

Olá!

Gostaríamos de realizar um teste que será usado para sabermos a sua proficiência nas habilidades para a vaga. O teste consiste na análise de um conjunto de dados e seus resultados deverão ser armazenados no GitHub. O link do seu repositório deverá ser compartilhado conosco ao final do teste.

Quando usar alguma referência ou biblioteca externa, informe no arquivo README do seu projeto. Se tiver alguma dúvida, use o bom senso e se precisar deixe isso registrado na documentação do seu projeto.

Dataset: Bank Marketing

Fonte oficial do dataset: <https://archive.ics.uci.edu/ml/datasets/bank+marketing>

Dados: <https://archive.ics.uci.edu/ml/machine-learning-databases/00222/bank.zip>

Arquivos incluídos no link acima:

- bank.csv: uma versão reduzida do conjunto de dados;
- bank-full.csv: o conjunto completo;
- bank-names.txt: com a descrição dos campos do conjunto de dados.

Para a resolução deste teste utilize o arquivo bank-full.csv e para uma descrição sobre o dataset use bank-names.txt e/ou a fonte oficial.

Questões

1. Qual profissão tem mais tendência a fazer um empréstimo? De qual tipo?
2. Fazendo uma relação entre número de contatos e sucesso da campanha quais são os pontos relevantes a serem observados?
3. Baseando-se nos resultados de adesão desta campanha qual o número médio e o máximo de ligações que você indica para otimizar a adesão?
4. O resultado da campanha anterior tem relevância na campanha atual?
5. Qual o fator determinante para que o banco exija um seguro de crédito?
6. Quais são as características mais proeminentes de um cliente que possua empréstimo imobiliário?

1. Qual profissão tem mais tendência a fazer um empréstimo? De qual tipo?

A palavra em inglês para designar empréstimo é *loan*. Assumindo que o passado recente é um bom preditor do futuro próximo, então tenho que ver as profissões com mais ocorrências de um dos dois tipos de empréstimo. Para isso, tabelas de contingência ajudam:

	no	yes	
admin.	1602	3569	5171
blue-collar	2133	7599	9732
entrepreneur	471	1016	1487
housemaid	750	490	1240
management	4233	5225	9458
retired	1565	699	2264
self-employed	715	864	1579
services	1091	3063	4154
student	681	257	938
technician	2989	4608	7597
unemployed	714	589	1303
unknown	260	28	288
	17204	28007	45211

Optei por combinar os dois tipos de *loan* (chamei de *anyloan*) e que fazendo isso, 62% dos clientes analisados contrataram um empréstimo. Destes, as profissões do tipo *blue-collar* (colarinho azul) são as que têm proporcionalmente mais empréstimos contratados (78% deste grupo tem algum empréstimo contratado). Os números estão expressos em percentuais nas linhas:

	no	yes	
admin.	30.980468	69.019532	100.000000
blue-collar	21.917386	78.082614	100.000000
entrepreneur	31.674512	68.325488	100.000000
housemaid	60.483871	39.516129	100.000000
management	44.755762	55.244238	100.000000
retired	69.125442	30.874558	100.000000
self-employed	45.281824	54.718176	100.000000
services	26.263842	73.736158	100.000000
student	72.601279	27.398721	100.000000
technician	39.344478	60.655522	100.000000
unemployed	54.796623	45.203377	100.000000
unknown	90.277778	9.722222	100.000000
	38.052686	61.947314	100.000000

Logo, vemos que os clientes com empregos do tipo colarinho azul são os que mais têm empréstimos contratados, podendo supor que eles são os que têm a maior tendência de tê-los.

Se desmembrarmos os tipos de empréstimo, podemos ver que 55% dos clientes analisados possuem crédito imobiliário e mais uma vez os clientes com tipo de emprego colarinho azul são os que proporcionalmente possuem mais contratações (72%):

	no	yes	
admin.	1989	3182	5171
blue-collar	2684	7048	9732
entrepreneur	618	869	1487
housemaid	842	398	1240
management	4780	4678	9458
retired	1773	491	2264
self-employed	814	765	1579
services	1388	2766	4154
student	689	249	938
technician	3482	4115	7597
unemployed	760	543	1303
unknown	262	26	288
	20081	25130	45211

