

Informações Iniciais dos Datasets

```
# Verificação dos arquivos:
\ / - Arquivo encontrado: data/car+evaluation/car.data
\ / - Arquivo encontrado: data/student/student-mat.csv
\ / - Arquivo encontrado: data/energy+efficiency/ENB2012_data.xlsx
```

==== Car Evaluation Data =====

Primeiras linhas:

	buying	maint	doors	persons	lug_boot	safety	class
0	vhigh	vhigh	2	2	small	low	unacc
1	vhigh	vhigh	2	2	small	med	unacc
2	vhigh	vhigh	2	2	small	high	unacc
3	vhigh	vhigh	2	2	med	low	unacc
4	vhigh	vhigh	2	2	med	med	unacc

Dimensões (linhas, colunas): (1728, 7)

Valores nulos por coluna:

buying	0
maint	0
doors	0
persons	0
lug_boot	0
safety	0
class	0

==== Student Math Data =====

Primeiras linhas:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	reason	guardian	traveltime	studytime	failures	schoolsup	famsup	paid	activities	nursery	higher	internet	romantic	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	course	mother	2	2	0	yes	no	no	no	yes	yes	no	no	4	3	4	1	1	3	6	5	6	6
1	GP	F	17	U	GT3	T	1	1	at_home	other	course	father	1	2	0	no	yes	no	no	no	yes	yes	no	5	3	3	1	1	3	4	5	5	6
2	GP	F	15	U	LE3	T	1	1	at_home	other	other	mother	1	2	3	yes	no	yes	no	yes	yes	no	4	3	2	2	3	3	10	7	8	10	
3	GP	F	15	U	GT3	T	4	2	health	services	home	mother	1	3	0	no	yes	yes	yes	yes	yes	yes	3	2	2	1	1	5	2	15	14	15	
4	GP	F	16	U	GT3	T	3	3	other	other	home	father	1	2	0	no	yes	yes	no	yes	no	no	4	3	2	1	2	5	4	6	10	10	

Dimensões (linhas, colunas): (395, 33)

Valores nulos por coluna:

school	0
sex	0
age	0
address	0
famsize	0
Pstatus	0
Medu	0
Fedu	0
Mjob	0
Fjob	0
reason	0
guardian	0
traveltime	0
studytime	0
failures	0
schoolsup	0
famsup	0
paid	0
activities	0
nursery	0
higher	0
internet	0
romantic	0
famrel	0
freetime	0
goout	0
Dalc	0
Walc	0
health	0
absences	0
G1	0
G2	0
G3	0

==== Energy Efficiency Data =====

Primeiras linhas:

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2
0	0.98	514.5	294.0	110.25	7.0	2	0.0	0	15.55	21.33
1	0.98	514.5	294.0	110.25	7.0	3	0.0	0	15.55	21.33
2	0.98	514.5	294.0	110.25	7.0	4	0.0	0	15.55	21.33
3	0.98	514.5	294.0	110.25	7.0	5	0.0	0	15.55	21.33
4	0.90	563.5	318.5	122.50	7.0	2	0.0	0	20.84	28.28

Dimensões (linhas, colunas): (768, 10)

Valores nulos por coluna:

X1	0
X2	0
X3	0
X4	0
X5	0
X6	0
X7	0
X8	0
Y1	0
Y2	0

==== Pré-processamento - Car Evaluation Data =====

Linhas antes: 1728
Linhas removidas (com valores nulos): 0
Linhas após pré-processamento: 1728

==== Pré-processamento - Student Math Data =====

Linhas antes: 395
Linhas removidas (com valores nulos): 0
Linhas após pré-processamento: 395

==== Pré-processamento - Energy Efficiency Data =====

Linhas antes: 768
Linhas removidas (com valores nulos): 0
Linhas após pré-processamento: 768

Divisão Treino/Teste - Car Evaluation:

X_train: (1209, 6), y_train: (1209,)
X_test : (519, 6), y_test : (519,)

Divisão Treino/Teste - Student Math:

X_train: (276, 30), y_train: (276,)
X_test : (119, 30), y_test : (119,)

Divisão Treino/Teste - Energy Efficiency:

X_train: (537, 8), y_train: (537,)
X_test : (231, 8), y_test : (231,)

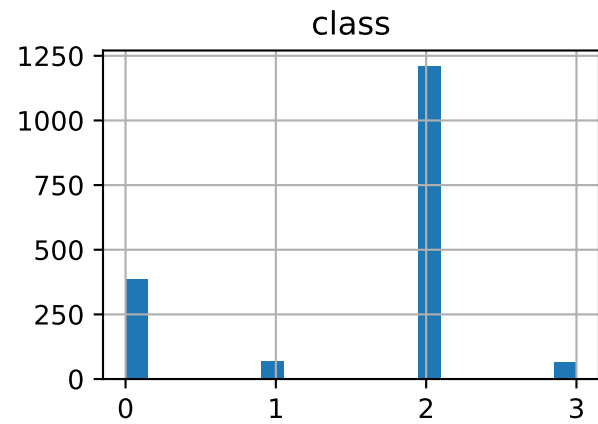
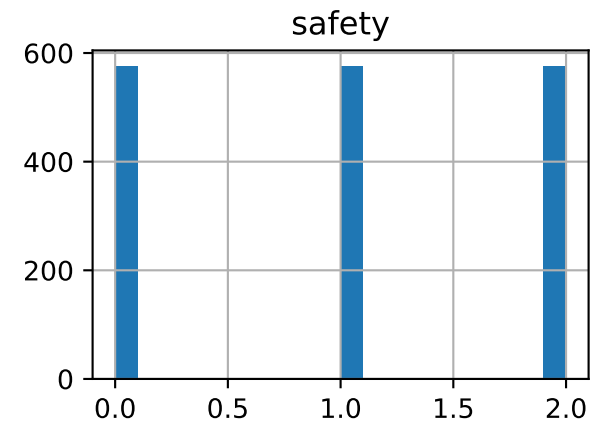
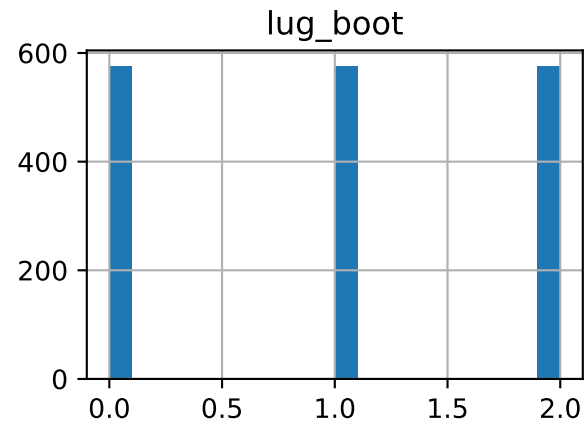
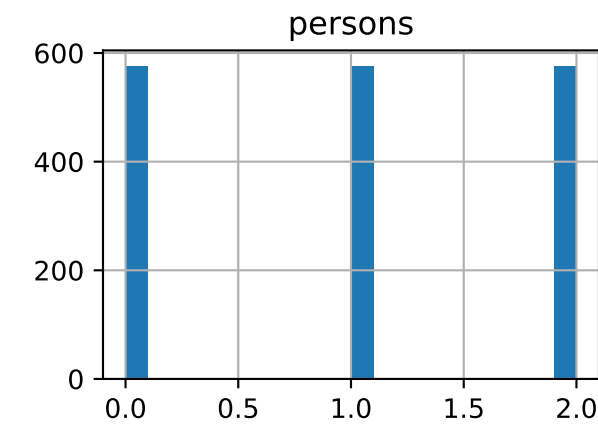
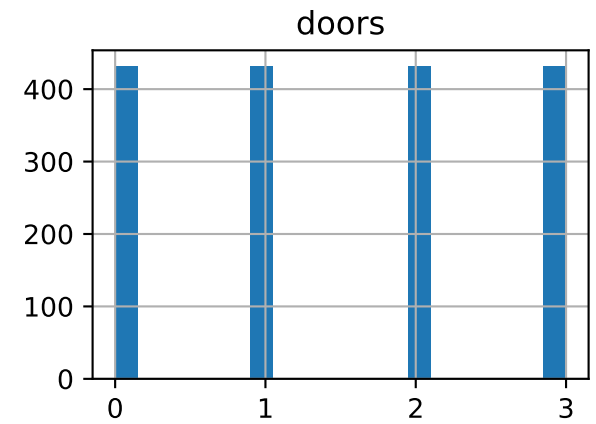
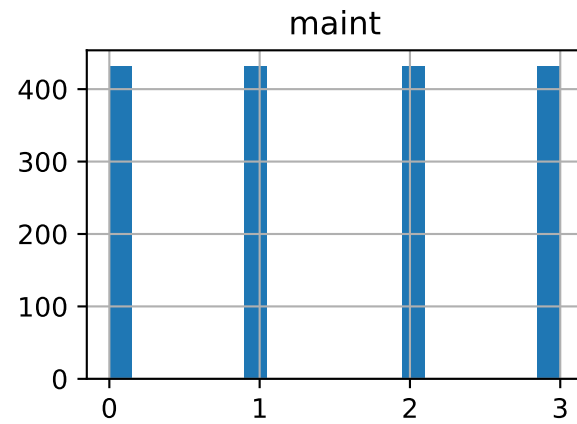
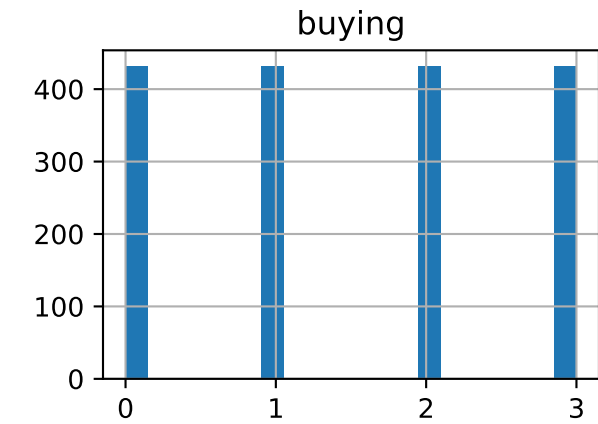
Análise Descritiva - Car Evaluation Data

