

Assignment 3 BankDataframe Insights

GundaSuman

November 29, 2018

R Markdown

10 insights from bank dataset . Shown in plots and tables.

1)Let's just start by seeing which job professions has most "single" status

Just taking this is a sample from population.Let's see what this data says Observation:Most of Technician,Management are single.May be psychologist might explain why with this job professions relative to others.

Table 1: Number of people with 'single' status for each profession:Sorted

marital	job	count
single	unknown	7
single	retired	11
single	housemaid	15
single	entrepreneur	20
single	unemployed	31
single	self-employed	41
single	student	74
single	services	119
single	admin.	143
single	blue-collar	174
single	technician	268
single	management	293

Table 2: Percentage of people with 'single' status for each profession:Sorted

marital	job	percentage
single	unknown	0.0058528
single	retired	0.0091973
single	housemaid	0.0125418
single	entrepreneur	0.0167224
single	unemployed	0.0259197
single	self-employed	0.0342809
single	student	0.0618729
single	services	0.0994983
single	admin.	0.1195652
single	blue-collar	0.1454849
single	technician	0.2240803
single	management	0.2449833

2)For each job,finding out Mean age of those who possess loan and housing . Repeat the same with all possible combinations of housing,loan

Just taking this is a sample from population.Let's see what this data says Observation:Most of Technician,Managemnt are single.May be psychologist might explain why with this job professions relative to others.

Table 3: MEAN AGE BASED ON JOB HAVING LOAN AND HOUSING

job	x
blue-collar	37.28155
services	37.43902
technician	38.10959
unemployed	38.11111
admin.	38.64286
entrepreneur	38.72414
management	39.00000
self-employed	41.93333
housemaid	46.20000
unknown	50.00000
retired	54.53846

Table 4: MEAN AGE BASED ON JOB HAVING HOUSING AND NO LOAN

job	x
student	26.25000
services	37.31967
admin.	38.52439
technician	39.04274
blue-collar	39.33953
management	39.64932
self-employed	40.61644
unemployed	41.65306
entrepreneur	41.87692
housemaid	44.38235
retired	57.08108

Table 5: MEAN AGE OF BASED ON JOB HAVING LOAN AND NO HOUSING

job	x
student	22.00000
admin.	40.85714
technician	41.13043
services	41.78788
blue-collar	42.96226
management	43.42373
unemployed	45.00000
entrepreneur	45.75000

job	x
self-employed	49.73333
housemaid	51.25000
retired	53.84211

Table 6: MEAN AGE OF BASED ON JOB WITH NO LOAN AND NO HOUSING

job	x
student	27.07937
technician	40.05034
unemployed	40.48485
self-employed	40.57500
services	41.05051
management	41.32187
admin.	41.82270
entrepreneur	42.96774
blue-collar	43.34343
unknown	48.05405
housemaid	48.49231
retired	64.50932

3)AVERAGE DURATION OF CALL - MONTHLY DATA

Observation:I find December to deviate from rest

Table 7: AVEARGE DURATION OF CALL FOR EACH MONTH

month	y
mar	198.8980
sep	215.7308
aug	243.7425
jun	252.4256
feb	255.6982
may	267.3162
jan	269.5405
jul	271.2054
nov	272.0668
oct	272.8000
apr	294.3549
dec	416.9000

4)BASED ON JOB PROFESSION,LET'S SEE WHICH EDUCATION BACK-GROUND MOST NUMBER OF PEOPLE HAVE

Observation:Understandably,most of the entrepreneur,manager,self employed have higher education back-ground.Also most of the housemaids just have primary education

Table 8: HIGHEST EDUCATION BACKGROUND FOR MAJORITY @ JOB

job	education	count
unknown	unknown	15
student	secondary	47
housemaid	primary	57
unemployed	secondary	68
entrepreneur	tertiary	73
self-employed	tertiary	88
retired	secondary	105
services	secondary	363
admin.	secondary	393
technician	secondary	520
blue-collar	secondary	524
management	tertiary	787

5) ANY CORRELATION BETWEEN BANK BALANCE AND EDUCATION

Intuitively we feel people with higher education background earn more. Let's see if it's true.

Observation: For Retired people job doesn't matter? NO. Wealth of individual cannot be decided based on bank data. May be people have just accumulated in bank instead investing as they get retired unlike young. Also interesting to see technician and management where primary education people earn more than secondary and tertiary. May be acumen and skillset matters in those sectors nothing to do with education background.