Olhando os números de empréstimo imobiliário expressos em percentuais nas linhas:

	no	yes	
admin.	38.464514	61.535486	100.000000

blue-collar	27.579120	72.420880	100.000000
entrepreneur	41.560188	58.439812	100.000000
housemaid	67.903226	32.096774	100.000000
management	50.539226	49.460774	100.000000
retired	78.312721	21.687279	100.000000
self-employed	51.551615	48.448385	100.000000
services	33.413577	66.586423	100.000000
student	73.454158	26.545842	100.000000
technician	45.833882	54.166118	100.000000
unemployed	58.326938	41.673062	100.000000
unknown	90.972222	9.027778	100.000000
	44.416182	55.583818	100.000000

Logo, vemos que os clientes com empregos do tipo colarinho azul são os que mais têm empréstimos imobiliários contratados, podendo supor que eles são os que têm a maior tendência de tê-los.

No caso dos outros tipos de empréstimos, vemos que eles são menos comuns (16% dos clientes analisados os têm) e neste outro grupo os clientes com empregos do tipo empreendedor (*entrepreneur*) são os que proporcionalmente possuem mais contratações:

	no	yes	
admin.	4180	991	5171
blue-collar	8048	1684	9732
entrepreneur	1131	356	1487
housemaid	1088	152	1240
management	8205	1253	9458
retired	1955	309	2264
self-employed	1350	229	1579
services	3318	836	4154
student	926	12	938
technician	6288	1309	7597
unemployed	1194	109	1303
unknown	284	4	288
	37967	7244	45211

Olhando os números de outros tipos de empréstimo em percentuais nas linhas:

	no	yes	
admin.	80.835428	19.164572	100.000000
blue-collar	82.696260	17.303740	100.000000
entrepreneur	76.059180	23.940820	100.000000
housemaid	87.741935	12.258065	100.000000
management	86.751956	13.248044	100.000000
retired	86.351590	13.648410	100.000000
self-employed	85.497150	14.502850	100.000000
services	79.874819	20.125181	100.000000
student	98.720682	1.279318	100.000000
technician	82.769514	17.230486	100.000000
unemployed	91.634689	8.365311	100.000000
unknown	98.611111	1.388889	100.000000
	83.977351	16.022649	100.000000

2. Fazendo uma relação entre número de contatos e sucesso da campanha, quais são os pontos relevantes a serem observados?

- A primeira coisa que chama a atenção é que apesar de um grande intervalo no número de contatos, como seria de se esperar existe uma grande concentração nos menores números de contato:

	no	yes	
1	14983	2561	17544
2	11104	1401	12505
3	4903	618	5521
4	3205	317	3522
5	1625	139	1764
6	1199	92	1291
7	688	47	735
8	508	32	540
9	306	21	327
10	252	14	266
11	185	16	201
12	151	4	155
13	127	6	133
14	89	4	93
15	80	4	84
16	77	2	79
17	63	6	69
18	51	0	51
19	44	0	44
20	42	1	43
21	34	1	35
22	23	0	23
23	22	0	22
24	19	1	20
25	22	0	22
26	13	0	13
27	10	0	10
28	16	0	16
29	15	1	16
30	8	0	8
31	12	0	12
32	8	1	9
33	6	0	6
34	5	0	5
35	4	0	4
36	4	0	4
37	2	0	2
38	3	0	3
39	1	0	1
41	2	0	2
43	3	0	3
44	1	0	1
46	1	0	1
50	2	0	2
51	1	0	1
55	1	0	1
58	1	0	1
63	1	0	1
	39922	5289	45211

- A segunda coisa que chama muito a atenção é que o percentual de sucesso da campanha é decrescente conforme se aumenta o número de contatos. A partir de 2 contatos, o percentual de sucesso é menor que a média geral (tabela resumida [números em percentuais nas linhas]):

	no	yes	
1	85.402417	14.597583	100.000000
2	88.796481	11.203519	100.000000
3	88.806376	11.193624	100.000000
4	90.999432	9.000568	100.000000
5	92.120181	7.879819	100.000000
6	92.873741	7.126259	100.000000
(...)			
	88.301520	11.698480	100.000000

3. Baseando-se nos resultados de adesão desta campanha qual o número médio e o máximo de ligações que você indica para otimizar a adesão?