==== Car Evaluation Data ====

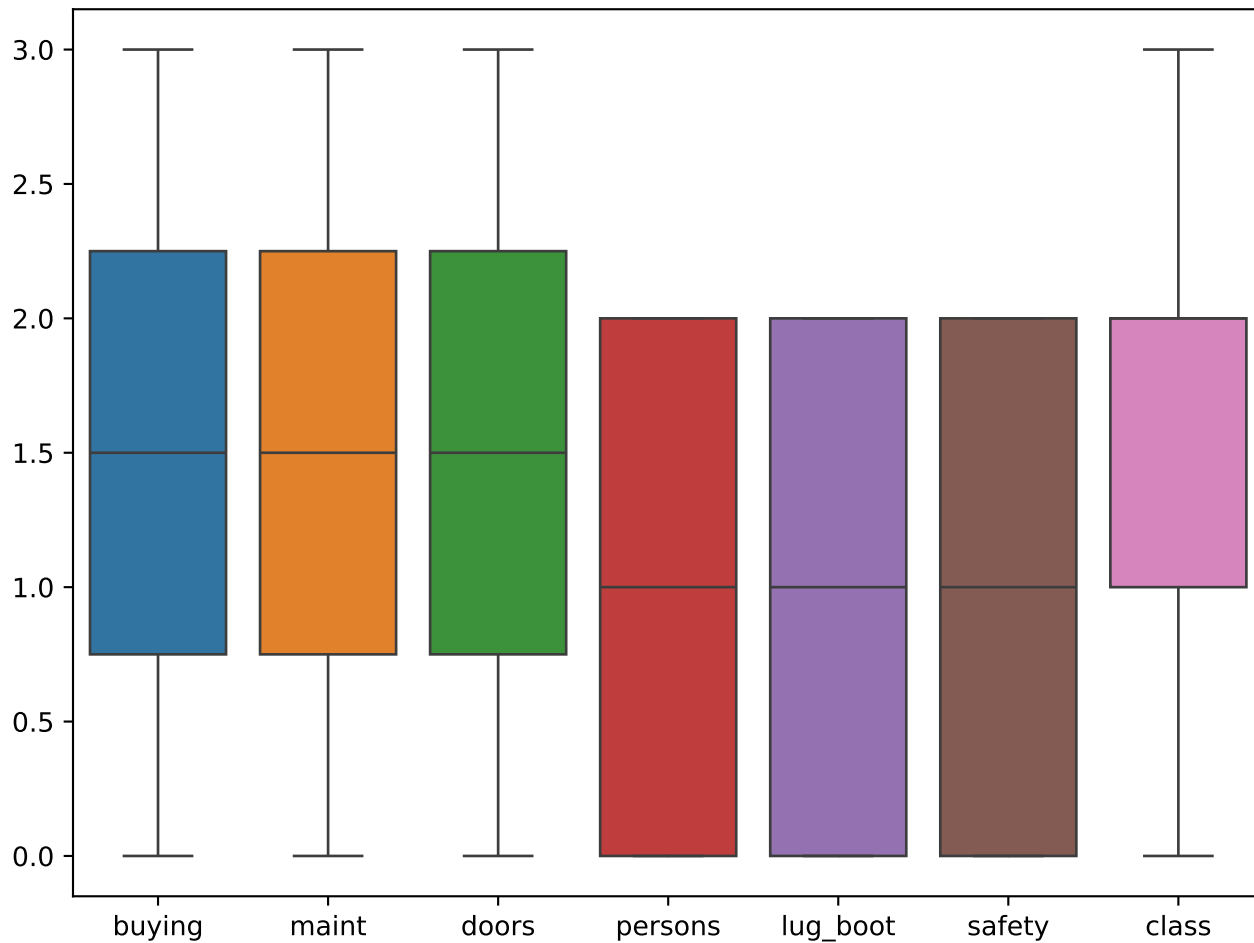
Estatísticas Descritivas:

	buying	maint	doors	persons	lug_boot	safety	class
count	1728.000000	1728.000000	1728.000000	1728.000000	1728.000000	1728.000000	1728.000000
mean	1.500000	1.500000	1.500000	1.000000	1.000000	1.000000	1.553241
std	1.118358	1.118358	1.118358	0.816733	0.816733	0.816733	0.875948
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.750000	0.750000	0.750000	0.000000	0.000000	0.000000	1.000000
50%	1.500000	1.500000	1.500000	1.000000	1.000000	1.000000	2.000000
75%	2.250000	2.250000	2.250000	2.000000	2.000000	2.000000	2.000000
max	3.000000	3.000000	3.000000	2.000000	2.000000	2.000000	3.000000

