

Data Table

Sat Nov 23 24, 01:09:43

Data instances: 614
Features: 11
Meta attributes: None

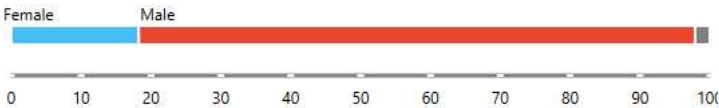
	Gender	Married	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
1	Male	No	Graduate	No	5849	0	?	360	1	Urban	Y
2	Male	Yes	Graduate	No	4583	1508	128	360	1	Rural	N
3	Male	Yes	Graduate	Yes	3000	0	66	360	1	Urban	Y
4	Male	Yes	Not Graduate	No	2583	2358	120	360	1	Urban	Y
5	Male	No	Graduate	No	6000	0	141	360	1	Urban	Y
6	Male	Yes	Graduate	Yes	5417	4196	267	360	1	Urban	Y
7	Male	Yes	Not Graduate	No	2333	1516	95	360	1	Urban	Y
8	Male	Yes	Graduate	No	3036	2504	158	360	0	Semiurban	N
9	Male	Yes	Graduate	No	4006	1526	168	360	1	Urban	Y
10	Male	Yes	Graduate	No	12841	10968	349	360	1	Semiurban	N
11	Male	Yes	Graduate	No	3200	700	70	360	1	Urban	Y
12	Male	Yes	Graduate	?	2500	1840	109	360	1	Urban	Y
13	Male	Yes	Graduate	No	3073	8106	200	360	1	Urban	Y
14	Male	No	Graduate	No	1853	2840	114	360	1	Rural	N
15	Male	Yes	Graduate	No	1299	1086	17	120	1	Urban	Y
16	Male	No	Graduate	No	4950	0	125	360	1	Urban	Y
17	Male	No	Not Graduate	No	3596	0	100	240	?	Urban	Y
18	Female	No	Graduate	No	3510	0	76	360	0	Urban	N
19	Male	Yes	Not Graduate	No	4887	0	133	360	1	Rural	N
20	Male	Yes	Graduate	?	2600	3500	115	?	1	Urban	Y
21	Male	Yes	Not Graduate	No	7660	0	104	360	0	Urban	N
22	Male	Yes	Graduate	No	5955	5625	315	360	1	Urban	Y
23	Male	Yes	Not Graduate	No	2600	1911	116	360	0	Semiurban	N
24	?	Yes	Not Graduate	No	3365	1917	112	360	0	Rural	N
25	Male	Yes	Graduate	?	3717	2925	151	360	?	Semiurban	N
26	Male	Yes	Graduate	Yes	9560	0	191	360	1	Semiurban	Y
27	Male	Yes	Graduate	No	2799	2253	122	360	1	Semiurban	Y
28	Male	Yes	Not Graduate	No	4226	1040	110	360	1	Urban	Y
29	Male	No	Not Graduate	No	1442	0	35	360	1	Urban	N
30	Female	No	Graduate	?	3750	2083	120	360	1	Semiurban	Y
31	Male	Yes	Graduate	?	4166	3369	201	360	?	Urban	N
32	Male	No	Graduate	No	3167	0	74	360	1	Urban	N
33	Male	No	Graduate	Yes	4692	0	106	360	1	Rural	N
34	Male	Yes	Graduate	No	3500	1667	114	360	1	Semiurban	Y
35	Male	No	Graduate	No	12500	3000	320	360	1	Rural	N
36	Male	Yes	Graduate	No	2275	2067	?	360	1	Urban	Y
37	Male	Yes	Graduate	No	1828	1330	100	?	0	Urban	N
38	Female	Yes	Graduate	No	3667	1459	144	360	1	Semiurban	Y
39	Male	No	Graduate	No	4166	7210	184	360	1	Urban	Y
40	Male	No	Not Graduate	No	3748	1668	110	360	1	Semiurban	Y
41	Male	No	Graduate	No	3600	0	80	360	1	Urban	N
42	Male	No	Graduate	No	1800	1213	47	360	1	Urban	Y
43	Male	Yes	Graduate	No	2400	0	75	360	?	Urban	Y
44	Male	Yes	Graduate	No	3941	2336	134	360	1	Semiurban	Y
45	Male	Yes	Not Graduate	Yes	4695	0	96	?	1	Urban	Y
46	Female	No	Graduate	No	3410	0	88	?	1	Urban	Y
47	Male	Yes	Graduate	No	5649	0	44	360	1	Urban	Y
48	Male	Yes	Graduate	No	5821	0	144	360	1	Urban	Y
49	Female	Yes	Graduate	No	2645	3440	120	360	0	Urban	N
50	Female	No	Graduate	No	4000	2275	144	360	1	Semiurban	Y
51	Female	Yes	Not Graduate	No	1928	1644	100	360	1	Semiurban	Y
52	Female	No	Graduate	No	3086	0	120	360	1	Semiurban	Y
53	Female	No	Graduate	No	4230	0	112	360	1	Semiurban	N
54	Male	Yes	Graduate	No	4616	0	134	360	1	Urban	N
55	Female	Yes	Graduate	Yes	11500	0	286	360	0	Urban	N
56	Male	Yes	Graduate	No	2708	1167	97	360	1	Semiurban	Y
57	Male	Yes	Graduate	No	2132	1591	96	360	1	Semiurban	Y
58	Male	Yes	Graduate	No	3366	2200	135	360	1	Rural	N