Resumindo então a tabela anterior para 3 linhas (1 contato, 2 contatos e mais de 2 contatos) temos:

	no	yes
1	14983	2561 17544
2	11104	1401 12505
2+	13835	1327 15162
	39922	5289 45211

Em percentuais de adesão à campanha temos (números em percentuais nas linhas):

	no	yes
1	85.402417	14.597583 100.000000
2	88.796481	11.203519 100.000000
2+	91.247856	8.752144 100.000000
	88.301520	11.698480 100.000000

Esses números mostram que, além do custo extra, não justifica fazer mais do que 2 contatos. Minha sugestão é a de fazer em média 1 ligação, no máximo 2, não mais do que isso.

4. O resultado da campanha anterior tem relevância na campanha atual?

Uma tabela cruzada vai mostrar algumas coisas:

	no	yes	
failure	4283	618	4901
other	1533	307	1840
success	533	978	1511
unknown	33573	3386	36959
	39922	5289	45211

O primeiro problema aqui é que uma quantidade muito grande de casos tem a marcação de *unknown* (desconhecido) e *other* (outro). Em percentuais nas colunas, temos:

	no	yes	
failure	10.728420	11.684628	10.840282
other	3.839988	5.804500	4.069806
success	1.335103	18.491208	3.342107
unknown	84.096488	64.019663	81.747805
	100.000000	100.000000	100.000000

Ou seja, quase 86% tem uma resposta da campanha anterior que não podemos deduzir como tendo obtido sucesso ou não. Ainda assim, em relação à campanha atual, os casos de sucesso na campanha anterior tem um percentual significativamente maior de adesão à campanha mais atual, cerca de 65% (enquanto a média geral é de 12%). Em valores de percentuais nas linhas:

	no	yes	
failure	87.390329	12.609671	100.000000
other	83.315217	16.684783	100.000000
success	35.274653	64.725347	100.000000
unknown	90.838497	9.161503	100.000000
	88.301520	11.698480	100.000000

Este percentual nos indica que sim, o sucesso na campanha anterior influencia a adesão à campanha atual. Ainda que com a ressalva de ser minoria na base.

5. Qual o fator determinante para que o banco exija um seguro de crédito?

Esta questão é antes de mais nada conceitual. O banco exige um seguro de crédito para evitar o calote, o *default* na base original. Assim, estou entendendo que devo descobrir a variável mais relevante para prever o *default*.

Para responder a esta questão implícita, vou modelar esta variável.

Ainda conceitualmente, dentre as variáveis do banco de dados, o saldo em conta (*balance*) deve ser o (senão um dos) melhor(es) preditor(es) do *default*. Além deste, o empréstimo “comum” (*loan*) e o empréstimo imobiliário (*housing*) devem ser outros bons preditores.

Algo relevante é o fato de termos muito poucos casos de *default*. É o que esperamos em termos do negócio bancário, mas constitui uma dificuldade a mais em termos de modelagem:

no	yes
44396	815

Em termos percentuais:

no	yes
98.197341	1.802659

Com apenas 1,8% de ocorrências de *default*, vou aplicar a técnica de balanceamento de amostra em modelos de regressão, que consiste em selecionar todos os casos de clientes com *default* e amostrar aleatoriamente igual quantidade de clientes sem *default*. Isso vai tornar o modelo mais robusto. Exigiria uma correção no termo livre da equação (intercepto), mas aqui estamos mais preocupados com as variáveis mais relevantes do que com a fórmula em si. Assim, o novo conjunto de dados para os modelos é menor e possui proporções iguais:

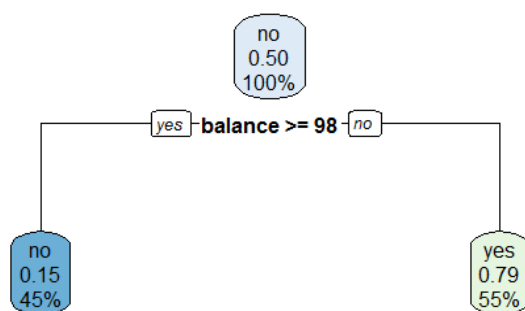
no	yes
815	815

Numa primeira modelagem, uma variável que se mostrou relevante (apenas menos relevante que o saldo em conta) foi a variável mês do último contato (*month* no banco de dados). É natural que o contato por parte do banco deva viabilizar alguma negociação de dívidas que impeça o *default*, mas este fato torna esta modelagem sazonalmente enviesada.

Assim, por conta do bom senso, eu decidi retirar da amostra balanceada as variáveis dia e mês do último contato (*day* e *month*, respectivamente).

