R Markdown assignment 3#

Abhishek Vijaykumar

November 26, 2018
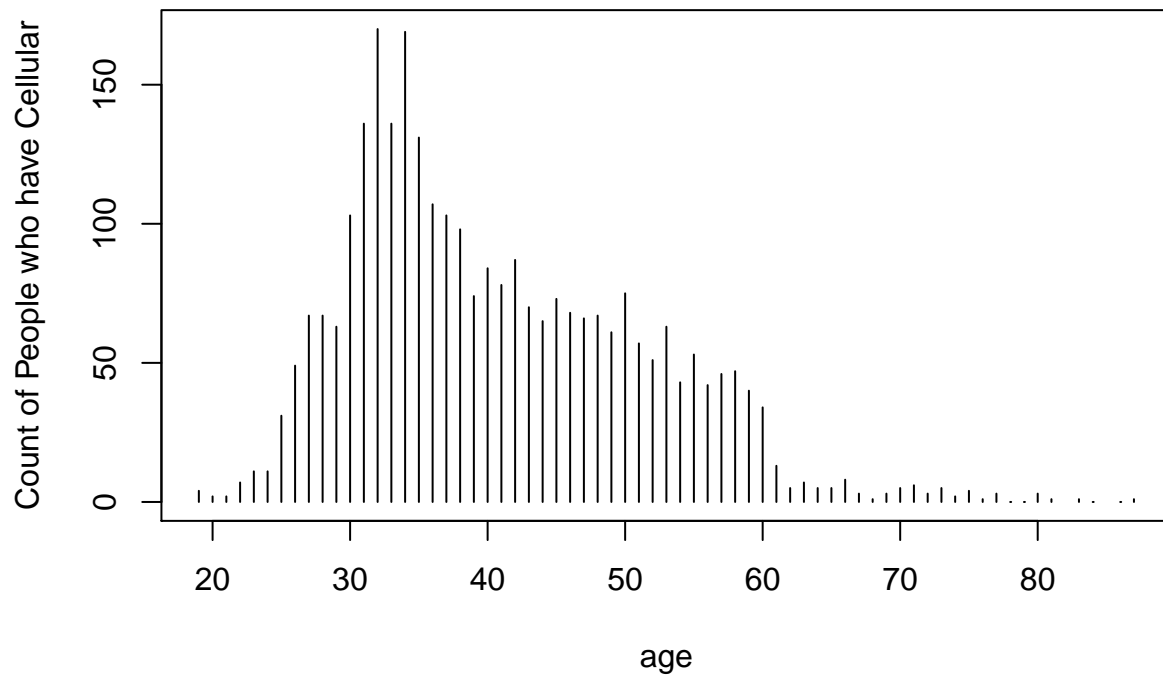
R Markdown

**Here are the ten insights into the Banking Data**

Including Plots

1) Distribution of People who have cellcular across each age

Observation - Mostly people between the ages of 30 and 40 have cellular


2) Average Balance of people based on profession

Observation - Retired people have the highest balance, believeable, but housemaids have the second highest? The next Query will try to find data outliers
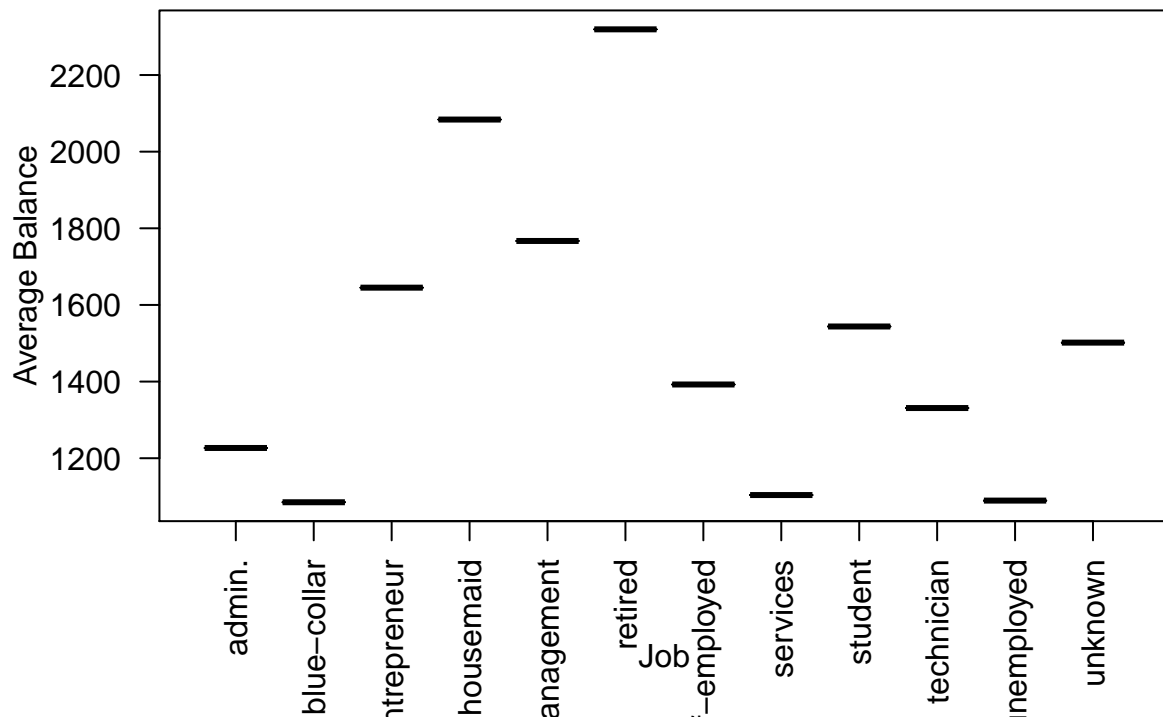
3) **Looking for Outliers in Bank balance in each profession by seeing the difference between mean and median for each profession**