59	Male	Yes	Graduate	No	8080	2250	180	360	1	Urban	Y
60	Male	Yes	Not Graduate	No	3357	2859	144	360	1	Urban	Y
61	Male	Yes	Graduate	No	2500	3796	120	360	1	Urban	Y
62	Male	Yes	Graduate	No	3029	0	99	360	1	Urban	Y
63	Male	Yes	Not Graduate	Yes	2609	3449	165	180	0	Rural	N
64	Male	Yes	Graduate	No	4945	0	?	360	0	Rural	N
65	Female	No	Graduate	No	4166	0	116	360	0	Semiurban	N
66	Male	Yes	Graduate	No	5726	4595	258	360	1	Semiurban	N
67	Male	No	Not Graduate	No	3200	2254	126	180	0	Urban	N
68	Male	Yes	Graduate	No	10750	0	312	360	1	Urban	Y
69	Male	Yes	Not Graduate	Yes	7100	0	125	60	1	Urban	Y
70	Female	No	Graduate	No	4300	0	136	360	0	Semiurban	N
71	Male	Yes	Graduate	No	3208	3066	172	360	1	Urban	Y
72	Male	Yes	Not Graduate	Yes	1875	1875	97	360	1	Semiurban	Y
73	Male	No	Graduate	No	3500	0	81	300	1	Semiurban	Y
74	Male	Yes	Not Graduate	No	4755	0	95	?	0	Semiurban	N
75	Male	Yes	Graduate	Yes	5266	1774	187	360	1	Semiurban	Y
76	Male	No	Graduate	No	3750	0	113	480	1	Urban	N
77	Male	No	Graduate	No	3750	4750	176	360	1	Urban	N
78	Male	Yes	Graduate	Yes	1000	3022	110	360	1	Urban	N
79	Male	Yes	Graduate	No	3167	4000	180	300	0	Semiurban	N
80	Male	Yes	Not Graduate	Yes	3333	2166	130	360	?	Semiurban	Y
81	Female	No	Graduate	No	3846	0	111	360	1	Semiurban	Y
82	Male	Yes	Graduate	Yes	2395	0	?	360	1	Semiurban	Y
83	Female	Yes	Graduate	No	1378	1881	167	360	1	Urban	N
84	Male	Yes	Graduate	No	6000	2250	265	360	?	Semiurban	N
85	Male	Yes	Graduate	No	3988	0	50	240	1	Urban	Y
86	Male	No	Graduate	No	2366	2531	136	360	1	Semiurban	Y
87	Male	Yes	Not Graduate	No	3333	2000	99	360	?	Semiurban	Y
88	Male	Yes	Graduate	No	2500	2118	104	360	1	Semiurban	Y
89	Male	No	Graduate	No	8566	0	210	360	1	Urban	Y
90	Male	Yes	Graduate	No	5695	4167	175	360	1	Semiurban	Y
91	Male	Yes	Graduate	No	2958	2900	131	360	1	Semiurban	Y
92	Male	Yes	Graduate	No	6250	5654	188	180	1	Semiurban	Y
93	Male	Yes	Not Graduate	No	3273	1820	81	360	1	Urban	Y
94	Male	No	Graduate	No	4133	0	122	360	1	Semiurban	Y
95	Male	No	Not Graduate	No	3620	0	25	120	1	Semiurban	Y
96	Male	No	Graduate	?	6782	0	?	360	?	Urban	N
97	Female	Yes	Graduate	No	2484	2302	137	360	1	Semiurban	Y
98	Male	Yes	Graduate	No	1977	997	50	360	1	Semiurban	Y
99	Male	Yes	Not Graduate	No	4188	0	115	180	1	Semiurban	Y
100	Male	Yes	Graduate	No	1759	3541	131	360	1	Semiurban	Y
+ 514 more											

Ao importar o arquivo csv “train loan” o “data_table” mostrou a seguinte tabela.

Box Plot

Sat Nov 23 24, 01:11:07



Box plot for attribute 'Gender'

Esse é o resultado do "box plot" para a coluna "gender". Como é possível perceber uma pequena parte é desconhecida, justamente os dados perdidos que precisamos tratar

Settings

Impute Missing Values: Remove rows with missing values.

Uma opção seria remover todas as linhas que possuem dados perdidos...

Data instances: 490
Features: 11
Meta attributes: None
Target: Class 'Selected'

