
s_frequency_of_evaluations	0.
↪077667	
s_frequency_of_internet_usage	0.
↪015617	
s_frequency_of_materials_in_class	0.
↪051873	
f_ESCS	-0.
↪086076	
s_extent_of_classmates_affinity	0.
↪123810	

s_extent_of_teacher_performance	0.
↪136448	
s_extent_of_class_env	0.
↪140228	
f_mother_age	-0.
↪014695	
s_frequency_of_computer_usage	0.
↪015375	
f_number_of_homework_hours_a_week	-0.
↪014983	
s_extent_of_school_satisfaction	0.
↪144083	
f_mother_education_level	-0.
↪067264	
f_extent_of_family_satisfaction	1.
↪000000	
f_number_of_tech_at_home	-0.
↪029384	
f_frequency_of_books_at_home	0.
↪061662	
f_frequency_of_parent_involved_in_school_activi...	0.137134
s_gender	-0.
↪017418	
level_MAT	0.
↪056351	

	f_number_of_tech_at_home \
s_frequency_of_work_with_teachers	-0.017440
s_frequency_of_evaluations	0.000708
s_frequency_of_internet_usage	0.045162
s_frequency_of_materials_in_class	-0.014363
f_ESCS	0.268389
s_extent_of_classmates_affinity	0.033261
s_extent_of_teacher_performance	0.047870
s_extent_of_class_env	0.014887
f_mother_age	0.056860
s_frequency_of_computer_usage	0.020655
f_number_of_homework_hours_a_week	0.013894
s_extent_of_school_satisfaction	0.035594
f_mother_education_level	0.239962
f_extent_of_family_satisfaction	-0.029384
f_number_of_tech_at_home	1.000000
f_frequency_of_books_at_home	0.066989
f_frequency_of_parent_involved_in_school_activi...	0.000961
s_gender	-0.013463
level_MAT	0.136941

	f_frequency_of_books_at_home	
↪ \		
s_frequency_of_work_with_teachers		0.023068
s_frequency_of_evaluations		0.028708
s_frequency_of_internet_usage		0.025437
s_frequency_of_materials_in_class		-0.007351
f_ESCS		0.177241
s_extent_of_classmates_affinity		0.024182
s_extent_of_teacher_performance		0.036880
s_extent_of_class_env		0.038779
f_mother_age		0.073622
s_frequency_of_computer_usage		0.001636
f_number_of_homework_hours_a_week		0.060948
s_extent_of_school_satisfaction		0.042360
f_mother_education_level		0.151003
f_extent_of_family_satisfaction		0.061662
f_number_of_tech_at_home		0.066989
f_frequency_of_books_at_home		1.000000
f_frequency_of_parent_involved_in_school_activi...		0.175514
s_gender		-0.039330
level_MAT		0.105491

	f_frequency_of_parent_involved_in_school_activities	\
s_frequency_of_work_with_teachers		
↪	0.017585	
s_frequency_of_evaluations		
↪	0.017756	
s_frequency_of_internet_usage		
↪	0.026814	
s_frequency_of_materials_in_class		
↪	0.038761	
f_ESCS		
↪	0.058164	
s_extent_of_classmates_affinity		
↪	0.034539	
s_extent_of_teacher_performance		
↪	0.023407	
s_extent_of_class_env		
↪	0.035032	
f_mother_age		
↪	0.060336	
s_frequency_of_computer_usage		
↪	0.043040	
f_number_of_homework_hours_a_week		
↪	0.025651	

s_extent_of_school_satisfaction	0.024147	□
f_mother_education_level	0.027669	□
f_extent_of_family_satisfaction	0.137134	□
f_number_of_tech_at_home	0.000961	□
f_frequency_of_books_at_home	0.175514	□
f_frequency_of_parent_involved_in_school_activi...	1.000000	□
s_gender	-0.002085	□
level_MAT	0.009686	□