Observation - The housemaid average balance has the highest deviation of mean with respect to median relative to any other job group
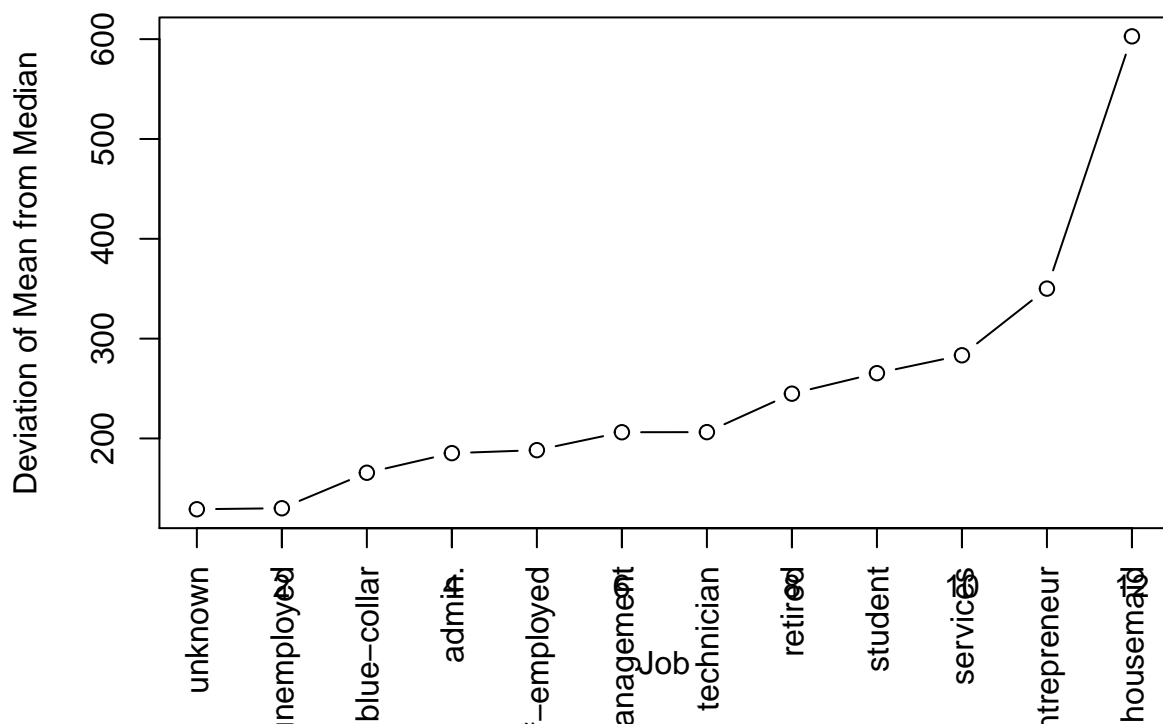
Interesting sidenote- The highest Bank balance for a housemaid is almost **27,000!** the sixth highest balance in the entire dataset!

This number is very suspicious for a single person with only a Primary level education, i'll leave the rest to the imagination