	Selected	Gender	Married	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property
1	No	Male	Yes	Graduate	No	4583	1508	128	360	1	Rural
2	No	Male	Yes	Graduate	Yes	3000	0	66	360	1	Urban
3	No	Male	Yes	Not Graduate	No	2583	2358	120	360	1	Urban
4	No	Male	No	Graduate	No	6000	0	141	360	1	Urban
5	No	Male	Yes	Graduate	Yes	5417	4196	267	360	1	Urban
6	No	Male	Yes	Not Graduate	No	2333	1516	95	360	1	Urban
7	No	Male	Yes	Graduate	No	3036	2504	158	360	0	Semiurb
8	No	Male	Yes	Graduate	No	4006	1526	168	360	1	Urban
9	No	Male	Yes	Graduate	No	12841	10968	349	360	1	Semiurb
10	No	Male	Yes	Graduate	No	3200	700	70	360	1	Urban
11	No	Male	Yes	Graduate	No	3073	8106	200	360	1	Urban
12	No	Male	No	Graduate	No	1853	2840	114	360	1	Rural
13	No	Male	Yes	Graduate	No	1299	1086	17	120	1	Urban
14	No	Male	No	Graduate	No	4950	0	125	360	1	Urban
15	No	Female	No	Graduate	No	3510	0	76	360	0	Urban
16	No	Male	Yes	Not Graduate	No	4887	0	133	360	1	Rural
17	No	Male	Yes	Not Graduate	No	7660	0	104	360	0	Urban
18	No	Male	Yes	Graduate	No	5955	5625	315	360	1	Urban
19	No	Male	Yes	Not Graduate	No	2600	1911	116	360	0	Semiurb
20	No	Male	Yes	Graduate	Yes	9560	0	191	360	1	Semiurb
21	No	Male	Yes	Graduate	No	2799	2253	122	360	1	Semiurb
22	No	Male	Yes	Not Graduate	No	4226	1040	110	360	1	Urban
23	No	Male	No	Not Graduate	No	1442	0	35	360	1	Urban
24	No	Male	No	Graduate	No	3167	0	74	360	1	Urban
25	No	Male	No	Graduate	Yes	4692	0	106	360	1	Rural
26	No	Male	Yes	Graduate	No	3500	1667	114	360	1	Semiurb
27	No	Male	No	Graduate	No	12500	3000	320	360	1	Rural
28	No	Female	Yes	Graduate	No	3667	1459	144	360	1	Semiurb
29	No	Male	No	Graduate	No	4166	7210	184	360	1	Urban
30	No	Male	No	Not Graduate	No	3748	1668	110	360	1	Semiurb
31	No	Male	No	Graduate	No	3600	0	80	360	1	Urban
32	No	Male	No	Graduate	No	1800	1213	47	360	1	Urban
33	No	Male	Yes	Graduate	No	3941	2336	134	360	1	Semiurb
34	No	Male	Yes	Graduate	No	5649	0	44	360	1	Urban
35	No	Male	Yes	Graduate	No	5821	0	144	360	1	Urban
36	No	Female	Yes	Graduate	No	2645	3440	120	360	0	Urban
37	No	Female	No	Graduate	No	4000	2275	144	360	1	Semiurb
38	No	Female	Yes	Not Graduate	No	1928	1644	100	360	1	Semiurb
39	No	Female	No	Graduate	No	3086	0	120	360	1	Semiurb
40	No	Female	No	Graduate	No	4230	0	112	360	1	Semiurb
41	No	Male	Yes	Graduate	No	4616	0	134	360	1	Urban
42	No	Female	Yes	Graduate	Yes	11500	0	286	360	0	Urban
43	No	Male	Yes	Graduate	No	2708	1167	97	360	1	Semiurb
44	No	Male	Yes	Graduate	No	2132	1591	96	360	1	Semiurb
45	No	Male	Yes	Graduate	No	3366	2200	135	360	1	Rural
46	No	Male	Yes	Graduate	No	8080	2250	180	360	1	Urban
47	No	Male	Yes	Not Graduate	No	3357	2859	144	360	1	Urban
48	No	Male	Yes	Graduate	No	2500	3796	120	360	1	Urban
49	No	Male	Yes	Graduate	No	3029	0	99	360	1	Urban
50	No	Male	Yes	Not Graduate	Yes	2609	3449	165	180	0	Rural
51	No	Female	No	Graduate	No	4166	0	116	360	0	Semiurb
52	No	Male	Yes	Graduate	No	5726	4595	258	360	1	Semiurb
53	No	Male	No	Not Graduate	No	3200	2254	126	180	0	Urban
54	No	Male	Yes	Graduate	No	10750	0	312	360	1	Urban
55	No	Male	Yes	Not Graduate	Yes	7100	0	125	60	1	Urban
56	No	Female	No	Graduate	No	4300	0	136	360	0	Semiurb
57	No	Male	Yes	Graduate	No	3208	3066	172	360	1	Urban