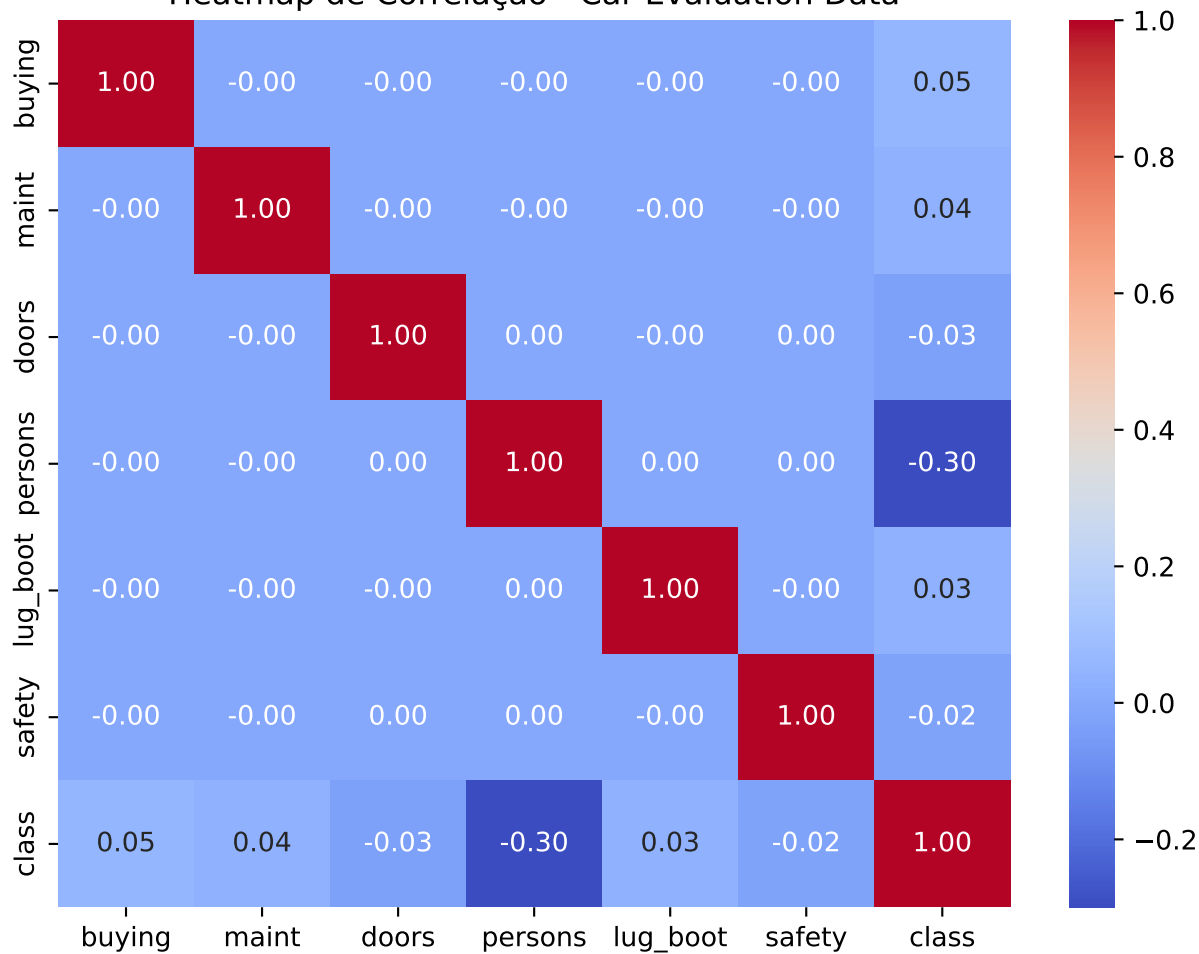
Histogramas - Car Evaluation Data



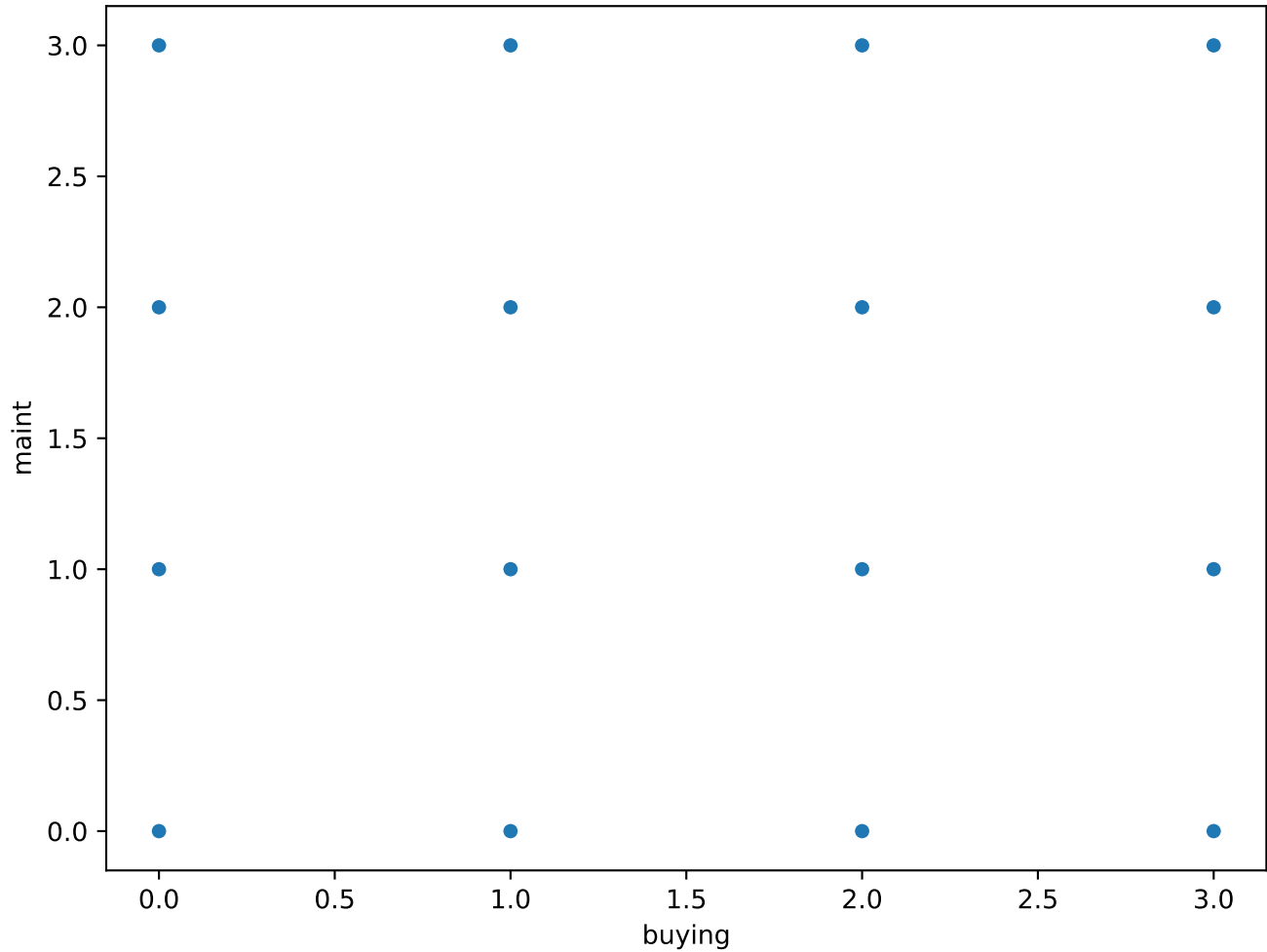
Boxplot - Car Evaluation Data



Heatmap de Correlação - Car Evaluation Data



Scatter Plot - Car Evaluation Data (buying vs maint)



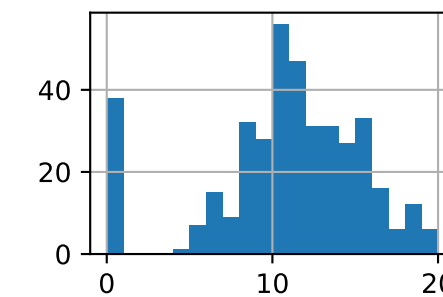
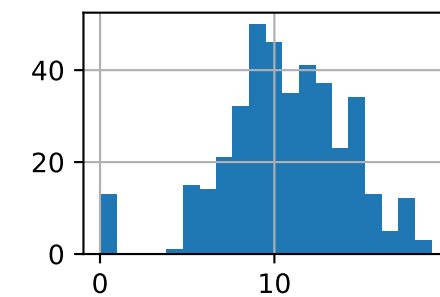
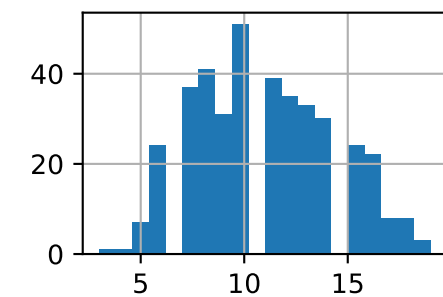
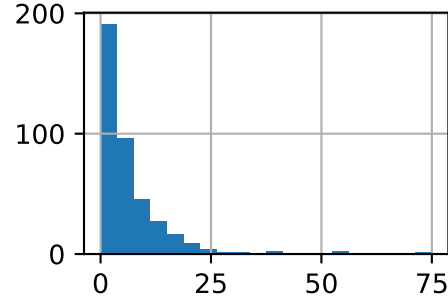
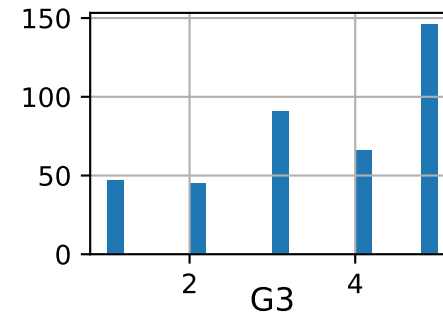
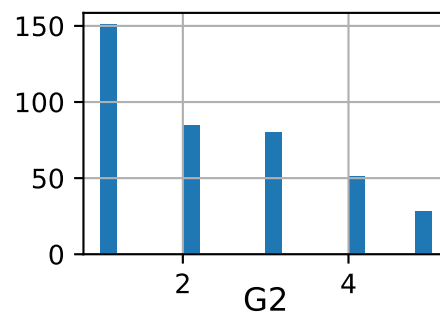
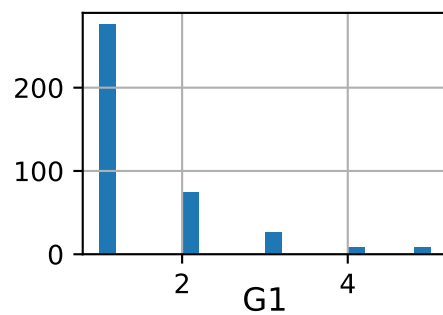
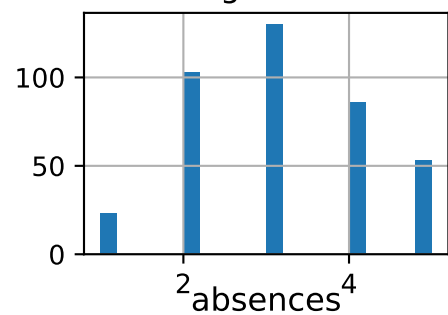
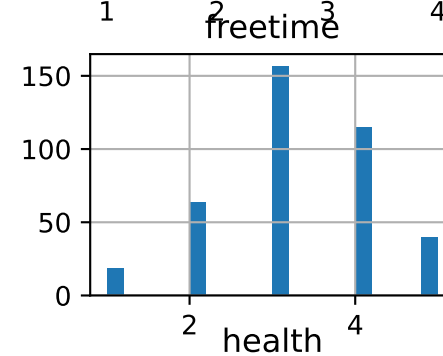
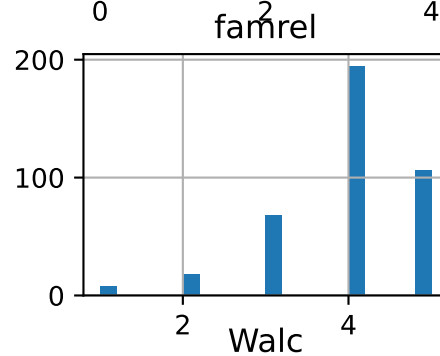
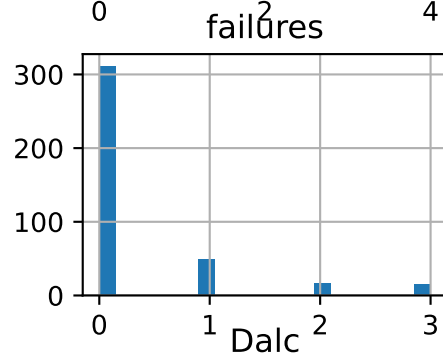
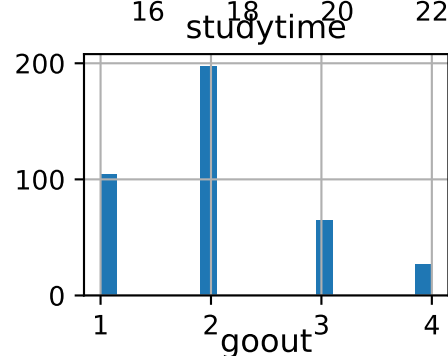
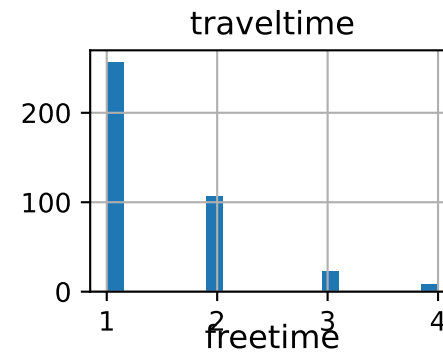
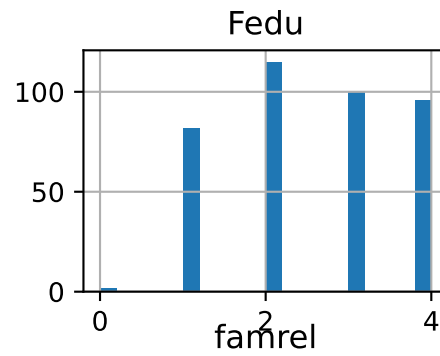
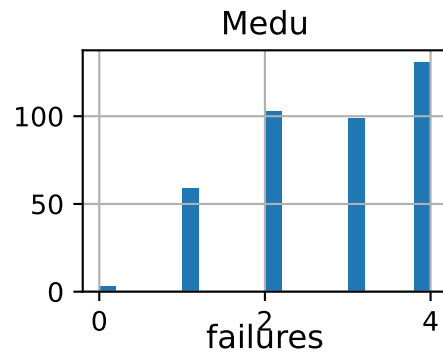
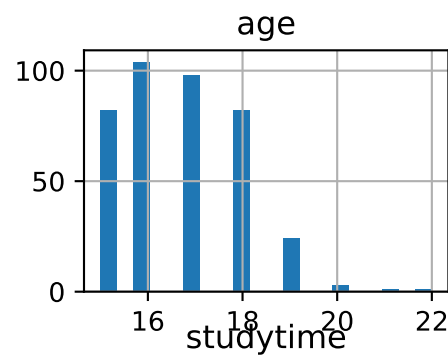
Análise Descritiva - Student Math Data