	s_gender	level_MAT
s_frequency_of_work_with_teachers	-0.018127	-0.052970
s_frequency_of_evaluations	-0.035529	0.014032
s_frequency_of_internet_usage	0.086346	-0.039068
s_frequency_of_materials_in_class	-0.002908	-0.065283
f_ESCS	0.006280	0.241087
s_extent_of_classmates_affinity	-0.056301	0.113418
s_extent_of_teacher_performance	-0.054769	0.094313
s_extent_of_class_env	-0.059657	0.035256
f_mother_age	0.014927	0.127296
s_frequency_of_computer_usage	0.021234	-0.072849
f_number_of_homework_hours_a_week	-0.035562	-0.049336
s_extent_of_school_satisfaction	-0.057120	0.099471
f_mother_education_level	0.002725	0.278425
f_extent_of_family_satisfaction	-0.017418	0.056351
f_number_of_tech_at_home	-0.013463	0.136941
f_frequency_of_books_at_home	-0.039330	0.105491
f_frequency_of_parent_involved_in_school_activi...	-0.002085	0.009686
s_gender	1.000000	0.013750
level_MAT	0.013750	1.000000

```
[204]: (
    id_questionnaire
    1          0.733333
    3          0.833333
    5          0.566667
    6          0.533333
    8          0.633333
    ...
    20421     0.566667
```

20422	0.700000
20423	0.400000
20424	0.500000
20427	0.666667

	s_frequency_of_evaluations	s_frequency_of_internet_usage \
id_questionnaire		
1	0.740741	0.066667
3	0.750000	0.666667
5	0.888889	0.133333
6	0.851852	0.133333
8	0.703704	0.266667
...
20421	0.777778	0.600000
20422	0.518519	0.266667
20423	0.555556	0.400000
20424	0.555556	0.200000
20427	0.740741	0.133333

	s_frequency_of_materials_in_class	f_ESCS \
id_questionnaire		
1	0.380952	0.235340
3	1.000000	0.261451
5	0.666667	3.151773
6	0.380952	-1.627731
8	0.666667	-0.358498
...
20421	0.476190	-0.764891
20422	0.380952	-0.048524
20423	0.571429	0.072922
20424	0.523810	2.755881
20427	0.476190	1.372627

	s_extent_of_classmates_affinity \
id_questionnaire	
1	0.714286
3	0.833333
5	0.952381
6	0.761905
8	0.857143
...	...
20421	0.904762
20422	0.952381
20423	0.809524
20424	0.904762
20427	0.857143

id_questionnaire	s_extent_of_teacher_performance	s_extent_of_class_env \
1	0.833333	0.791667
3	0.800000	0.833333
5	1.000000	1.000000
6	0.866667	1.000000
8	0.966667	0.833333
...
20421	1.000000	1.000000
20422	0.888889	0.916667
20423	1.000000	0.666667
20424	0.766667	0.791667
20427	0.900000	0.916667

id_questionnaire	f_mother_age	s_frequency_of_computer_usage \
1	43.0	0.111111
3	45.0	0.666667
5	39.0	0.000000
6	25.0	0.222222
8	37.0	0.333333
...
20421	41.0	0.111111
20422	36.0	0.222222
20423	46.0	0.111111
20424	41.0	0.111111
20427	37.0	0.000000

id_questionnaire	f_number_of_homework_hours_a_week \
1	14.0
3	9.0
5	6.0
6	5.0
8	10.0
...	...
20421	14.0
20422	2.0
20423	8.0
20424	2.0
20427	1.0

id_questionnaire	s_extent_of_school_satisfaction	f_mother_education_level \
1	0.944444	7.0
3	0.888889	3.0
5	1.000000	9.0

6	1.000000	3.0
8	0.833333	7.0
...
20421	1.000000	5.0
20422	0.944444	4.0
20423	1.000000	3.0
20424	0.833333	9.0
20427	1.000000	5.0

id_questionnaire	f_extent_of_family_satisfaction	f_number_of_tech_at_home \
1	0.700000	3.0
3	0.291667	4.0
5	0.833333	7.0
6	0.833333	3.0
8	1.000000	3.0
...
20421	1.000000	4.0
20422	0.708333	5.0
20423	0.833333	10.0
20424	0.666667	10.0
20427	0.291667	5.0

id_questionnaire	f_frequency_of_books_at_home \
1	1.000000
3	0.444444
5	0.555556
6	0.333333
8	0.555556
...	...
20421	0.222222
20422	0.555556
20423	0.444444
20424	0.555556
20427	0.444444

id_questionnaire	f_frequency_of_parent_involved_in_school_activities \
1	0.333333
3	0.166667
5	0.250000
6	0.083333
8	0.333333
...	...
20421	0.333333
20422	0.000000