58	No	Male	Yes	Not Graduate	Yes	1875	1875	97	360	1	Semiurb
59	No	Male	No	Graduate	No	3500	0	81	300	1	Semiurb
60	No	Male	Yes	Graduate	Yes	5266	1774	187	360	1	Semiurb
61	No	Male	No	Graduate	No	3750	0	113	480	1	Urban
62	No	Male	No	Graduate	No	3750	4750	176	360	1	Urban
63	No	Male	Yes	Graduate	Yes	1000	3022	110	360	1	Urban
64	No	Male	Yes	Graduate	No	3167	4000	180	300	0	Semiurb
65	No	Female	No	Graduate	No	3846	0	111	360	1	Semiurb
66	No	Female	Yes	Graduate	No	1378	1881	167	360	1	Urban
67	No	Male	Yes	Graduate	No	3988	0	50	240	1	Urban
68	No	Male	No	Graduate	No	2366	2531	136	360	1	Semiurb
69	No	Male	Yes	Graduate	No	2500	2118	104	360	1	Semiurb
70	No	Male	No	Graduate	No	8566	0	210	360	1	Urban
71	No	Male	Yes	Graduate	No	5695	4167	175	360	1	Semiurb
72	No	Male	Yes	Graduate	No	2958	2900	131	360	1	Semiurb
73	No	Male	Yes	Graduate	No	6250	5654	188	180	1	Semiurb
74	No	Male	Yes	Not Graduate	No	3273	1820	81	360	1	Urban
75	No	Male	No	Graduate	No	4133	0	122	360	1	Semiurb
76	No	Male	No	Not Graduate	No	3620	0	25	120	1	Semiurb
77	No	Female	Yes	Graduate	No	2484	2302	137	360	1	Semiurb
78	No	Male	Yes	Graduate	No	1977	997	50	360	1	Semiurb
79	No	Male	Yes	Not Graduate	No	4188	0	115	180	1	Semiurb
80	No	Male	Yes	Graduate	No	1759	3541	131	360	1	Semiurb
81	No	Male	Yes	Not Graduate	No	4288	3263	133	180	1	Urban
82	No	Male	No	Graduate	No	4843	3806	151	360	1	Semiurb
83	No	Male	Yes	Graduate	No	3052	1030	100	360	1	Urban
84	No	Male	Yes	Graduate	No	11417	1126	225	360	1	Urban
85	No	Male	Yes	Graduate	No	3800	3600	216	360	0	Urban
86	No	Male	Yes	Not Graduate	No	2071	754	94	480	1	Semiurb
87	No	Male	No	Graduate	No	5316	0	136	360	1	Urban
88	No	Male	Yes	Graduate	No	14583	0	185	180	1	Rural
89	No	Female	Yes	Graduate	No	3167	2283	154	360	1	Semiurb
90	No	Male	Yes	Graduate	No	5568	2142	175	360	1	Rural
91	No	Female	No	Graduate	No	10408	0	259	360	1	Urban
92	No	Male	Yes	Graduate	No	5667	2667	180	360	1	Rural
93	No	Female	No	Graduate	No	4166	0	44	360	1	Semiurb
94	No	Female	No	Graduate	No	2137	8980	137	360	0	Semiurb
95	No	Male	Yes	Graduate	No	2957	0	81	360	1	Semiurb
96	No	Male	Yes	Not Graduate	No	4300	2014	194	360	1	Rural
97	No	Male	Yes	Graduate	Yes	10513	3850	160	180	0	Urban
98	No	Male	No	Graduate	No	2014	1929	74	360	1	Urban
99	No	Male	No	Graduate	No	2718	0	70	360	1	Semiurb
100	No	Male	Yes	Graduate	Yes	3459	0	25	120	1	Semiurb
+ 390 more											

...Porém se perdeu muitos dados, como pode-se perceber foi de 614 para 490.

Settings

Impute Missing Values: Average/Most frequent
Normalize Features: Standardize to $\mu=0$, $\sigma^2=1$

Então decidi substituir por valores que seriam a média de resultados anteriores, e os numericos alterei para ficar dessa forma.