===== Student Math Data =====

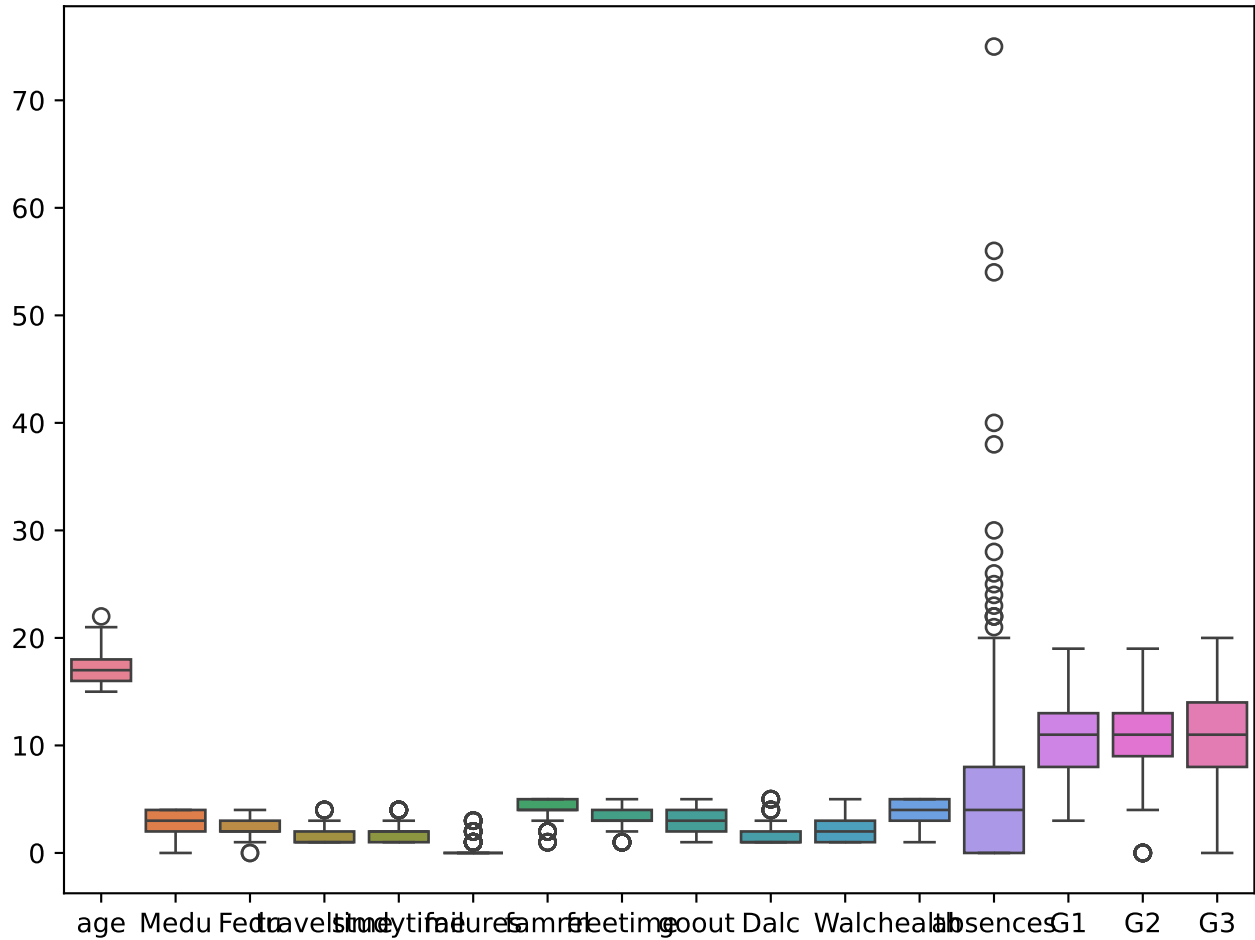
Estatísticas Descritivas:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	reason	guardian	traveltime	studytime	failures	schoolsup	famsup	paid	activities	nursery	higher	internet	romantic	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
count	395	395	395.000000	395	395	395	395.000000	395.000000	395	395	395	395	395.000000	395.000000	395.000000	395	395	395	395	395	395	395	395	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000
unique	2	2	NaN	2	2	2	NaN	NaN	5	5	4	3	NaN	NaN	NaN	2	2	2	2	2	2	2	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
top	GP	F	NaN	U	GT3	T	NaN	NaN	other	other	course	mother	NaN	NaN	NaN	no	yes	no	yes	yes	yes	yes	no	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
freq	349	208	NaN	307	281	354	NaN	NaN	141	217	145	273	NaN	NaN	NaN	344	242	214	201	314	375	329	263	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
mean	NaN	NaN	16.696203	NaN	NaN	NaN	2.749367	2.521519	NaN	NaN	NaN	NaN	1.448101	2.035443	0.334177	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	3.944304	3.235443	3.108861	1.481013	2.291139	3.554430	5.708861	10.908861	10.713924	10.415190
std	NaN	NaN	1.276043	NaN	NaN	NaN	1.094735	1.088201	NaN	NaN	NaN	NaN	0.697505	0.839240	0.743651	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.896659	0.998862	1.113278	0.890741	1.287897	1.390303	8.003096	3.319195	3.761505	4.581443
min	NaN	NaN	15.000000	NaN	NaN	NaN	0.000000	0.000000	NaN	NaN	NaN	NaN	1.000000	1.000000	0.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	3.000000	0.000000	0.000000
25%	NaN	NaN	16.000000	NaN	NaN	NaN	2.000000	2.000000	NaN	NaN	NaN	NaN	1.000000	1.000000	0.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	4.000000	3.000000	2.000000	1.000000	1.000000	3.000000	0.000000	8.000000	9.000000	8.000000
50%	NaN	NaN	17.000000	NaN	NaN	NaN	3.000000	2.000000	NaN	NaN	NaN	NaN	1.000000	2.000000	0.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	4.000000	3.000000	3.000000	1.000000	2.000000	4.000000	4.000000	11.000000	11.000000	11.000000
75%	NaN	NaN	18.000000	NaN	NaN	NaN	4.000000	3.000000	NaN	NaN	NaN	NaN	2.000000	2.000000	0.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5.000000	4.000000	4.000000	2.000000	3.000000	5.000000	8.000000	13.000000	13.000000	14.000000
max	NaN	NaN	22.000000	NaN	NaN	NaN	4.000000	4.000000	NaN	NaN	NaN	NaN	4.000000	4.000000	3.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	75.000000	19.000000	19.000000	20.000000