20423	0.000000
20424	0.416667
20427	0.000000

	s_gender	level_MAT
id_questionnaire		
1	FEMALE	3.0
3	FEMALE	1.0
5	FEMALE	4.0
6	FEMALE	1.0
8	MALE	2.0
...
20421	MALE	2.0
20422	FEMALE	3.0
20423	MALE	2.0
20424	MALE	3.0
20427	FEMALE	2.0

```
[10735 rows x 19 columns],
s_frequency_of_work_with_teachers \
s_frequency_of_work_with_teachers
1.000000
s_frequency_of_evaluations
0.454881
s_frequency_of_internet_usage
0.194741
s_frequency_of_materials_in_class
0.409276
f_ESCS
-0.054629
s_extent_of_classmates_affinity
0.125635
s_extent_of_teacher_performance
0.284706
s_extent_of_class_env
0.256114
f_mother_age
-0.037594
s_frequency_of_computer_usage
0.126924
f_number_of_homework_hours_a_week
0.024993
s_extent_of_school_satisfaction
0.136178
f_mother_education_level
-0.068564
f_extent_of_family_satisfaction
```


0.056726
 f_number_of_tech_at_home
 -0.017440
 f_frequency_of_books_at_home
 0.023068
 f_frequency_of_parent_involved_in_school_activi...
 0.017585
 s_gender
 -0.018127
 level_MAT
 -0.052970

	s_frequency_of_evaluations
\	
s_frequency_of_work_with_teachers	0.454881
s_frequency_of_evaluations	1.000000
s_frequency_of_internet_usage	0.106232
s_frequency_of_materials_in_class	0.351169
f_ESCS	-0.000939
s_extent_of_classmates_affinity	0.221110
s_extent_of_teacher_performance	0.397563
s_extent_of_class_env	0.342555
f_mother_age	-0.012102
s_frequency_of_computer_usage	0.080795
f_number_of_homework_hours_a_week	0.014494
s_extent_of_school_satisfaction	0.237172
f_mother_education_level	0.003207
f_extent_of_family_satisfaction	0.077667
f_number_of_tech_at_home	0.000708
f_frequency_of_books_at_home	0.028708
f_frequency_of_parent_involved_in_school_activi...	0.017756
s_gender	-0.035529
level_MAT	0.014032

s_frequency_of_internet_usage \
 s_frequency_of_work_with_teachers
 0.194741
 s_frequency_of_evaluations
 0.106232
 s_frequency_of_internet_usage
 1.000000
 s_frequency_of_materials_in_class
 0.206295
 f_ESCS
 -0.012901
 s_extent_of_classmates_affinity
 -0.043188

```

s_extent_of_teacher_performance
0.022901
s_extent_of_class_env
0.018326
f_mother_age
-0.008697
s_frequency_of_computer_usage
0.366037
f_number_of_homework_hours_a_week
-0.019355
s_extent_of_school_satisfaction
-0.033456
f_mother_education_level
-0.005847
f_extent_of_family_satisfaction
0.015617
f_number_of_tech_at_home
0.045162
f_frequency_of_books_at_home
0.025437
f_frequency_of_parent_involved_in_school_activi...
0.026814
s_gender
0.086346
level_MAT
-0.039068

s_frequency_of_materials_in_class \
s_frequency_of_work_with_teachers
0.409276
s_frequency_of_evaluations
0.351169
s_frequency_of_internet_usage
0.206295
s_frequency_of_materials_in_class
1.000000
f_ESCS
-0.068555
s_extent_of_classmates_affinity
0.099953
s_extent_of_teacher_performance
0.201062
s_extent_of_class_env
0.211431
f_mother_age
-0.033654
s_frequency_of_computer_usage