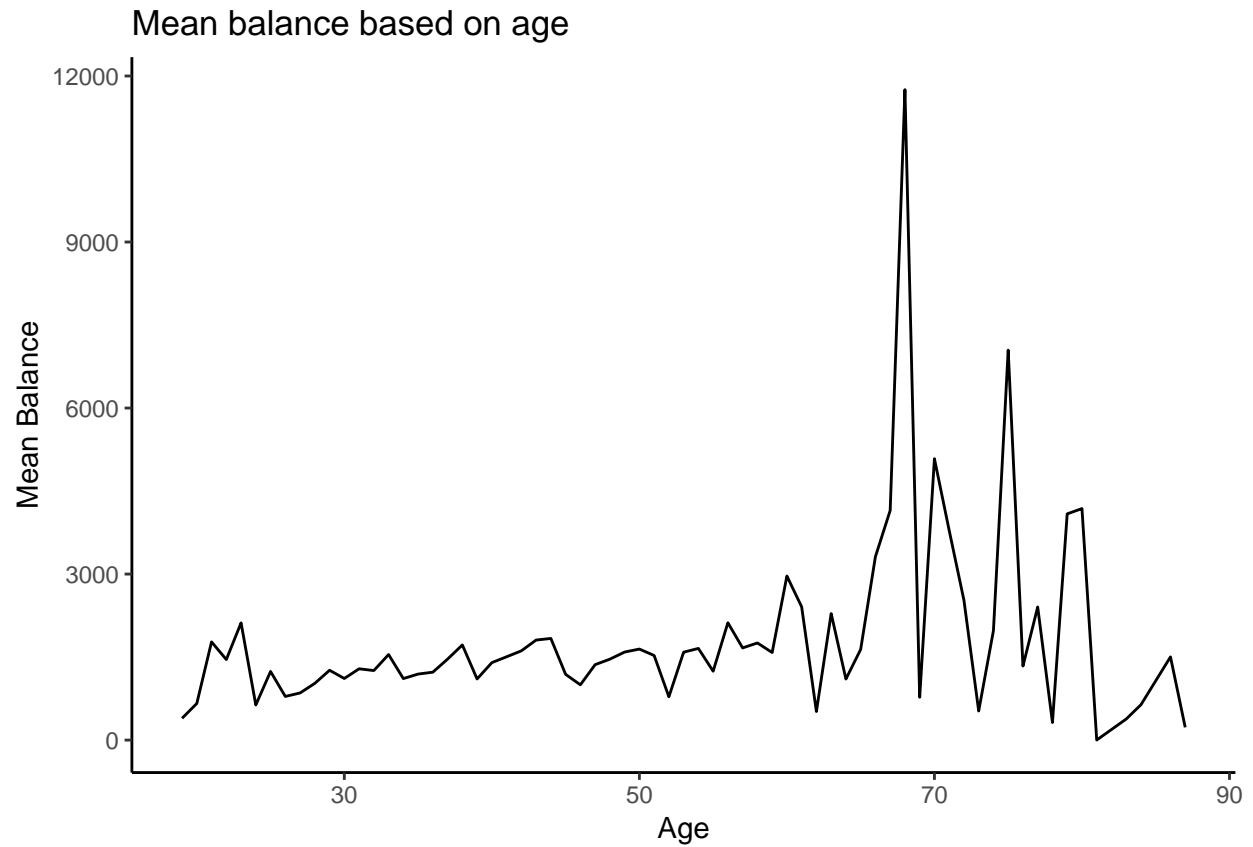
Table 9: MEAN BALANCE FOR EACH PROFESSION BASED ON RESPECTIVE JOB EDUCATION

job	education	Meanbalance
admin.	primary	390.5882
admin.	secondary	1269.6794
admin.	tertiary	1053.2941
admin.	unknown	1590.4706
blue-collar	primary	1072.2141
blue-collar	secondary	1068.5916
blue-collar	tertiary	2385.5000
blue-collar	unknown	1032.8780
entrepreneur	primary	383.9231
entrepreneur	secondary	1276.1724
entrepreneur	tertiary	2585.9041
entrepreneur	unknown	328.1818
housemaid	primary	1807.1053
housemaid	secondary	2011.8929
housemaid	tertiary	2392.5455
housemaid	unknown	4282.4000
management	primary	2685.4103
management	secondary	1250.1034
management	tertiary	1776.3431
management	unknown	2386.2593
retired	primary	2744.6000
retired	secondary	2089.0952
retired	tertiary	2476.7419