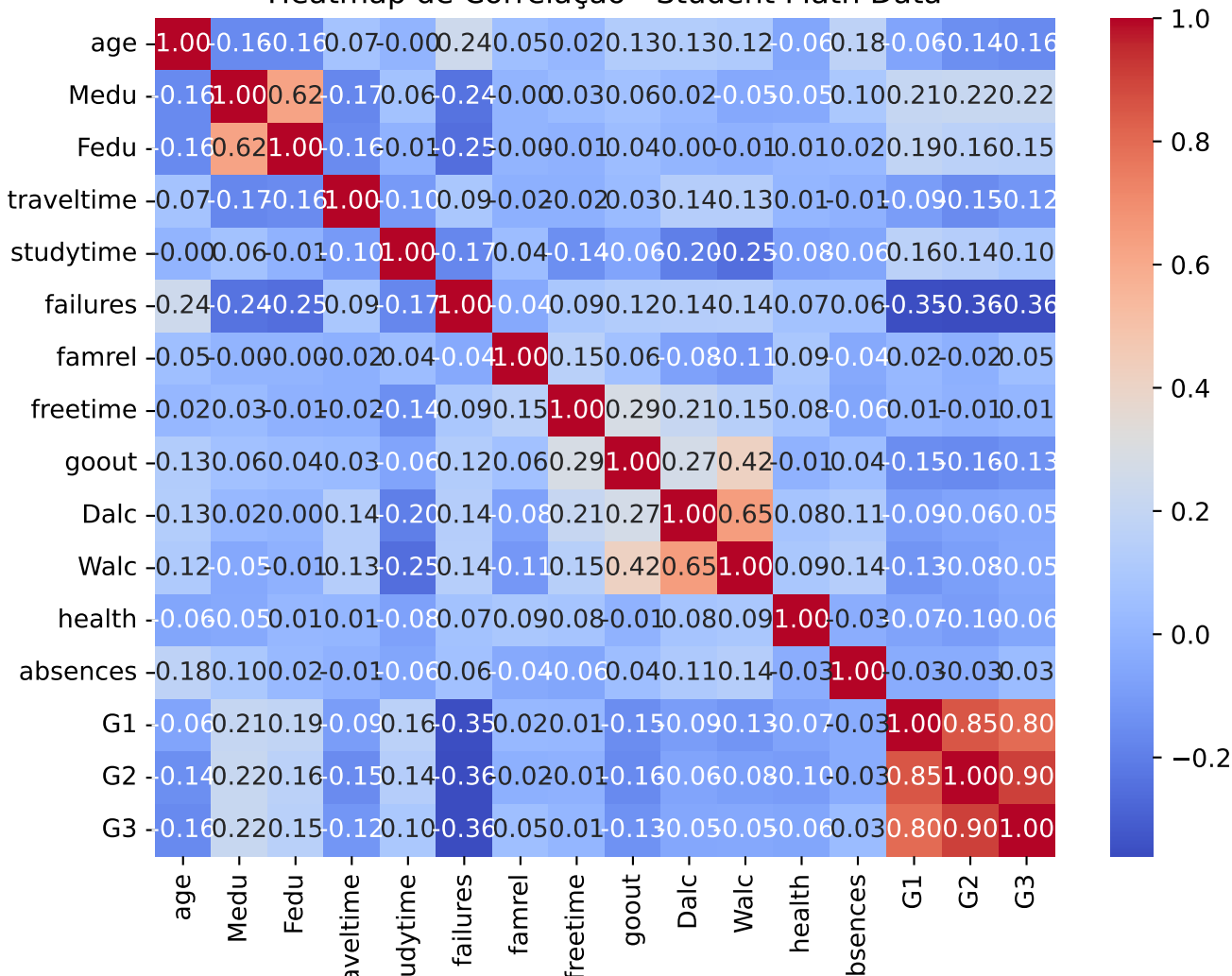
Histogramas - Student Math Data



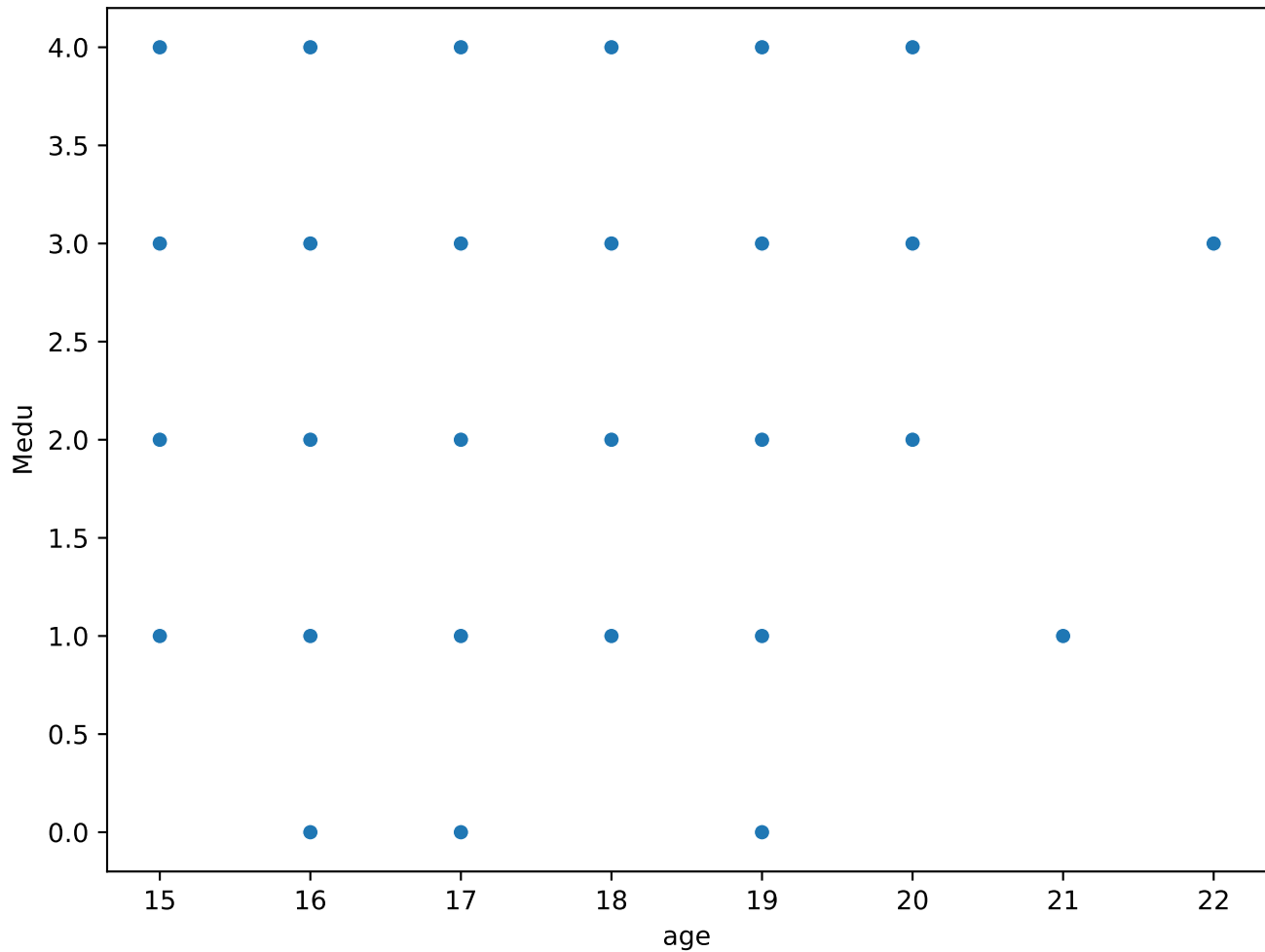
Boxplot - Student Math Data



Heatmap de Correlação - Student Math Data



Scatter Plot - Student Math Data (age vs Medu)