```

0.229184	
f_number_of_homework_hours_a_week	
0.013167	
s_extent_of_school_satisfaction	
0.099923	
f_mother_education_level	
-0.069164	
f_extent_of_family_satisfaction	
0.051873	
f_number_of_tech_at_home	
-0.014363	
f_frequency_of_books_at_home	
-0.007351	
f_frequency_of_parent_involved_in_school_activi...	
0.038761	
s_gender	
-0.002908	
level_MAT	
-0.065283	
	f_ESCS \
s_frequency_of_work_with_teachers	-0.054629
s_frequency_of_evaluations	-0.000939
s_frequency_of_internet_usage	-0.012901
s_frequency_of_materials_in_class	-0.068555
f_ESCS	1.000000
s_extent_of_classmates_affinity	0.039126
s_extent_of_teacher_performance	0.042015
s_extent_of_class_env	0.006251
f_mother_age	0.266225
s_frequency_of_computer_usage	-0.037662
f_number_of_homework_hours_a_week	-0.003042
s_extent_of_school_satisfaction	0.056717
f_mother_education_level	0.621918
f_extent_of_family_satisfaction	-0.086076
f_number_of_tech_at_home	0.268389
f_frequency_of_books_at_home	0.177241
f_frequency_of_parent_involved_in_school_activi...	0.058164
s_gender	0.006280
level_MAT	0.241087
s_extent_of_classmates_affinity \	
s_frequency_of_work_with_teachers	
0.125635	
s_frequency_of_evaluations	
0.221110	
s_frequency_of_internet_usage	

-0.043188
 s_frequency_of_materials_in_class
 0.099953
 f_ESCS
 0.039126
 s_extent_of_classmates_affinity
 1.000000
 s_extent_of_teacher_performance
 0.457112
 s_extent_of_class_env
 0.518564
 f_mother_age
 0.038003
 s_frequency_of_computer_usage
 -0.035870
 f_number_of_homework_hours_a_week
 -0.010784
 s_extent_of_school_satisfaction
 0.459867
 f_mother_education_level
 0.045788
 f_extent_of_family_satisfaction
 0.123810
 f_number_of_tech_at_home
 0.033261
 f_frequency_of_books_at_home
 0.024182
 f_frequency_of_parent_involved_in_school_activi...
 0.034539
 s_gender
 -0.056301
 level_MAT
 0.113418

 s_extent_of_teacher_performance \
 s_frequency_of_work_with_teachers
 0.284706
 s_frequency_of_evaluations
 0.397563
 s_frequency_of_internet_usage
 0.022901
 s_frequency_of_materials_in_class
 0.201062
 f_ESCS
 0.042015
 s_extent_of_classmates_affinity
 0.457112

s_extent_of_teacher_performance
 1.000000
 s_extent_of_class_env
 0.700419
 f_mother_age
 0.017997
 s_frequency_of_computer_usage
 -0.001503
 f_number_of_homework_hours_a_week
 -0.010408
 s_extent_of_school_satisfaction
 0.572681
 f_mother_education_level
 0.036255
 f_extent_of_family_satisfaction
 0.136448
 f_number_of_tech_at_home
 0.047870
 f_frequency_of_books_at_home
 0.036880
 f_frequency_of_parent_involved_in_school_activi...
 0.023407
 s_gender
 -0.054769
 level_MAT
 0.094313

	s_extent_of_class_env \
s_frequency_of_work_with_teachers	0.256114
s_frequency_of_evaluations	0.342555
s_frequency_of_internet_usage	0.018326
s_frequency_of_materials_in_class	0.211431
f_ESCS	0.006251
s_extent_of_classmates_affinity	0.518564
s_extent_of_teacher_performance	0.700419
s_extent_of_class_env	1.000000
f_mother_age	0.002250
s_frequency_of_computer_usage	0.021759
f_number_of_homework_hours_a_week	-0.003077
s_extent_of_school_satisfaction	0.605949
f_mother_education_level	-0.002109
f_extent_of_family_satisfaction	0.140228
f_number_of_tech_at_home	0.014887
f_frequency_of_books_at_home	0.038779
f_frequency_of_parent_involved_in_school_activi...	0.035032
s_gender	-0.059657
level_MAT	0.035256

s_frequency_of_work_with_teachers	f_mother_age \	-0.037594
s_frequency_of_evaluations		-0.012102
s_frequency_of_internet_usage		-0.008697
s_frequency_of_materials_in_class		-0.033654
f_ESCS		0.266225
s_extent_of_classmates_affinity		0.038003
s_extent_of_teacher_performance		0.017997
s_extent_of_class_env		0.002250
f_mother_age		1.000000
s_frequency_of_computer_usage		-0.021918
f_number_of_homework_hours_a_week		-0.027460
s_extent_of_school_satisfaction		0.031414
f_mother_education_level		0.257911
f_extent_of_family_satisfaction		-0.014695
f_number_of_tech_at_home		0.056860
f_frequency_of_books_at_home		0.073622
f_frequency_of_parent_involved_in_school_activi...		0.060336
s_gender		0.014927
level_MAT		0.127296