Data instances: 614
Features: 11
Meta attributes: None

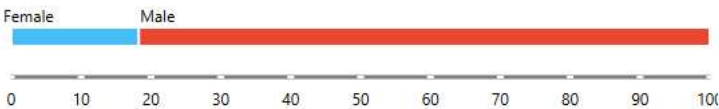
	Gender	Married	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
1	Male	No	Graduate	No	0.072991	-0.554487	0.00	0.2799	1	Urban	Y
2	Male	Yes	Graduate	No	-0.134412	-0.0387315	-0.2193	0.2799	1	Rural	N
3	Male	Yes	Graduate	Yes	-0.393747	-0.554487	-0.9576	0.2799	1	Urban	Y
4	Male	Yes	Not Graduate	No	-0.462062	0.25198	-0.3145	0.2799	1	Urban	Y
5	Male	No	Graduate	No	0.097728	-0.554487	-0.0645	0.2799	1	Urban	Y
6	Male	Yes	Graduate	Yes	0.002218	0.8806	1.4361	0.2799	1	Urban	Y
7	Male	Yes	Not Graduate	No	-0.503019	-0.0359954	-0.6123	0.2799	1	Urban	Y
8	Male	Yes	Graduate	No	-0.387850	0.301914	0.1380	0.2799	0	Semiurban	N
9	Male	Yes	Graduate	No	-0.228939	-0.0325753	0.2571	0.2799	1	Urban	Y
10	Male	Yes	Graduate	No	1.218457	3.19671	2.4127	0.2799	1	Semiurban	N
11	Male	Yes	Graduate	No	-0.360982	-0.315078	-0.9100	0.2799	1	Urban	Y
12	Male	Yes	Graduate	No	-0.475660	0.0748168	-0.4455	0.2799	1	Urban	Y
13	Male	Yes	Graduate	No	-0.381788	2.21787	0.6382	0.2799	1	Urban	Y
14	Male	No	Graduate	No	-0.581655	0.41683	-0.3860	0.2799	1	Rural	N
15	Male	Yes	Graduate	No	-0.672414	-0.183061	-1.5412	-3.4515	1	Urban	Y
16	Male	No	Graduate	No	-0.074288	-0.554487	-0.2550	0.2799	1	Urban	Y
17	Male	No	Not Graduate	No	-0.296107	-0.554487	-0.5527	-1.5858	1	Urban	Y
18	Female	No	Graduate	No	-0.310196	-0.554487	-0.8385	0.2799	0	Urban	N
19	Male	Yes	Not Graduate	No	-0.084609	-0.554487	-0.1597	0.2799	1	Rural	N
20	Male	Yes	Graduate	No	-0.459277	0.642559	-0.3741	0.00	1	Urban	Y
21	Male	Yes	Not Graduate	No	0.369678	-0.554487	-0.5051	0.2799	0	Urban	N
22	Male	Yes	Graduate	No	0.090356	1.36934	2.0077	0.2799	1	Urban	Y
23	Male	Yes	Not Graduate	No	-0.459277	0.0990997	-0.3622	0.2799	0	Semiurban	N
24	Male	Yes	Not Graduate	No	-0.333951	0.101152	-0.4098	0.2799	0	Rural	N
25	Male	Yes	Graduate	No	-0.276285	0.445901	0.0546	0.2799	1	Semiurban	N
26	Male	Yes	Graduate	Yes	0.680946	-0.554487	0.5310	0.2799	1	Semiurban	Y
27	Male	Yes	Graduate	No	-0.426676	0.216068	-0.2907	0.2799	1	Semiurban	Y
28	Male	Yes	Not Graduate	No	-0.192898	-0.198794	-0.4336	0.2799	1	Urban	Y
29	Male	No	Not Graduate	No	-0.648987	-0.554487	-1.3268	0.2799	1	Urban	N
30	Female	No	Graduate	No	-0.270878	0.157926	-0.3145	0.2799	1	Semiurban	Y
31	Male	Yes	Graduate	No	-0.202727	0.597755	0.6501	0.2799	1	Urban	N
32	Male	No	Graduate	No	-0.366389	-0.554487	-0.8624	0.2799	1	Urban	N
33	Male	No	Graduate	Yes	-0.116555	-0.554487	-0.4813	0.2799	1	Rural	N
34	Male	Yes	Graduate	No	-0.311835	0.0156485	-0.3860	0.2799	1	Semiurban	Y
35	Male	No	Graduate	No	1.162593	0.471552	2.0673	0.2799	1	Rural	N
36	Male	Yes	Graduate	No	-0.512521	0.152454	0.00	0.2799	1	Urban	Y
37	Male	Yes	Graduate	No	-0.585751	-0.0996099	-0.5527	0.00	0	Urban	N
38	Female	Yes	Graduate	No	-0.284476	-0.0554902	-0.0287	0.2799	1	Semiurban	Y
39	Male	No	Graduate	No	-0.202727	1.91143	0.4476	0.2799	1	Urban	Y
40	Male	No	Not Graduate	No	-0.271206	0.0159906	-0.4336	0.2799	1	Semiurban	Y
41	Male	No	Graduate	No	-0.295452	-0.554487	-0.7909	0.2799	1	Urban	N
42	Male	No	Graduate	No	-0.590338	-0.139625	-1.1839	0.2799	1	Urban	Y
43	Male	Yes	Graduate	No	-0.492042	-0.554487	-0.8505	0.2799	1	Urban	Y
44	Male	Yes	Graduate	No	-0.239588	0.244455	-0.1478	0.2799	1	Semiurban	Y
45	Male	Yes	Not Graduate	Yes	-0.116064	-0.554487	-0.6004	0.00	1	Urban	Y
46	Female	No	Graduate	No	-0.326579	-0.554487	-0.6956	0.00	1	Urban	Y
47	Male	Yes	Graduate	No	0.040226	-0.554487	-1.2196	0.2799	1	Urban	Y
48	Male	Yes	Graduate	No	0.068404	-0.554487	-0.0287	0.2799	1	Urban	Y
49	Female	Yes	Graduate	No	-0.451905	0.622038	-0.3145	0.2799	0	Urban	N
50	Female	No	Graduate	No	-0.229922	0.223593	-0.0287	0.2799	1	Semiurban	Y
51	Female	Yes	Not Graduate	No	-0.569368	0.00778224	-0.5527	0.2799	1	Semiurban	Y
52	Female	No	Graduate	No	-0.379658	-0.554487	-0.3145	0.2799	1	Semiurban	Y
53	Female	No	Graduate	No	-0.192242	-0.554487	-0.4098	0.2799	1	Semiurban	N
54	Male	Yes	Graduate	No	-0.129006	-0.554487	-0.1478	0.2799	1	Urban	N
55	Female	Yes	Graduate	Yes	0.998767	-0.554487	1.6624	0.2799	0	Urban	N
56	Male	Yes	Graduate	No	-0.441584	-0.155358	-0.5885	0.2799	1	Semiurban	Y
57	Male	Yes	Graduate	No	-0.535948	-0.0103445	-0.6004	0.2799	1	Semiurban	Y
58	Male	Yes	Graduate	No	-0.333787	0.197942	-0.1359	0.2799	1	Rural	N
59	Male	Yes	Graduate	No	0.438485	0.215042	0.4000	0.2799	1	Urban	Y