E para ter a certeza de qual seria o fator determinante (ou as principais variáveis), eu optei por usar dois tipos completamente diferentes de modelos.

Um primeiro modelo mais simples é de uma árvore de decisão:



Por este modelo, temos que o fator determinante para o *default* (o que levaria o banco a exigir um seguro de crédito) é o saldo em conta (*balance* no banco de dados). Pelo menos por esta técnica nossa lógica estava sendo validada. Mais que isso, esta técnica ainda sugere um corte ótimo do saldo em 98 unidades monetárias... Com apenas este corte, a amostra balanceada parte de 50% de *default* para 2 grupos, um com 55% de *default* e outro com 45% de *default*...

Devido ao pequeno tamanho da amostra, a árvore de decisão não ficou maior, com mais variáveis preditivas.

Mas para ter certeza (e procurar outras variáveis relevantes), também farei uso de regressões logísticas. Sim, no plural, por que utilizarei regressões logísticas com 3 tipos de seleção de variáveis. Esses algoritmos *stepwise* vão nos ajudar a identificar quais variáveis (pela ordem) seriam mais relevantes para inclusão nos melhores modelos e quais seriam menos relevantes (para exclusão nos melhores modelos).

Primeiro, o algoritmo de seleção *backward* (retira de um modelo inicialmente com todas as variáveis, pela ordem apresentada, aquelas que são menos significantes):

	Step	Df	Deviance	Resid.	Df	Resid.	Dev	AIC
1			NA	NA	1599		1588.652	1650.652
2	- age	1	0.3661301		1600		1589.018	1649.018
3	- y	1	0.7521032		1601		1589.770	1647.770
4	- contact	2	3.3020826		1603		1593.072	1647.072
5	- previous	1	1.2271403		1604		1594.299	1646.299
6	- pdays	1	1.7291476		1605		1596.028	1646.028

Depois, o algoritmo de seleção *forward* (acrescenta de um modelo inicialmente sem nenhuma variável, pela ordem apresentada, aquelas que são mais significantes):

	Step	Df	Deviance	Resid.	Df	Resid.	Dev	AIC
1			NA	NA	1629		2259.660	2261.660
2	+ balance	-1	577.672925		1628		1681.987	1685.987
3	+ loan	-1	29.631440		1627		1652.355	1658.355
4	+ poutcome	-3	17.365801		1624		1634.990	1646.990
5	+ marital	-2	5.802669		1622		1629.187	1645.187
6	+ duration	-1	3.141187		1621		1626.046	1644.046
7	+ housing	-1	2.703832		1620		1623.342	1643.342
8	+ pdays	-1	2.163205		1619		1621.179	1643.179
9	+ campaign	-1	2.035817		1618		1619.143	1643.143

Finalmente, o algoritmo de seleção *stepwise* (acrescenta e retira de um modelo inicialmente sem nenhuma variável, pela ordem apresentada, aquelas que são mais significantes):

	Step	Df	Deviance	Resid.	Df	Resid.	Dev	AIC
1			NA	NA	1629		2259.660	2261.660
2	+ balance	-1	577.672925		1628		1681.987	1685.987
3	+ loan	-1	29.631440		1627		1652.355	1658.355
4	+ poutcome	-3	17.365801		1624		1634.990	1646.990
5	+ marital	-2	5.802669		1622		1629.187	1645.187
6	+ duration	-1	3.141187		1621		1626.046	1644.046
7	+ housing	-1	2.703832		1620		1623.342	1643.342
8	+ pdays	-1	2.163205		1619		1621.179	1643.179
9	+ campaign	-1	2.035817		1618		1619.143	1643.143

Os 2 últimos algoritmos produziram o mesmo resultado e mostram como a variável saldo em conta (*balance*) é a mais relevante e em que proporção (na quarta coluna da tabela acima).

Além delas, ter um empréstimo que não o imobiliário (*loan*) é a segunda variável mais relevante, o resultado da campanha de marketing anterior (*poutcome*) é a terceira variável mais relevante, o estado civil (*marital*) é a quarta variável mais relevante e assim por diante.