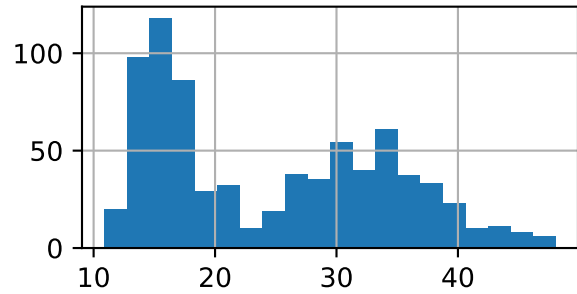
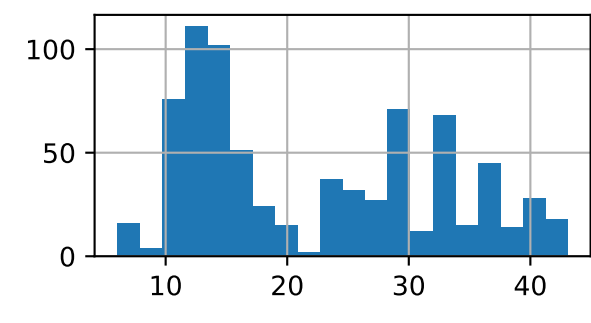
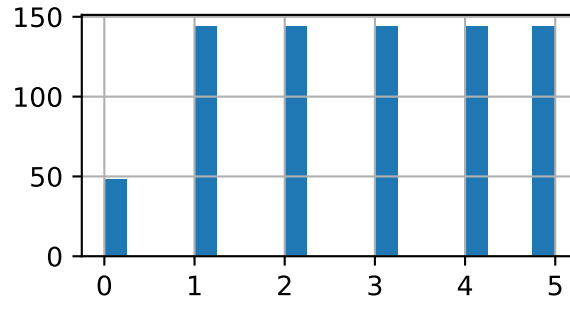
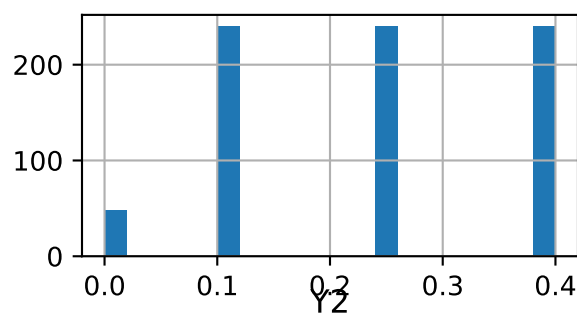
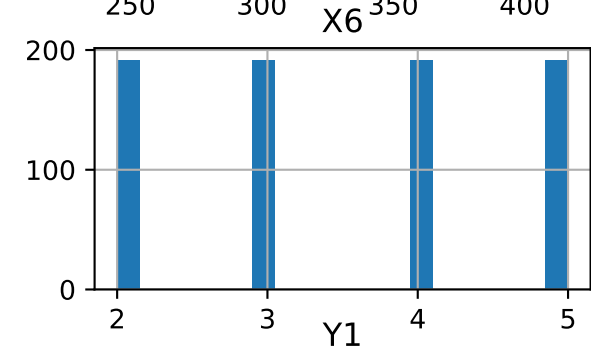
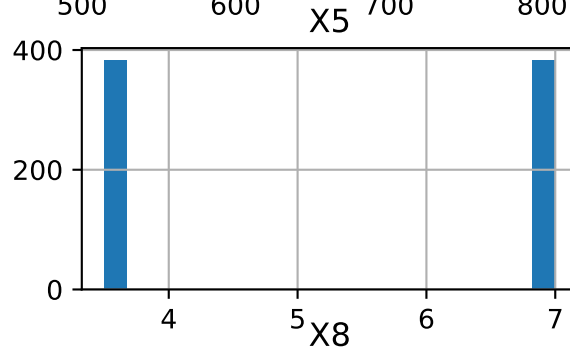
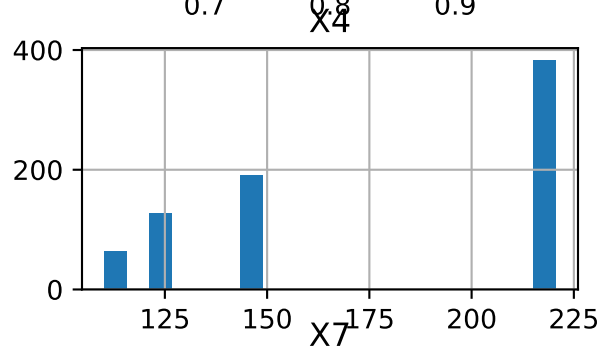
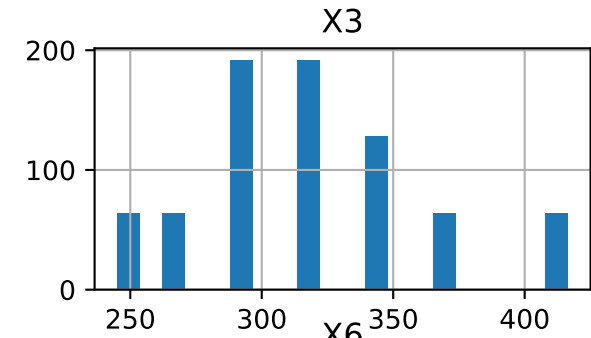
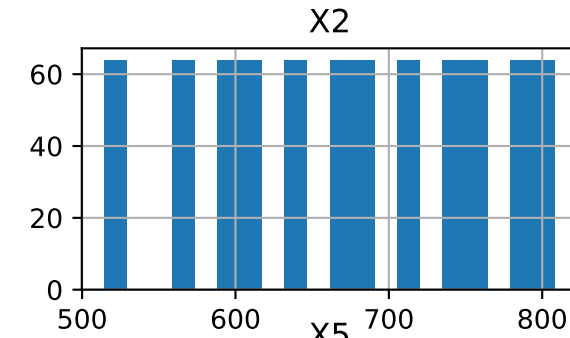
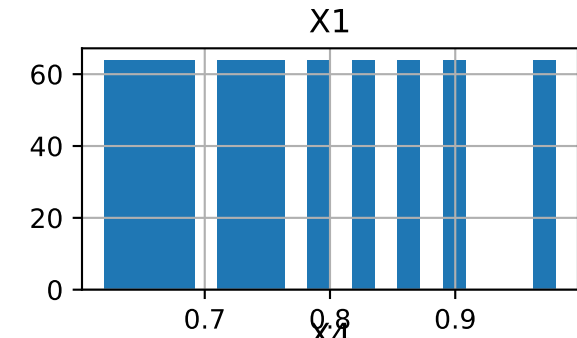


Análise Descritiva - Energy Efficiency Data

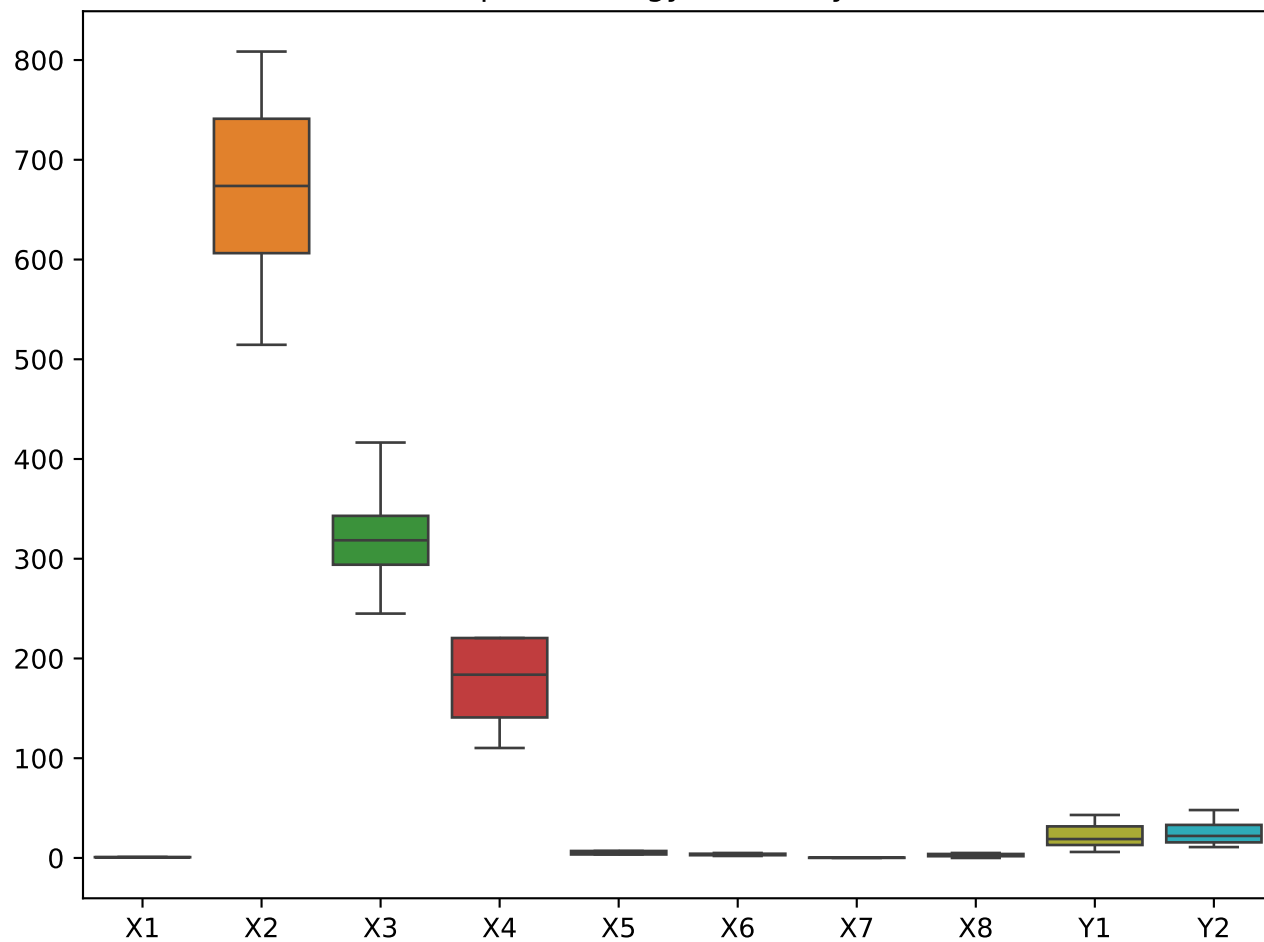
===== Energy Efficiency Data =====

Estatísticas Descritivas:											
	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	class
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768
unique	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	3
top	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	low
freq	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	377
mean	0.764167	671.708333	318.500000	176.604167	5.25000	3.500000	0.234375	2.81250	22.307195	24.587760	NaN
std	0.105777	88.086116	43.626481	45.165950	1.75114	1.118763	0.133221	1.55096	10.090204	9.513306	NaN
min	0.620000	514.500000	245.000000	110.250000	3.50000	2.000000	0.000000	0.00000	6.010000	10.900000	NaN
25%	0.682500	606.375000	294.000000	140.875000	3.50000	2.750000	0.100000	1.75000	12.992500	15.620000	NaN
50%	0.750000	673.750000	318.500000	183.750000	5.25000	3.500000	0.250000	3.00000	18.950000	22.080000	NaN
75%	0.830000	741.125000	343.000000	220.500000	7.00000	4.250000	0.400000	4.00000	31.667500	33.132500	NaN
max	0.980000	808.500000	416.500000	220.500000	7.00000	5.000000	0.400000	5.00000	43.100000	48.030000	NaN