60	Male	Yes	Not Graduate	No	-0.335262	0.423328	-0.0287	0.2799	1	Urban	Y
61	Male	Yes	Graduate	No	-0.475660	0.743794	-0.3145	0.2799	1	Urban	Y
62	Male	Yes	Graduate	No	-0.388996	-0.554487	-0.5646	0.2799	1	Urban	Y
63	Male	Yes	Not Graduate	Yes	-0.457803	0.625116	0.2214	-2.5187	0	Rural	N
64	Male	Yes	Graduate	No	-0.075107	-0.554487	0.00	0.2799	0	Rural	N
65	Female	No	Graduate	No	-0.202727	-0.554487	-0.3622	0.2799	0	Semiurban	N
66	Male	Yes	Graduate	No	0.052840	1.01706	1.3289	0.2799	1	Semiurban	N
67	Male	No	Not Graduate	No	-0.360982	0.21641	-0.2431	-2.5187	0	Urban	N
68	Male	Yes	Graduate	No	0.875898	-0.554487	1.9720	0.2799	1	Urban	Y
69	Male	Yes	Not Graduate	Yes	0.277936	-0.554487	-0.2550	-4.3843	1	Urban	Y
70	Female	No	Graduate	No	-0.180775	-0.554487	-0.1240	0.2799	0	Semiurban	N
71	Male	Yes	Graduate	No	-0.359672	0.494125	0.3047	0.2799	1	Urban	Y
72	Male	Yes	Not Graduate	Yes	-0.578051	0.0867873	-0.5885	0.2799	1	Semiurban	Y
73	Male	No	Graduate	No	-0.311835	-0.554487	-0.7790	-0.6530	1	Semiurban	Y
74	Male	Yes	Not Graduate	No	-0.106234	-0.554487	-0.6123	0.00	0	Semiurban	N
75	Male	Yes	Graduate	Yes	-0.022519	0.0522439	0.4834	0.2799	1	Semiurban	Y
76	Male	No	Graduate	No	-0.270878	-0.554487	-0.3979	2.1455	1	Urban	N
77	Male	No	Graduate	No	-0.270878	1.07007	0.3524	0.2799	1	Urban	N
78	Male	Yes	Graduate	Yes	-0.721398	0.479076	-0.4336	0.2799	1	Urban	N
79	Male	Yes	Graduate	No	-0.366389	0.813565	0.4000	-0.6530	0	Semiurban	N
80	Male	Yes	Not Graduate	Yes	-0.339194	0.186313	-0.1955	0.2799	1	Semiurban	Y
81	Female	No	Graduate	No	-0.255151	-0.554487	-0.4217	0.2799	1	Semiurban	Y
82	Male	Yes	Graduate	Yes	-0.492862	-0.554487	0.00	0.2799	1	Semiurban	Y
83	Female	Yes	Graduate	No	-0.659472	0.0888393	0.2452	0.2799	1	Urban	N
84	Male	Yes	Graduate	No	0.097728	0.215042	1.4123	0.2799	1	Semiurban	N
85	Male	Yes	Graduate	No	-0.231888	-0.554487	-1.1482	-1.5858	1	Urban	Y
86	Male	No	Graduate	No	-0.497613	0.311148	-0.1240	0.2799	1	Semiurban	Y
87	Male	Yes	Not Graduate	No	-0.339194	0.129539	-0.5646	0.2799	1	Semiurban	Y
88	Male	Yes	Graduate	No	-0.475660	0.169896	-0.5051	0.2799	1	Semiurban	Y
89	Male	No	Graduate	No	0.518104	-0.554487	0.7573	0.2799	1	Urban	Y
90	Male	Yes	Graduate	No	0.047762	0.870681	0.3405	0.2799	1	Semiurban	Y
91	Male	Yes	Graduate	No	-0.400628	0.437351	-0.1835	0.2799	1	Semiurban	Y
92	Male	Yes	Graduate	No	0.138685	1.37925	0.4953	-2.5187	1	Semiurban	Y
93	Male	Yes	Not Graduate	No	-0.349023	0.0679765	-0.7790	0.2799	1	Urban	Y
94	Male	No	Graduate	No	-0.208133	-0.554487	-0.2907	0.2799	1	Semiurban	Y
95	Male	No	Not Graduate	No	-0.292176	-0.554487	-1.4459	-3.4515	1	Semiurban	Y
96	Male	No	Graduate	No	0.225840	-0.554487	0.00	0.2799	1	Urban	N
97	Female	Yes	Graduate	No	-0.478281	0.232827	-0.1121	0.2799	1	Semiurban	Y
98	Male	Yes	Graduate	No	-0.561341	-0.2135	-1.1482	0.2799	1	Semiurban	Y
99	Male	Yes	Not Graduate	No	-0.199123	-0.554487	-0.3741	-2.5187	1	Semiurban	Y
100	Male	Yes	Graduate	No	-0.597054	0.656581	-0.1835	0.2799	1	Semiurban	Y
+ 514 more											

Agora não existem mais dados perdidos.

Box Plot

Sat Nov 23 24, 01:16:18



Box plot for attribute 'Gender'

No gender, por exemplo, agora só contém "female" e "male"

Categorical variables

Preset: first value as base
Credit_History: keep categorical
Property_Area: keep categorical
Loan_Status: keep categorical

Numeric variables

keep as it is

Unlisted

Any unlisted attributes default to preset option, and unlisted meta attributes and target variables are kept as they are

Aqui pensei em tirar os valores "string" por numericos.