job	education	Meanbalance
retired	unknown	1265.1429
self-employed	primary	1471.7333
self-employed	secondary	1164.5526
self-employed	tertiary	1615.9659
self-employed	unknown	506.0000
services	primary	1107.3200
services	secondary	998.8843
services	tertiary	1894.8750
services	unknown	3058.0000
student	primary	1787.5000
student	secondary	1610.4255
student	tertiary	1175.6842
student	unknown	1754.8750
technician	primary	2593.0000
technician	secondary	1153.6077
technician	tertiary	1631.6303
technician	unknown	1780.0000
unemployed	primary	873.1923
unemployed	secondary	1025.1618
unemployed	tertiary	1224.7812
unemployed	unknown	3919.5000
unknown	primary	360.2857
unknown	secondary	1229.0000
unknown	tertiary	2497.7500
unknown	unknown	1648.6000

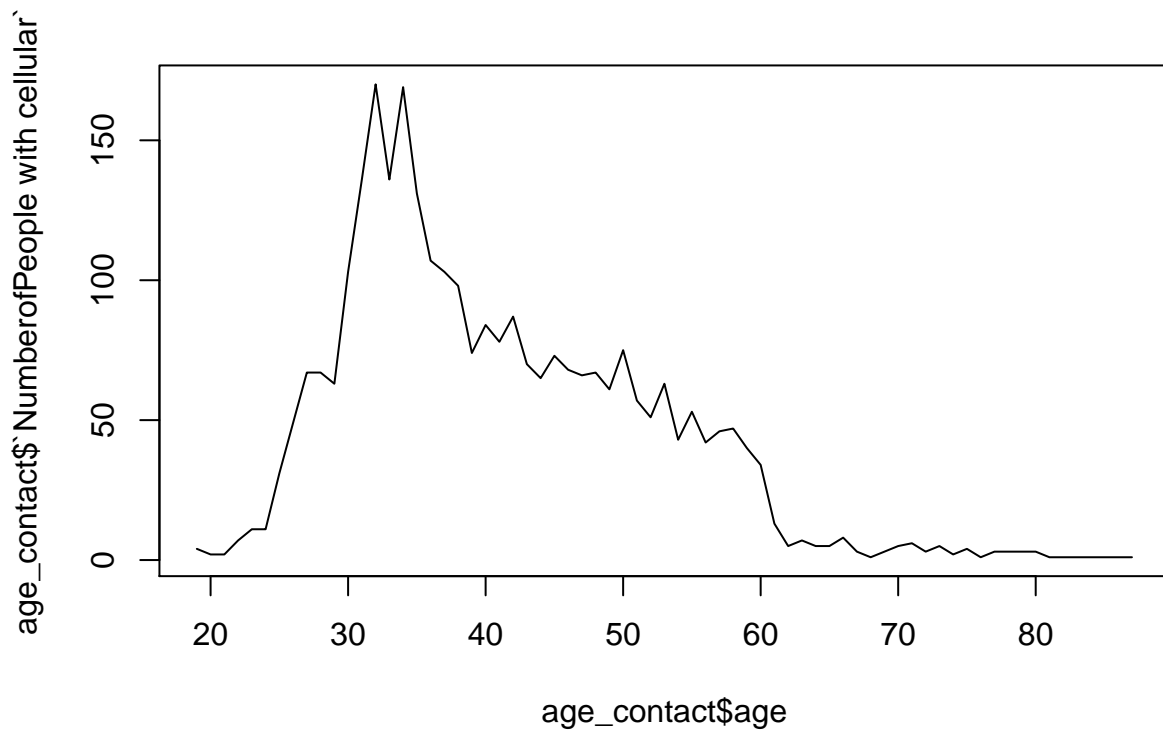
6)AVERAGE BANK BALANCE BASED ON AGE

Observation: Intuitively we feel people with higher age might be having good bank balance ,data reconfirms the same