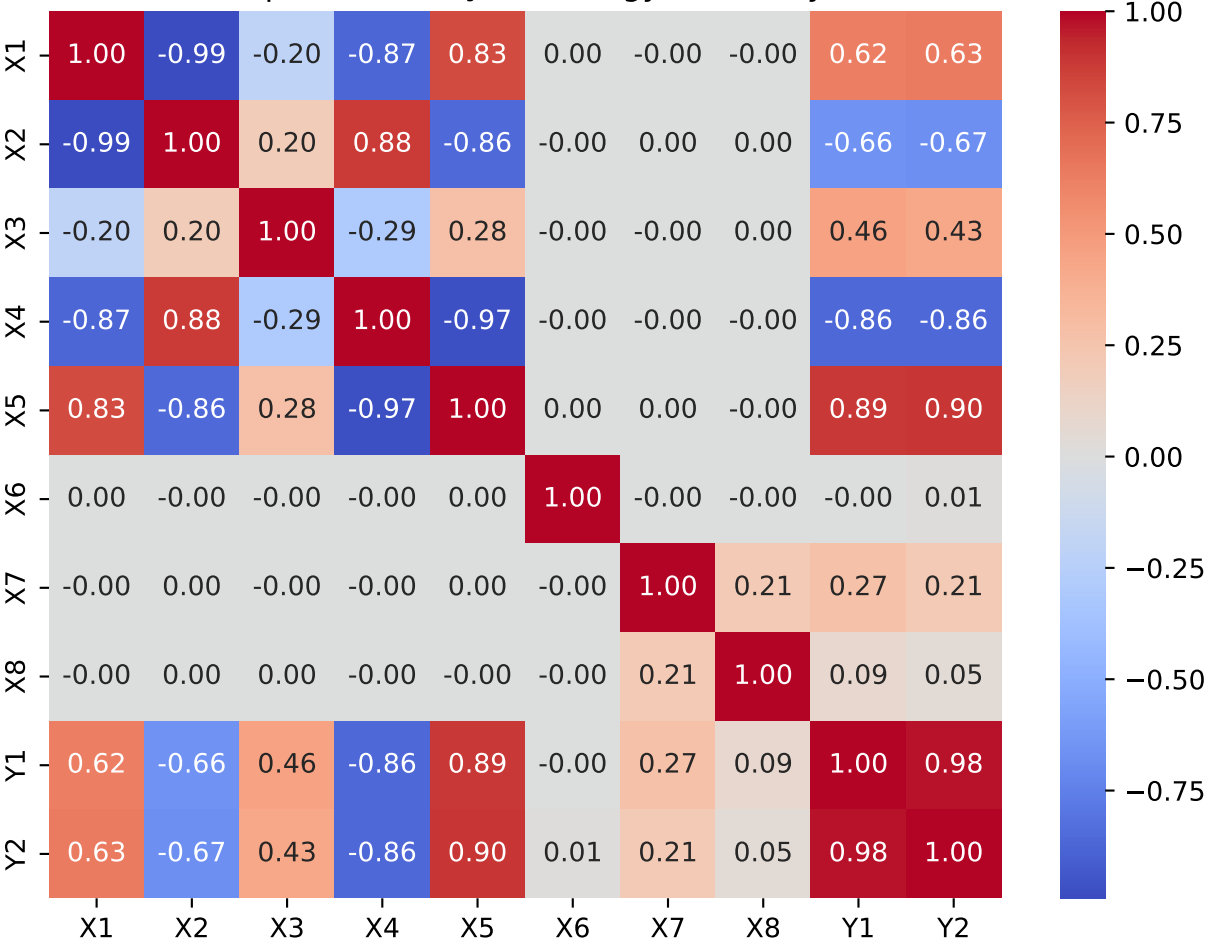
Histogramas - Energy Efficiency Data



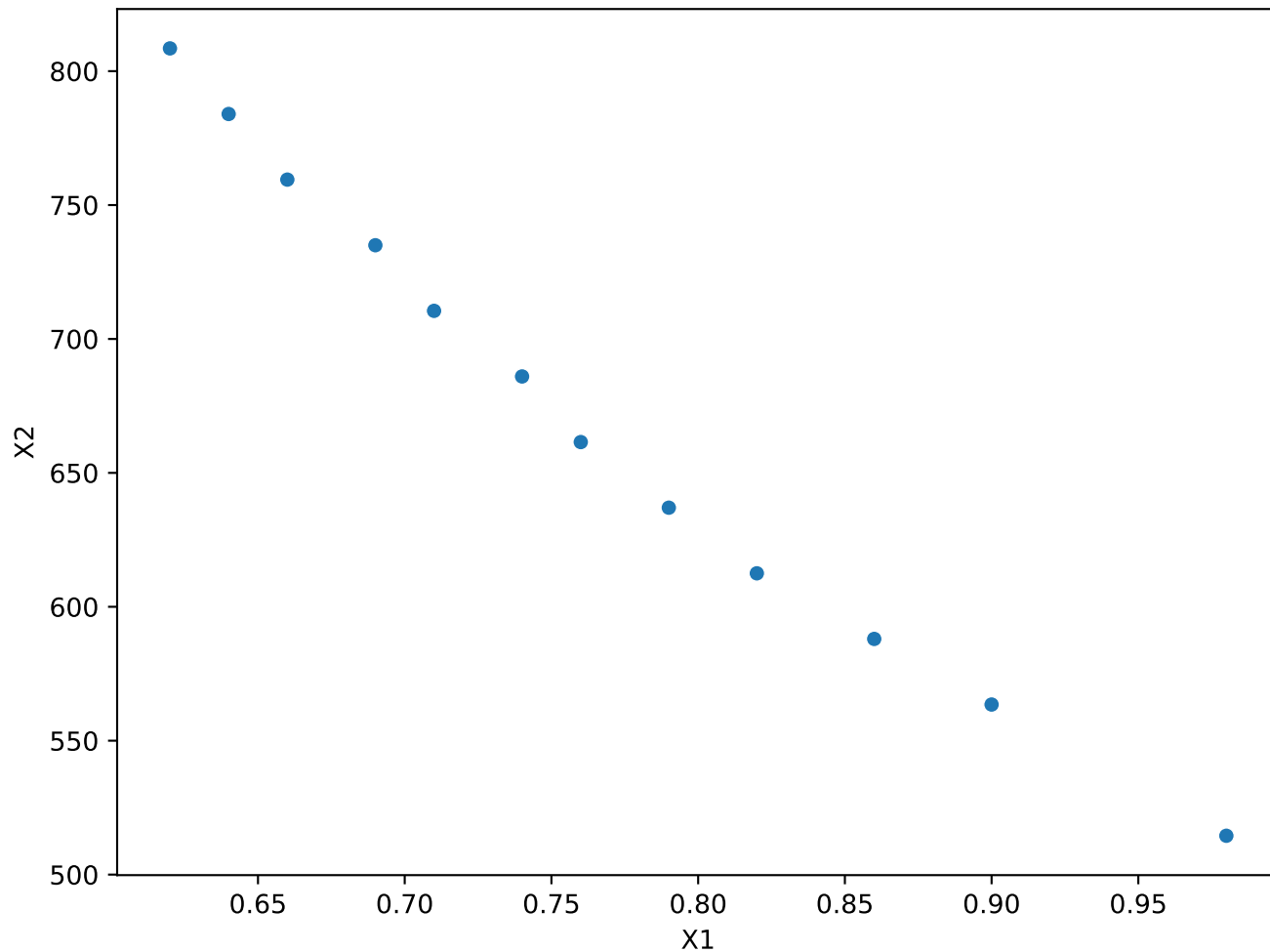
Boxplot - Energy Efficiency Data



Heatmap de Correlação - Energy Efficiency Data



Scatter Plot - Energy Efficiency Data (X1 vs X2)



Resultados - Car Evaluation

==== Avaliação - Car Evaluation ====

Random Forest:
Precisão: 0.9697
CV (média): 0.9734

	precision	recall	f1-score	support
0	0.93	0.92	0.93	118
1	0.71	0.89	0.79	19
2	1.00	0.99	0.99	358
3	0.95	0.88	0.91	24
accuracy			0.97	519
macro avg	0.90	0.92	0.91	519
weighted avg	0.97	0.97	0.97	519