Data instances: 614
Features: 11
Meta attributes: None

	Gender=Male	Married=Yes	Education=Not Graduate	Self_Employed=Yes	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
1	1	0	0	0	0.072991	-0.554487	0.00	0.2799	1
2	1	1	0	0	-0.134412	-0.0387315	-0.2193	0.2799	1
3	1	1	0	1	-0.393747	-0.554487	-0.9576	0.2799	1
4	1	1	1	0	-0.462062	0.25198	-0.3145	0.2799	1
5	1	0	0	0	0.097728	-0.554487	-0.0645	0.2799	1
6	1	1	0	1	0.002218	0.8806	1.4361	0.2799	1
7	1	1	1	0	-0.503019	-0.0359954	-0.6123	0.2799	1
8	1	1	0	0	-0.387850	0.301914	0.1380	0.2799	0
9	1	1	0	0	-0.228939	-0.0325753	0.2571	0.2799	1
10	1	1	0	0	1.218457	3.19671	2.4127	0.2799	1
11	1	1	0	0	-0.360982	-0.315078	-0.9100	0.2799	1
12	1	1	0	0	-0.475660	0.0748168	-0.4455	0.2799	1
13	1	1	0	0	-0.381788	2.21787	0.6382	0.2799	1
14	1	0	0	0	-0.581655	0.41683	-0.3860	0.2799	1
15	1	1	0	0	-0.672414	-0.183061	-1.5412	-3.4515	1
16	1	0	0	0	-0.074288	-0.554487	-0.2550	0.2799	1
17	1	0	1	0	-0.296107	-0.554487	-0.5527	-1.5858	1
18	0	0	0	0	-0.310196	-0.554487	-0.8385	0.2799	0
19	1	1	1	0	-0.084609	-0.554487	-0.1597	0.2799	1
20	1	1	0	0	-0.459277	0.642559	-0.3741	0.00	1
21	1	1	1	0	0.369678	-0.554487	-0.5051	0.2799	0
22	1	1	0	0	0.090356	1.36934	2.0077	0.2799	1
23	1	1	1	0	-0.459277	0.0990997	-0.3622	0.2799	0
24	1	1	1	0	-0.333951	0.101152	-0.4098	0.2799	0
25	1	1	0	0	-0.276285	0.445901	0.0546	0.2799	1
26	1	1	0	1	0.680946	-0.554487	0.5310	0.2799	1
27	1	1	0	0	-0.426676	0.216068	-0.2907	0.2799	1
28	1	1	1	0	-0.192898	-0.198794	-0.4336	0.2799	1
29	1	0	1	0	-0.648987	-0.554487	-1.3268	0.2799	1
30	0	0	0	0	-0.270878	0.157926	-0.3145	0.2799	1
31	1	1	0	0	-0.202727	0.597755	0.6501	0.2799	1
32	1	0	0	0	-0.366389	-0.554487	-0.8624	0.2799	1
33	1	0	0	1	-0.116555	-0.554487	-0.4813	0.2799	1
34	1	1	0	0	-0.311835	0.0156485	-0.3860	0.2799	1
35	1	0	0	0	1.162593	0.471552	2.0673	0.2799	1
36	1	1	0	0	-0.512521	0.152454	0.00	0.2799	1
37	1	1	0	0	-0.585751	-0.0996099	-0.5527	0.00	0
38	0	1	0	0	-0.284476	-0.0554902	-0.0287	0.2799	1
39	1	0	0	0	-0.202727	1.91143	0.4476	0.2799	1
40	1	0	1	0	-0.271206	0.0159906	-0.4336	0.2799	1
41	1	0	0	0	-0.295452	-0.554487	-0.7909	0.2799	1
42	1	0	0	0	-0.590338	-0.139625	-1.1839	0.2799	1
43	1	1	0	0	-0.492042	-0.554487	-0.8505	0.2799	1
44	1	1	0	0	-0.239588	0.244455	-0.1478	0.2799	1
45	1	1	1	1	-0.116064	-0.554487	-0.6004	0.00	1
46	0	0	0	0	-0.326579	-0.554487	-0.6956	0.00	1
47	1	1	0	0	0.040226	-0.554487	-1.2196	0.2799	1
48	1	1	0	0	0.068404	-0.554487	-0.0287	0.2799	1
49	0	1	0	0	-0.451905	0.622038	-0.3145	0.2799	0
50	0	0	0	0	-0.229922	0.223593	-0.0287	0.2799	1
51	0	1	1	0	-0.569368	0.00778224	-0.5527	0.2799	1
52	0	0	0	0	-0.379658	-0.554487	-0.3145	0.2799	1
53	0	0	0	0	-0.192242	-0.554487	-0.4098	0.2799	1
54	1	1	0	0	-0.129006	-0.554487	-0.1478	0.2799	1
55	0	1	0	1	0.998767	-0.554487	1.6624	0.2799	0
56	1	1	0	0	-0.441584	-0.155358	-0.5885	0.2799	1
57	1	1	0	0	-0.535948	-0.0103445	-0.6004	0.2799	1
58	1	1	0	0	-0.333787	0.197942	-0.1359	0.2799	1
59	1	1	0	0	0.438485	0.215042	0.4000	0.2799	1
60	1	1	1	0	-0.335262	0.423328	-0.0287	0.2799	1
61	1	1	0	0	-0.475660	0.743794	-0.3145	0.2799	1
62	1	1	0	0	-0.388996	-0.554487	-0.5646	0.2799	1
63	1	1	1	1	-0.457803	0.625116	0.2214	-2.5187	0
64	1	1	0	0	-0.075107	-0.554487	0.00	0.2799	0
65	0	0	0	0	-0.202727	-0.554487	-0.3622	0.2799	0
66	1	1	0	0	0.052840	1.01706	1.3289	0.2799	1
67	1	0	1	0	-0.360982	0.21641	-0.2431	-2.5187	0