s_frequency_of_computer_usage \	
s_frequency_of_work_with_teachers	0.126924
s_frequency_of_evaluations	0.080795
s_frequency_of_internet_usage	0.366037
s_frequency_of_materials_in_class	0.229184
f_ESCS	-0.037662
s_extent_of_classmates_affinity	-0.035870
s_extent_of_teacher_performance	-0.001503
s_extent_of_class_env	0.021759
f_mother_age	-0.021918
s_frequency_of_computer_usage	1.000000
f_number_of_homework_hours_a_week	-0.007628
s_extent_of_school_satisfaction	-0.040981

```

f_mother_education_level
-0.034161
f_extent_of_family_satisfaction
0.015375
f_number_of_tech_at_home
0.020655
f_frequency_of_books_at_home
0.001636
f_frequency_of_parent_involved_in_school_activi...
0.043040
s_gender
0.021234
level_MAT
-0.072849

f_number_of_homework_hours_a_week \
s_frequency_of_work_with_teachers
0.024993
s_frequency_of_evaluations
0.014494
s_frequency_of_internet_usage
-0.019355
s_frequency_of_materials_in_class
0.013167
f_ESCS
-0.003042
s_extent_of_classmates_affinity
-0.010784
s_extent_of_teacher_performance
-0.010408
s_extent_of_class_env
-0.003077
f_mother_age
-0.027460
s_frequency_of_computer_usage
-0.007628
f_number_of_homework_hours_a_week
1.000000
s_extent_of_school_satisfaction
-0.007675
f_mother_education_level
-0.021123
f_extent_of_family_satisfaction
-0.014983
f_number_of_tech_at_home
0.013894
f_frequency_of_books_at_home

```

0.060948
 f_frequency_of_parent_involved_in_school_activi...
 0.025651
 s_gender
 -0.035562
 level_MAT
 -0.049336

 s_extent_of_school_satisfaction \\
 s_frequency_of_work_with_teachers
 0.136178
 s_frequency_of_evaluations
 0.237172
 s_frequency_of_internet_usage
 -0.033456
 s_frequency_of_materials_in_class
 0.099923
 f_ESCS
 0.056717
 s_extent_of_classmates_affinity
 0.459867
 s_extent_of_teacher_performance
 0.572681
 s_extent_of_class_env
 0.605949
 f_mother_age
 0.031414
 s_frequency_of_computer_usage
 -0.040981
 f_number_of_homework_hours_a_week
 -0.007675
 s_extent_of_school_satisfaction
 1.000000
 f_mother_education_level
 0.046073
 f_extent_of_family_satisfaction
 0.144083
 f_number_of_tech_at_home
 0.035594
 f_frequency_of_books_at_home
 0.042360
 f_frequency_of_parent_involved_in_school_activi...
 0.024147
 s_gender
 -0.057120
 level_MAT
 0.099471

	f_mother_education_level \
s_frequency_of_work_with_teachers	-0.068564
s_frequency_of_evaluations	0.003207
s_frequency_of_internet_usage	-0.005847
s_frequency_of_materials_in_class	-0.069164
f_ESCS	0.621918
s_extent_of_classmates_affinity	0.045788
s_extent_of_teacher_performance	0.036255
s_extent_of_class_env	-0.002109
f_mother_age	0.257911
s_frequency_of_computer_usage	-0.034161
f_number_of_homework_hours_a_week	-0.021123
s_extent_of_school_satisfaction	0.046073
f_mother_education_level	1.000000
f_extent_of_family_satisfaction	-0.067264
f_number_of_tech_at_home	0.239962
f_frequency_of_books_at_home	0.151003
f_frequency_of_parent_involved_in_school_activi...	0.027669
s_gender	0.002725
level_MAT	0.278425

f_extent_of_family_satisfaction \
s_frequency_of_work_with_teachers
0.056726
s_frequency_of_evaluations
0.077667
s_frequency_of_internet_usage
0.015617
s_frequency_of_materials_in_class
0.051873
f_ESCS
-0.086076
s_extent_of_classmates_affinity
0.123810
s_extent_of_teacher_performance
0.136448
s_extent_of_class_env
0.140228
f_mother_age
-0.014695
s_frequency_of_computer_usage
0.015375
f_number_of_homework_hours_a_week
-0.014983
s_extent_of_school_satisfaction
0.144083