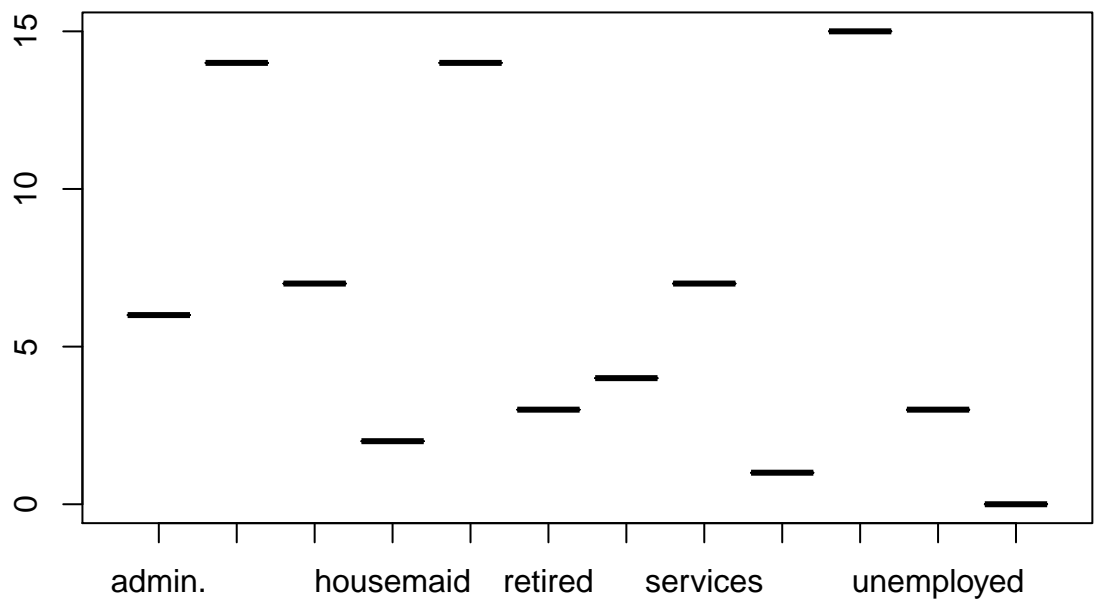
7)PLOT: PEOPLE WITH CELLULAR VS AGE

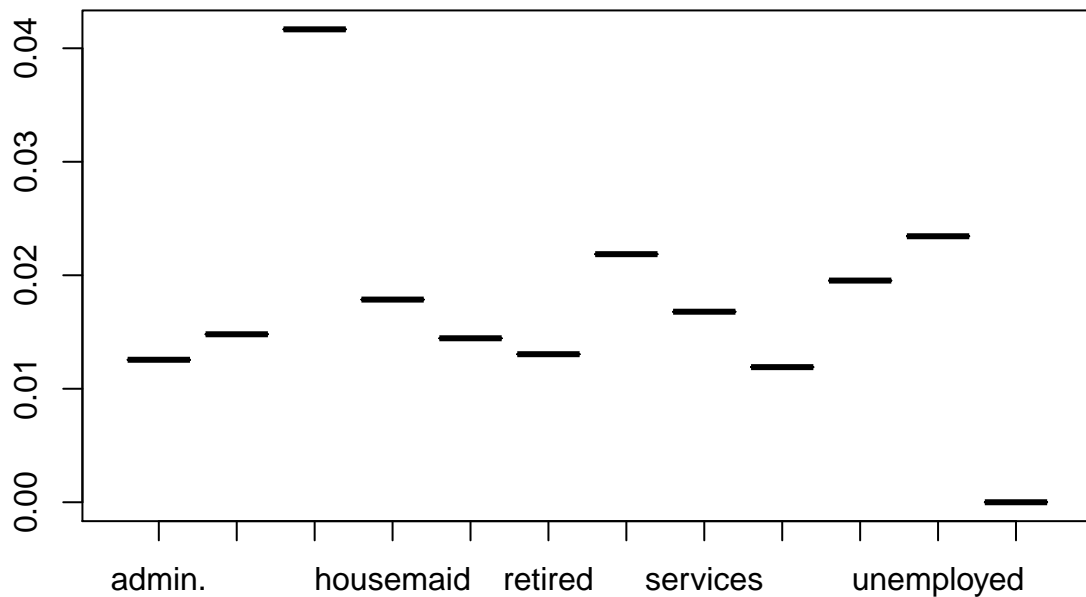
Observation:I see that mostly people in between 30 to 40 have cellular



8) ANALYSING DEFAULT COUNT BASED ON JOB PROFESSION

Following plots shows Default count based on job for given sample data and also percentage of default for each profession. observation-Entrepreneurs who takes more risk has highest default rate which is understandable.





9) Difference of balance between people having loans and not

Observation: Interestingly, I find huge difference interestingly between people having loans and those who doesn't from given sample. Anyways shouldn't jump into any concrete conclusion based on this sample

loan	sum(balance)
no	5798076
yes	633760

10) Which jobclients are contacted more by bank employees?

Observation: As a whole bank employees has contacted more Management, blue collar, technician.