68	1	1	0	0	0.875898	-0.554487	1.9720	0.2799	1
69	1	1	1	1	0.277936	-0.554487	-0.2550	-4.3843	1
70	0	0	0	0	-0.180775	-0.554487	-0.1240	0.2799	0
71	1	1	0	0	-0.359672	0.494125	0.3047	0.2799	1
72	1	1	1	1	-0.578051	0.0867873	-0.5885	0.2799	1
73	1	0	0	0	-0.311835	-0.554487	-0.7790	-0.6530	1
74	1	1	1	0	-0.106234	-0.554487	-0.6123	0.00	0
75	1	1	0	1	-0.022519	0.0522439	0.4834	0.2799	1
76	1	0	0	0	-0.270878	-0.554487	-0.3979	2.1455	1
77	1	0	0	0	-0.270878	1.07007	0.3524	0.2799	1
78	1	1	0	1	-0.721398	0.479076	-0.4336	0.2799	1
79	1	1	0	0	-0.366389	0.813565	0.4000	-0.6530	0
80	1	1	1	1	-0.339194	0.186313	-0.1955	0.2799	1
81	0	0	0	0	-0.255151	-0.554487	-0.4217	0.2799	1
82	1	1	0	1	-0.492862	-0.554487	0.00	0.2799	1
83	0	1	0	0	-0.659472	0.0888393	0.2452	0.2799	1
84	1	1	0	0	0.097728	0.215042	1.4123	0.2799	1
85	1	1	0	0	-0.231888	-0.554487	-1.1482	-1.5858	1
86	1	0	0	0	-0.497613	0.311148	-0.1240	0.2799	1
87	1	1	1	0	-0.339194	0.129539	-0.5646	0.2799	1
88	1	1	0	0	-0.475660	0.169896	-0.5051	0.2799	1
89	1	0	0	0	0.518104	-0.554487	0.7573	0.2799	1
90	1	1	0	0	0.047762	0.870681	0.3405	0.2799	1
91	1	1	0	0	-0.400628	0.437351	-0.1835	0.2799	1
92	1	1	0	0	0.138685	1.37925	0.4953	-2.5187	1
93	1	1	1	0	-0.349023	0.0679765	-0.7790	0.2799	1
94	1	0	0	0	-0.208133	-0.554487	-0.2907	0.2799	1
95	1	0	1	0	-0.292176	-0.554487	-1.4459	-3.4515	1
96	1	0	0	0	0.225840	-0.554487	0.00	0.2799	1
97	0	1	0	0	-0.478281	0.232827	-0.1121	0.2799	1
98	1	1	0	0	-0.561341	-0.2135	-1.1482	0.2799	1
99	1	1	1	0	-0.199123	-0.554487	-0.3741	-2.5187	1
100	1	1	0	0	-0.597054	0.656581	-0.1835	0.2799	1

+ 514 more

Com o continueze ele pega o primeiro valor e valoriza ele com 0 ou 1, no caso gender o "Male" ficou como número 1 e "female" 0. Porém não irei utilizar esses dados.

Test and Score

Sat Nov 23 24, 01:19:16

Settings

Sampling type: Stratified 10-fold Cross validation
Target class: None, show average over classes

Scores

Model	AUC	CA	F1	Prec	Recall	MCC
kNN (1)	0.638	0.770	0.754	0.742	0.770	0.146
Logistic Regression	0.740	0.809	0.766	0.761	0.809	0.173
SVM	0.626	0.805	0.740	0.727	0.805	0.060

Esse foi o resultado sem o tratamento de dados com o kNN, LR e SVM. Com acuracia de 0.770, 0.809 e 0.805 respectivamente

Test and Score

Sat Nov 23 24, 01:20:24

Settings

Sampling type: Stratified Shuffle split, 10 random samples with 80% data
Target class: None, show average over classes

Scores

Model	AUC	CA	F1	Prec	Recall	MCC
kNN	0.665	0.794	0.774	0.762	0.794	0.182
Logistic Regression (1)	0.762	0.815	0.770	0.764	0.815	0.158
SVM (1)	0.636	0.820	0.754	0.760	0.820	0.097

Já esse foi o resultado depois dos dados serem tratados, muito melhor em comparação a antes. Foi usado 80% para treinamento e 20% para verificação.

Confusion matrix for kNN (1) (showing number of instances)

		Predicted		
		Female	Male	Σ
Actual	Female	26	86	112
	Male	52	437	489
	Σ	78	523	601

Antes do tratamento o acerto foi de 26 (Female) e 437 (Male) sendo um total de 601

Confusion matrix for kNN (showing number of instances)

		Predicted		
		Female	Male	Σ
Actual	Female	51	169	220
	Male	84	926	1,010
	Σ	135	1.095	1.230

Após o treinamento aumentou consideravelmente, chegando a 1.230