f_mother_education_level
 -0.067264
 f_extent_of_family_satisfaction
 1.000000
 f_number_of_tech_at_home
 -0.029384
 f_frequency_of_books_at_home
 0.061662
 f_frequency_of_parent_involved_in_school_activi...
 0.137134
 s_gender
 -0.017418
 level_MAT
 0.056351

	f_number_of_tech_at_home \
s_frequency_of_work_with_teachers	-0.017440
s_frequency_of_evaluations	0.000708
s_frequency_of_internet_usage	0.045162
s_frequency_of_materials_in_class	-0.014363
f_ESCS	0.268389
s_extent_of_classmates_affinity	0.033261
s_extent_of_teacher_performance	0.047870
s_extent_of_class_env	0.014887
f_mother_age	0.056860
s_frequency_of_computer_usage	0.020655
f_number_of_homework_hours_a_week	0.013894
s_extent_of_school_satisfaction	0.035594
f_mother_education_level	0.239962
f_extent_of_family_satisfaction	-0.029384
f_number_of_tech_at_home	1.000000
f_frequency_of_books_at_home	0.066989
f_frequency_of_parent_involved_in_school_activi...	0.000961
s_gender	-0.013463
level_MAT	0.136941

f_frequency_of_books_at_home \
 s_frequency_of_work_with_teachers
 0.023068
 s_frequency_of_evaluations
 0.028708
 s_frequency_of_internet_usage
 0.025437
 s_frequency_of_materials_in_class
 -0.007351
 f_ESCS
 0.177241

```

s_extent_of_classmates_affinity
0.024182
s_extent_of_teacher_performance
0.036880
s_extent_of_class_env
0.038779
f_mother_age
0.073622
s_frequency_of_computer_usage
0.001636
f_number_of_homework_hours_a_week
0.060948
s_extent_of_school_satisfaction
0.042360
f_mother_education_level
0.151003
f_extent_of_family_satisfaction
0.061662
f_number_of_tech_at_home
0.066989
f_frequency_of_books_at_home
1.000000
f_frequency_of_parent_involved_in_school_activi...
0.175514
s_gender
-0.039330
level_MAT
0.105491

f_frequency_of_parent_involved_in_school_activities \
s_frequency_of_work_with_teachers
0.017585
s_frequency_of_evaluations
0.017756
s_frequency_of_internet_usage
0.026814
s_frequency_of_materials_in_class
0.038761
f_ESCS
0.058164
s_extent_of_classmates_affinity
0.034539
s_extent_of_teacher_performance
0.023407
s_extent_of_class_env
0.035032
f_mother_age

```

```

0.060336
  s_frequency_of_computer_usage
0.043040
  f_number_of_homework_hours_a_week
0.025651
  s_extent_of_school_satisfaction
0.024147
  f_mother_education_level
0.027669
  f_extent_of_family_satisfaction
0.137134
  f_number_of_tech_at_home
0.000961
  f_frequency_of_books_at_home
0.175514
  f_frequency_of_parent_involved_in_school_activi...
1.000000
  s_gender
-0.002085
  level_MAT
0.009686

```

	s_gender	level_MAT
s_frequency_of_work_with_teachers	-0.018127	-0.052970
s_frequency_of_evaluations	-0.035529	0.014032
s_frequency_of_internet_usage	0.086346	-0.039068
s_frequency_of_materials_in_class	-0.002908	-0.065283
f_ESCS	0.006280	0.241087
s_extent_of_classmates_affinity	-0.056301	0.113418
s_extent_of_teacher_performance	-0.054769	0.094313
s_extent_of_class_env	-0.059657	0.035256
f_mother_age	0.014927	0.127296
s_frequency_of_computer_usage	0.021234	-0.072849
f_number_of_homework_hours_a_week	-0.035562	-0.049336
s_extent_of_school_satisfaction	-0.057120	0.099471
f_mother_education_level	0.002725	0.278425
f_extent_of_family_satisfaction	-0.017418	0.056351
f_number_of_tech_at_home	-0.013463	0.136941
f_frequency_of_books_at_home	-0.039330	0.105491
f_frequency_of_parent_involved_in_school_activi...	-0.002085	0.009686
s_gender	1.000000	0.013750
level_MAT	0.013750	1.000000)

```
[356]: final_df.to_csv(os.path.join(ull_final_path, "final_df_v1.csv"))
```