6. Quais são as características mais proeminentes de um cliente que possua empréstimo imobiliário?

É possível responder a esta pergunta de 2 maneiras: Apresentando tabelas de frequência das variáveis para comparação entre todos os clientes analisados e os clientes com empréstimo imobiliário (*housing:yes*):

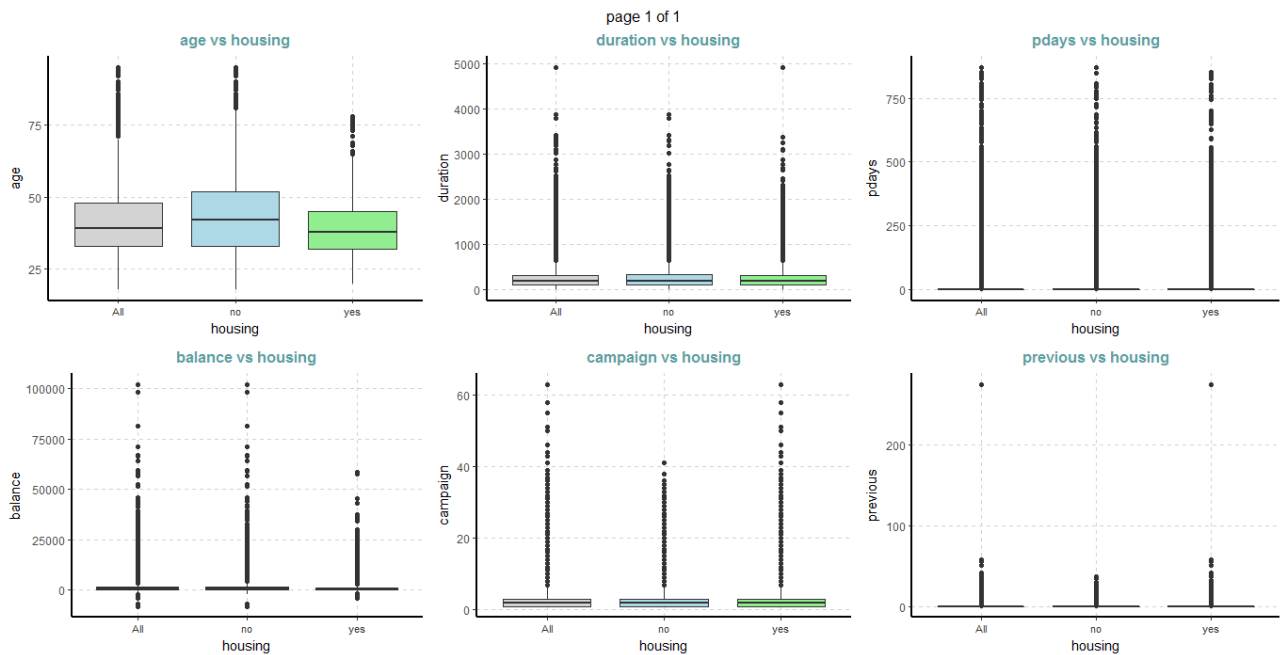
	VARIABLE	CATEGORY	housing:no	housing:yes	TOTAL	housing.no	housing.yes	housing.tot
1	job	admin.	1989	3182	5171	9.9048852	12.6621568	11.4374820
2	job	blue-collar	2684	7048	9732	13.3658682	28.0461600	21.5257349
3	job	entrepreneur	618	869	1487	3.0775360	3.4580183	3.2890226
4	job	housemaid	842	398	1240	4.1930183	1.5837644	2.7426954
5	job	management	4780	4678	9458	23.8035954	18.6152010	20.9196877
6	job	retired	1773	491	2264	8.8292416	1.9538400	5.0076309
7	job	self-employed	814	765	1579	4.0535830	3.0441703	3.4925129
8	job	services	1388	2766	4154	6.9120064	11.0067648	9.1880295
9	job	student	689	249	938	3.4311040	0.9908476	2.0747163
10	job	technician	3482	4115	7597	17.3397739	16.3748508	16.8034328
11	job	unemployed	760	543	1303	3.7846721	2.1607640	2.8820420
12	job	unknown	262	26	288	1.3047159	0.1034620	0.6370131
13	job	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
14	marital	divorced	2300	2907	5207	11.4536129	11.5678472	11.5171087
15	marital	married	11893	15321	27214	59.2251382	60.9669717	60.1933158
16	marital	single	5888	6902	12790	29.3212489	27.4651811	28.2895755
17	marital	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
18	education	primary	2957	3894	6851	14.7253623	15.4954238	15.1533919
19	education	secondary	9164	14038	23202	45.6351775	55.8615201	51.3193692
20	education	tertiary	6923	6378	13301	34.4753747	25.3800239	29.4198315
21	education	unknown	1037	820	1857	5.1640855	3.2630322	4.1074075
22	education	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
23	default	no	19701	24695	44396	98.1076640	98.2690012	98.1973414
24	default	yes	380	435	815	1.8923360	1.7309988	1.8026586
25	default	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
26	loan	no	17204	20763	37967	85.6730243	82.6223637	83.9773506
27	loan	yes	2877	4367	7244	14.3269757	17.3776363	16.0226494
28	loan	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
29	contact	cellular	14682	14603	29285	73.1138888	58.1098289	64.7740594
30	contact	telephone	1736	1170	2906	8.6449878	4.6557899	6.4276393
31	contact	unknown	3663	9357	13020	18.2411235	37.2343812	28.7983013
32	contact	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
33	month	apr	863	2069	2932	4.2975947	8.2331874	6.4851474
34	month	aug	5144	1103	6247	25.6162542	4.3891763	13.8174338
35	month	dec	173	41	214	0.8615109	0.1631516	0.4733361
36	month	feb	1517	1132	2649	7.5544047	4.5045762	5.8591936
37	month	jan	882	521	1403	4.3922115	2.0732193	3.1032271
38	month	jul	3557	3338	6895	17.7132613	13.2829288	15.2507133
39	month	jun	3114	2227	5341	15.5071959	8.8619180	11.8134967
40	month	mar	364	113	477	1.8126587	0.4496618	1.0550530
41	month	may	1690	12076	13766	8.4159155	48.0541186	30.4483422
42	month	nov	1756	2214	3970	8.7445844	8.8101870	8.7810489
43	month	oct	571	167	738	2.8434839	0.6645444	1.6323461
44	month	sep	450	129	579	2.2409243	0.5133307	1.2806618
45	month	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
46	poutcome	failure	1431	3470	4901	7.1261391	13.8081974	10.8402822
47	poutcome	other	641	1199	1840	3.1920721	4.7711898	4.0698060
48	poutcome	success	1040	471	1511	5.1790249	1.8742539	3.3421070
49	poutcome	unknown	16969	19990	36959	84.5027638	79.5463589	81.7478047
50	poutcome	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000
51	y	no	16727	23195	39922	83.2976445	92.3000398	88.3015195
52	y	yes	3354	1935	5289	16.7023555	7.6999602	11.6984805
53	y	TOTAL	20081	25130	45211	100.0000000	100.0000000	100.0000000

