Preprocessing

Admin

24 4 2019

Raw Dataset

text graphs are displayed; set 'tmp.img.dir' parameter to activate png graphs

Data Frame Summary

data

Dimensions: 4521×17

Duplicates: 0

			Freqs (% of			
No	Variable	Stats / Values	Valid)	Graph	Valid	Missing
1	age	Mean (sd): 41.2	67 distinct	:	4521	0
	[integer]	(10.6)	values	::.	(100%)	(0%)
		$\min < \max < \max$:		::::		
		19 < 39 < 87		:::::		
		IQR (CV) : 16 (0.3)	1=0 (10 00 ()	.:::::		
2	job	1. admin.	478 (10.6%)	II	4521	0
	[factor]	2. blue-collar	946 (20.9%)	IIII	(100%)	(0%)
		3. entrepreneur	168 (3.7%)			
		4. housemaid	112 (2.5%)			
		5. management6. retired	969 (21.4%)	TTTT		
		7. self-employed	230 (5.1%) 183 (4.0%)	IIII I		
		8. services	417 (9.2%)	1		
		9. student	84 (1.9%)			
		10. technician	768 (17.0%)	I		
		[2 others]	166 (3.7%)	1		
				III		
3	marital	1. divorced	528 (11.7%)	II	4521	0
	[factor]	2. married	2797 (61.9%)	IIIIIIIIIII	(100%)	(0%)
	[]	3. single	1196 (26.5%)	IIIII	(===,0)	(0,0)
4	education	1. primary	678 (15.0%)	II	4521	0
	[factor]	2. secondary	2306 (51.0%)	IIIIIIIII	(100%)	(0%)
		3. tertiary	1350 (29.9%)	IIIII	,	` /
		4. unknown	187 (4.1%)			
5	default	1. no	4445~(98.3%)	IIIIIIIIIIIIIIIII	4521	0
	[factor]	2. yes	76 (1.7%)		(100%)	(0%)
6	balance	Mean (sd) : 1422.7	2353 distinct	:	4521	0
	[integer]	(3009.6)	values	:	(100%)	(0%)
		$\min < \max < \max$:		:		
		-3313 < 444 < 71188		:		
		IQR (CV) : 1411 (2.1)		: .		

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
7	housing	1. no	1962 (43.4%)	IIIIIIII	4521	0
	[factor]	2. yes	2559~(56.6%)	IIIIIIIIII	(100%)	(0%)
8	loan	1. no	3830 (84.7%)	IIIIIIIIIIIIII	4521	0
	[factor]	2. yes	691 (15.3%)	III	(100%)	(0%)
9	contact	1. cellular	2896(64.1%)	IIIIIIIIIII	4521	ò
	[factor]	2. telephone	301 (6.7%)	I	(100%)	(0%)
	. ,	3. unknown	1324 (29.3%)	IIIII	,	,
10	day	Mean (sd) : 15.9 (8.2)	31 distinct	:.	4521	0
	[integer]	$\min < \max < \max$	values	: ::::	(100%)	(0%)
	[0.]	1 < 16 < 31		:::::::::::	(, -,	(-, -)
		IQR (CV) : 12 (0.5)		:::::::::		
		1410 (0.7) 112 (0.0)		:::::::::		
11	month	1. apr	293 (6.5%)	I	4521	0
	[factor]	2. aug	633 (14.0%)	II	(100%)	(0%)
	[ractor]	3. dec	20 (0.4%)	11	(10070)	(070)
		4. feb	222 (4.9%)			
		5. jan	148 (3.3%)			
		6. jul	706 (15.6%)			
		7. jun	531 (11.7%)	III		
		8. mar	49 (1.1%)	II		
		9. may	1398 (30.9%)	11		
		9. may 10. nov	389 (8.6%)			
		[2 others]	132 (2.9%)	IIIIII		
		[2 Others]	132 (2.970)	I		
12	duration	Mean (sd) : 264	875 distinct	:	4521	0
	[integer]	(259.9)	values	:	(100%)	(0%)
	. 0 1	$\min < \max < \max$:	,	,
		4 < 185 < 3025		: .		
		IQR (CV) : 225 (1)		::.		
13	campaign	Mean (sd) : $2.8 (3.1)$	32 distinct	:	4521	0
10	[integer]	$\min < \max < \max$	values	:	(100%)	(0%)
	[0.]	1 < 2 < 50		:	(, -,	(-, -,
		IQR (CV) : 2 (1.1)		· •		
		1410 (01) 12 (111)		· : .		
14	pdays	Mean (sd): 39.8	292 distinct		4521	0
	[integer]	(100.1)	values	· •	(100%)	(0%)
	[meger]	$\min < \max < \max$	varaes	· •	(10070)	(070)
		-1 < -1 < 871		· •		
		IQR (CV) : 0 (2.5)		•		
15	previous	Mean (sd) : 0.5 (1.7)	24 distinct		4521	0
	[integer]	$\min < \max < \max$:	values		(100%)	(0%)
	[meger]	0 < 0 < 25	varaes		(10070)	(070)
		IQR (CV) : 0 (3.1)				
		1Q1t (OV) : 0 (3.1)		· :.		
16	poutcome	1. failure	490 (10.8%)	II	4521	0
	[factor]	2. other	197 (4.4%)	11	(100%)	(0%)
		3. success	129 (2.9%)		(10070)	(070)
		4. unknown	3705 (82.0%)			
		T. UIIKIIUWII	3103 (04.070)	1111111111111111		
				IIIIIIIIIIIIII		

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
17	y [factor]	1. no 2. yes	4000 (88.5%) 521 (11.5%)	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	4521 (100%)	0 (0%)

Including Plots

You can also embed plots, for example:

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.