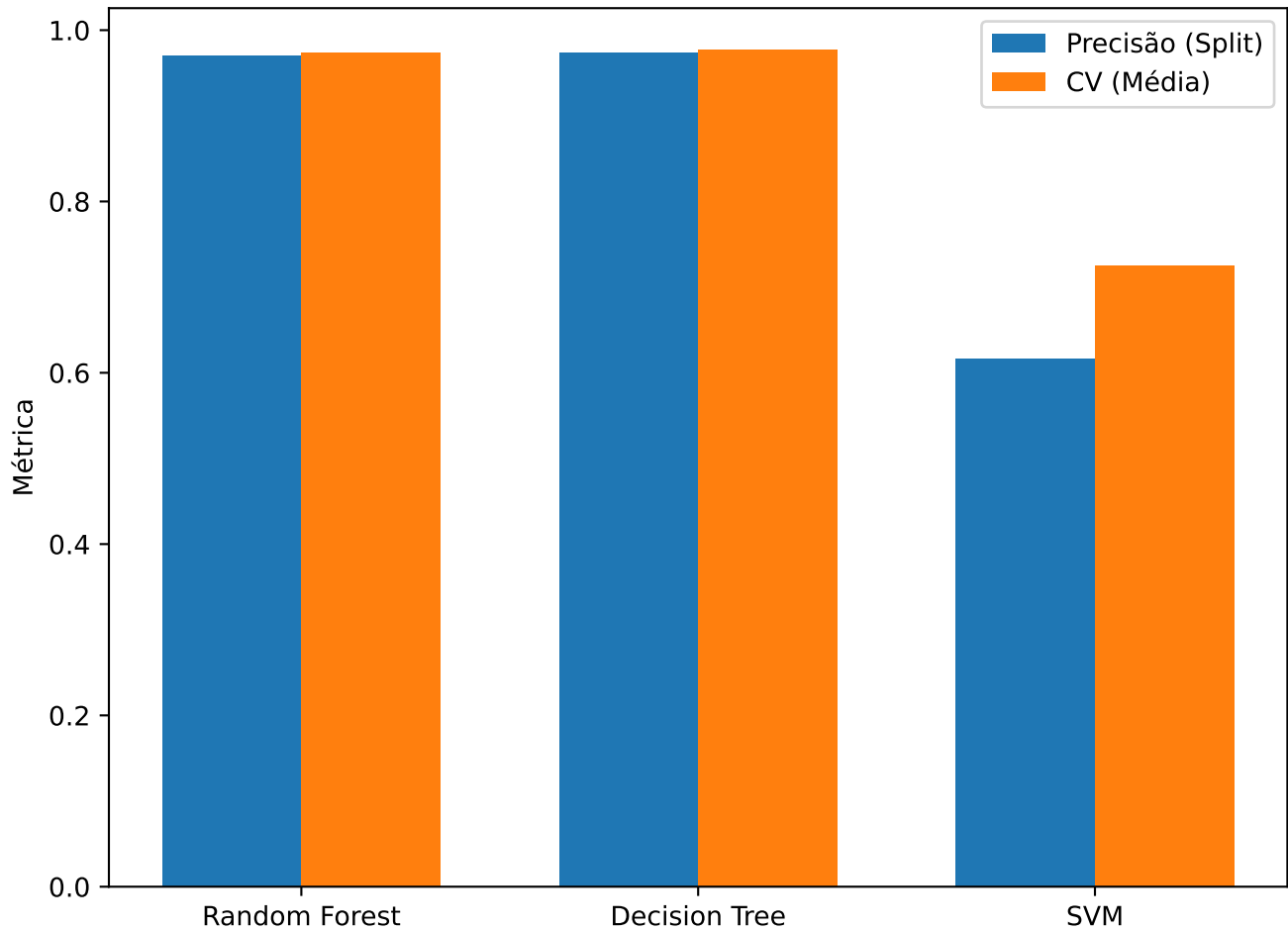
Decision Tree:
Precisão: 0.9736
CV (média): 0.9768

	precision	recall	f1-score	support
0	0.93	0.94	0.94	118
1	0.75	0.95	0.84	19
2	1.00	0.99	0.99	358
3	1.00	0.88	0.93	24
accuracy			0.97	519
macro avg	0.92	0.94	0.93	519
weighted avg	0.97	0.97	0.97	519

SVM:
Precisão: 0.6163
CV (média): 0.7251

	precision	recall	f1-score	support
0	0.52	0.14	0.21	118
1	0.00	0.00	0.00	19
2	0.72	0.99	0.83	358
3	0.00	0.00	0.00	24
accuracy			0.71	519
macro avg	0.31	0.28	0.26	519
weighted avg	0.62	0.71	0.62	519

Comparativo de Resultados - Car Evaluation



Resultados - Student Math

==== Avaliação - Student Math ====

Random Forest:

Precisão: 0.6541

CV (média): 0.6304

	precision	recall	f1-score	support
high	0.69	0.32	0.44	34
low	0.80	0.17	0.29	23
medium	0.58	0.92	0.71	62
accuracy			0.61	119
macro avg	0.69	0.47	0.48	119
weighted avg	0.65	0.61	0.55	119

Decision Tree:

Precisão: 0.5222

CV (média): 0.5367

	precision	recall	f1-score	support
high	0.43	0.35	0.39	34
low	0.42	0.35	0.38	23
medium	0.61	0.71	0.66	62
accuracy			0.54	119
macro avg	0.49	0.47	0.47	119
weighted avg	0.52	0.54	0.53	119

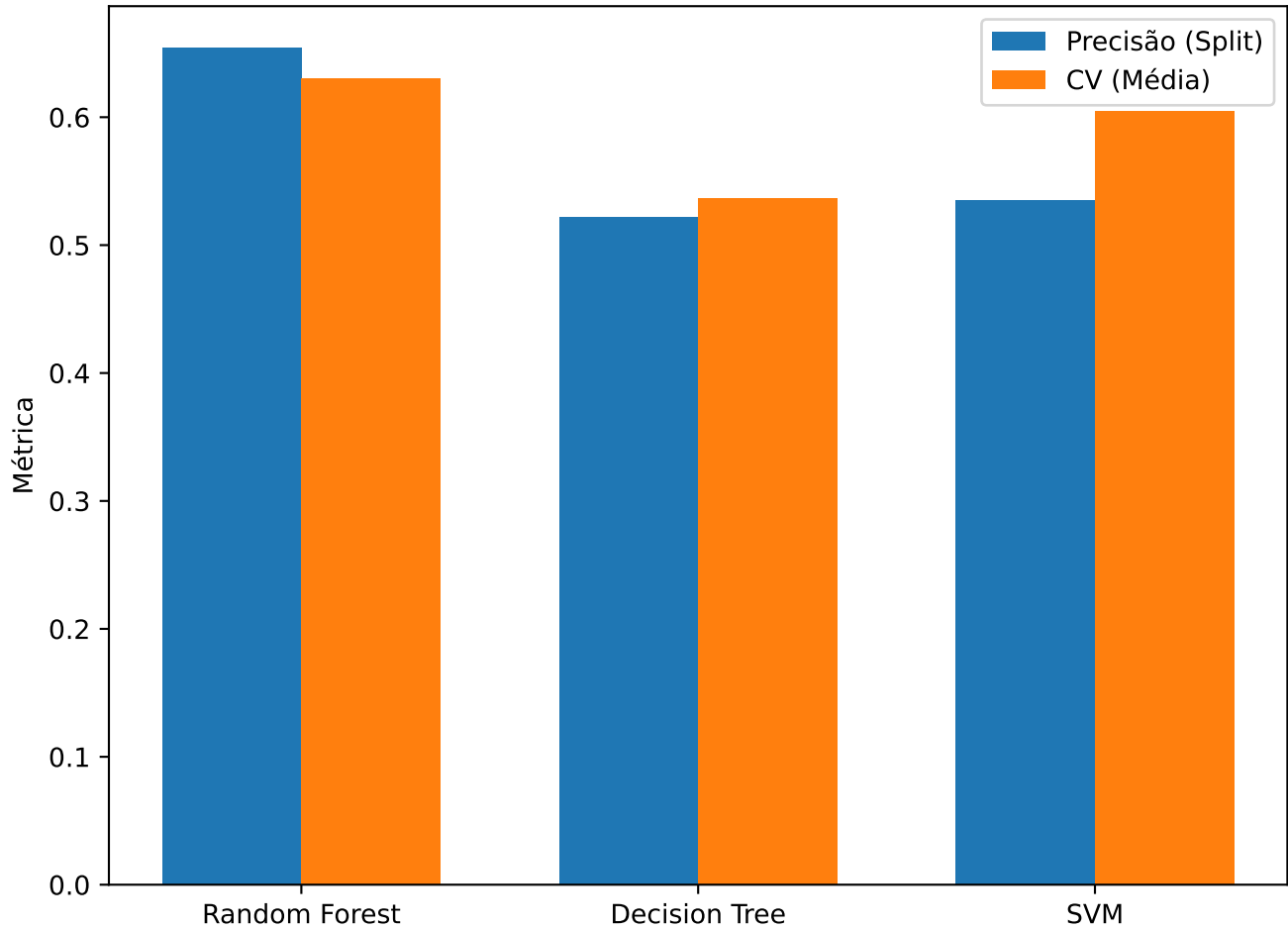
SVM:

Precisão: 0.5347

CV (média): 0.6051

	precision	recall	f1-score	support
high	0.56	0.44	0.49	34
low	0.40	0.17	0.24	23
medium	0.57	0.76	0.65	62
accuracy			0.55	119
macro avg	0.51	0.46	0.46	119
weighted avg	0.53	0.55	0.53	119

Comparativo de Resultados - Student Math



Resultados - Energy Efficiency

==== Avaliação - Energy Efficiency ====

Random Forest:

Precisão: 1.0000

CV (média): 0.9987

	precision	recall	f1-score	support
high	1.00	1.00	1.00	64
low	1.00	1.00	1.00	103
medium	1.00	1.00	1.00	64
accuracy			1.00	231
macro avg	1.00	1.00	1.00	231
weighted avg	1.00	1.00	1.00	231

Decision Tree:

Precisão: 0.9957

CV (média): 0.9974

	precision	recall	f1-score	support
high	1.00	1.00	1.00	64
low	0.99	1.00	1.00	103
medium	1.00	0.98	0.99	64
accuracy			1.00	231
macro avg	1.00	0.99	1.00	231
weighted avg	1.00	1.00	1.00	231

SVM:

Precisão: 0.9075

CV (média): 0.9062

	precision	recall	f1-score	support
high	0.90	0.88	0.89	64
low	0.94	0.99	0.97	103
medium	0.85	0.81	0.83	64
accuracy			0.91	231
macro avg	0.90	0.89	0.90	231
weighted avg	0.91	0.91	0.91	231

Comparativo de Resultados - Energy Efficiency