No tipo de trabalho, vemos que os clientes dos tipos colarinho azul (*blue-collar*) e serviços (*services*) estão mais concentrados entre os que têm empréstimo imobiliário, enquanto os clientes dos tipos empregados domésticos (*housemaid*), aposentados (*retired*), estudantes (*student*) estão bem menos concentrados entre os que têm empréstimo imobiliário.

Em termos de estado civil não há diferenças significativas em relação ao empréstimo imobiliário, assim como em relação ao calote (*default*). Em termos de educação, há algumas diferenças não muito significativas.

Há diferenças significativas no caso da forma de contato das campanhas de marketing (*contact*), mas como as opções são telefone/celular/sem informação, é difícil concluir alguma coisa. Difícil de interpretar também são as diferenças significativas em relação ao mês de contato. Possivelmente por conta de alguma campanha ou condição bastante específica (particularmente no mês de maio, mês em que foram feitas as maiores quantidades de contatos).

Interessante é a relação inversa entre o status da última campanha de marketing (*poutcome*). Há uma maior concentração de quem tem crédito imobiliário dentre quem não aderiu à campanha anterior (*failure*) e uma concentração menor em quem aderiu à campanha (*sucess*). De maneira similar, há uma maior concentração de quem tem crédito imobiliário dentre quem não aderiu à campanha mais recente (*no*) e uma concentração menor em quem aderiu à campanha (*yes*).



No caso das variáveis numéricas, é preciso estudar comparativamente as distribuições das mesmas:

A análise neste caso fica muito prejudicada pela presença de *outliers*, mas dentre os *boxplots* é possível ver que o público com empréstimo imobiliário é geralmente mais jovem que o público sem empréstimo imobiliário (no primeiro grupo de *boxplots*).

Outra análise é tentar entender via modelagem quais são as variáveis mais significantes para prever a contratação do empréstimo imobiliário, como fizemos anteriormente.

A diferença é que agora não será necessário utilizar o balanceamento de amostra, uma vez que o crédito imobiliário não é algo tão raro:

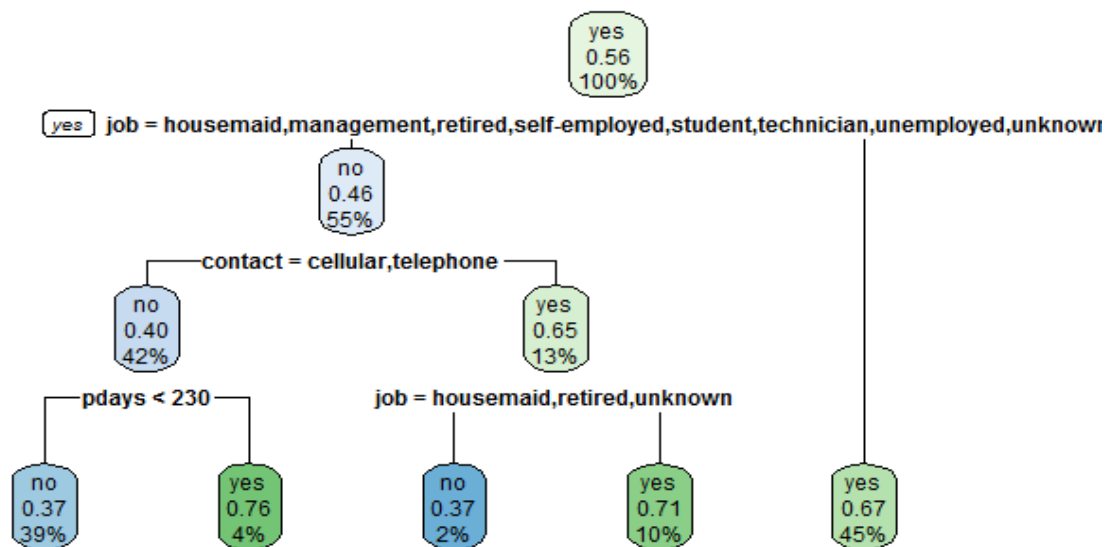
no	yes
20081	25130

Em valores percentuais:

no	yes
44.41618	55.58382

A árvore de decisão nos sugere as características mais relevantes para o empréstimo imobiliário:

- O tipo de trabalho
- A forma de contato do banco com seus clientes
- O número de dias passados desde que o cliente foi contatado anteriormente.



No caso dos modelos mais complexos de regressão logística, os 3 algoritmos não chegaram a gerar modelos diferentes, no sentido de que a seleção *backward* não eliminou nenhuma variável como sendo pouco significativa e as seleções *forward* e *stepwise* acrescentaram todas as variáveis.

Mas o importante mais uma vez é a ordem em que as variáveis foram acrescentadas:

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
1		NA	NA	45210	62110.72	62112.72
2	+ job	-11	3713.995720	45199	58396.73	58420.73
3	+ contact	-2	1585.543513	45197	56811.18	56839.18
4	+ pdays	-1	1393.122997	45196	55418.06	55448.06
5	+ age	-1	842.610016	45195	54575.45	54607.45
6	+ y	-1	528.430606	45194	54047.02	54081.02
7	+ marital	-2	240.330580	45192	53806.69	53844.69
8	+ poutcome	-3	226.585173	45189	53580.10	53624.10
9	+ duration	-1	120.894874	45188	53459.21	53505.21
10	+ education	-3	83.828908	45185	53375.38	53427.38
11	+ balance	-1	32.555972	45184	53342.82	53396.82
12	+ default	-1	13.779587	45183	53329.04	53385.04
13	+ campaign	-1	12.645431	45182	53316.40	53374.40
14	+ loan	-1	8.267298	45181	53308.13	53368.13
15	+ previous	-1	5.482079	45180	53302.65	53364.65

Os modelos de regressão não apenas confirmam o que já havíamos visto no modelo de árvore de decisão como a análise nas tabelas feita inicialmente:

- Que o tipo de trabalho é a variável mais proeminente na explicação do crédito imobiliário.
- Que a forma de contato é a segunda variável mais proeminente na explicação do crédito imobiliário.
- Que o número de dias desde o último contato é a terceira variável mais proeminente na explicação do crédito imobiliário.