age	job	marital	education	balance
31	housemaid	single	primary	26965

4) Which employment category is most frequently contacted by bank employees

The highest is the "admin" group, the student is second highest, which makes sense as banks see younger people having a higher future potential
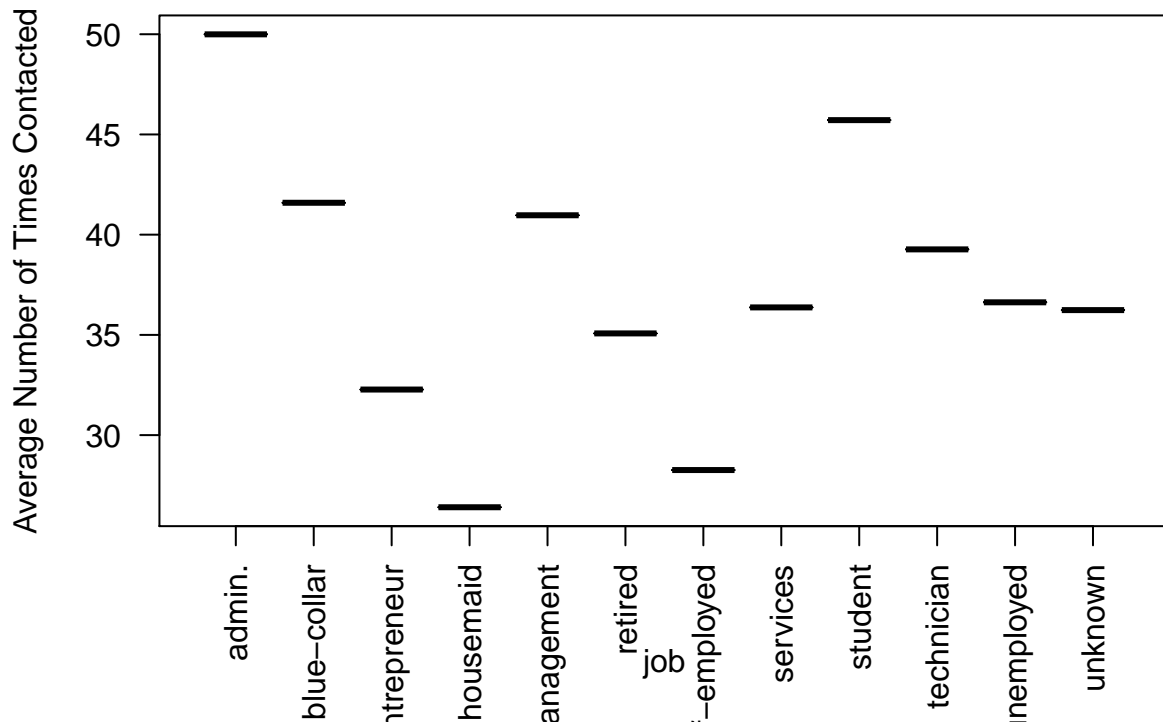
5) **We look at the average number of times each job group is contacted by the bank to see whether the campaigning is effective by checking the percent of those groups of people who have taken a housing/personal loan**

We do see a correlation between the number of calls a particular job group receives and the percentage of that group that has taken a loan

job	avgcontact	percentHousing	percentLoan
retired	168.3265	0.1224490	0.0408163
unknown	234.8333	0.1666667	0.1666667
student	177.3636	0.1818182	0.0000000
housemaid	169.5000	0.3888889	0.0000000
unemployed	228.3333	0.4761905	0.0476190
entrepreneur	242.0435	0.5652174	0.3043478
self-employed	190.2143	0.6071429	0.1071429
management	217.6398	0.6344086	0.1075269
admin.	224.6944	0.6666667	0.1388889
technician	218.3191	0.6666667	0.1773050
services	250.3548	0.7741935	0.0806452
blue-collar	264.0724	0.8289474	0.1315789

6) Which bucket of Call duration has the highest probability of getting a customer

Here we confirm that the longer a person is on a call, the more likely he is to agree to opening a term deposit (Column "y")

duration	percentyes
1-10 Seconds	0.0000000
11-20 Seconds	0.0000000
21-30 Seconds	0.0000000
41-50 Seconds	0.0000000
50s-1minutes	0.0000000
31-40 Seconds	0.0147059
1-2minutes	0.0235690
2-3minutes	0.0408654
3-4minutes	0.0848000
4-5minutes	0.1288056
5-10minutes	0.1969158
10-15minutes	0.3992674
15-20minutes	0.5157895
20-30minutes	0.5961538
>30minutes	0.6666667

7) Cost of Marriage!

Here we see the effect of bank balance on married people grouped by their job (Marriage is definetly expensive, Divorce is even worse!)

job	marital	avg_bal
admin.	divorced	977.8116
admin.	married	1353.2669
admin.	single	1111.4825
blue-collar	divorced	744.7975
blue-collar	married	1064.3608
blue-collar	single	1322.5402
entrepreneur	divorced	889.7500
entrepreneur	married	1612.8258
entrepreneur	single	2462.6000
housemaid	divorced	1166.2308
housemaid	married	1759.4643
housemaid	single	4695.3333
management	divorced	1432.3529
management	married	1858.0826
management	single	1729.5290
retired	divorced	1771.4651
retired	married	2383.8580
retired	single	3425.6364
self-employed	divorced	2572.0667
self-employed	married	1350.9449
self-employed	single	1089.2683
services	divorced	931.5645
services	married	1228.0254
services	single	947.7227
student	married	468.3000
student	single	1689.1622
technician	divorced	845.7865
technician	married	1436.9197
technician	single	1329.6866
unemployed	divorced	842.9091
unemployed	married	1131.5867
unemployed	single	1162.3548
unknown	divorced	137.0000
unknown	married	1275.9333
unknown	single	2664.2857

**8) Which age group is most likely to open a term deposit account/ Which age group should the bank target?**

Here we see that the target group should be people over the age of 60

age	percentyes
<24	0.2093023
24-28	0.1464435
28-32	0.1129326
32-36	0.1108404
36-40	0.0971787
40-44	0.0863039
44-48	0.1148649
48-52	0.0980861
52-56	0.1260997
56-60	0.0809969
60-70	0.2743363
70-80	0.3829787
80-90	0.4285714

9) Percentage of defaults according to profession

Here we see that the profession group that defaults the most is entrepreneurs
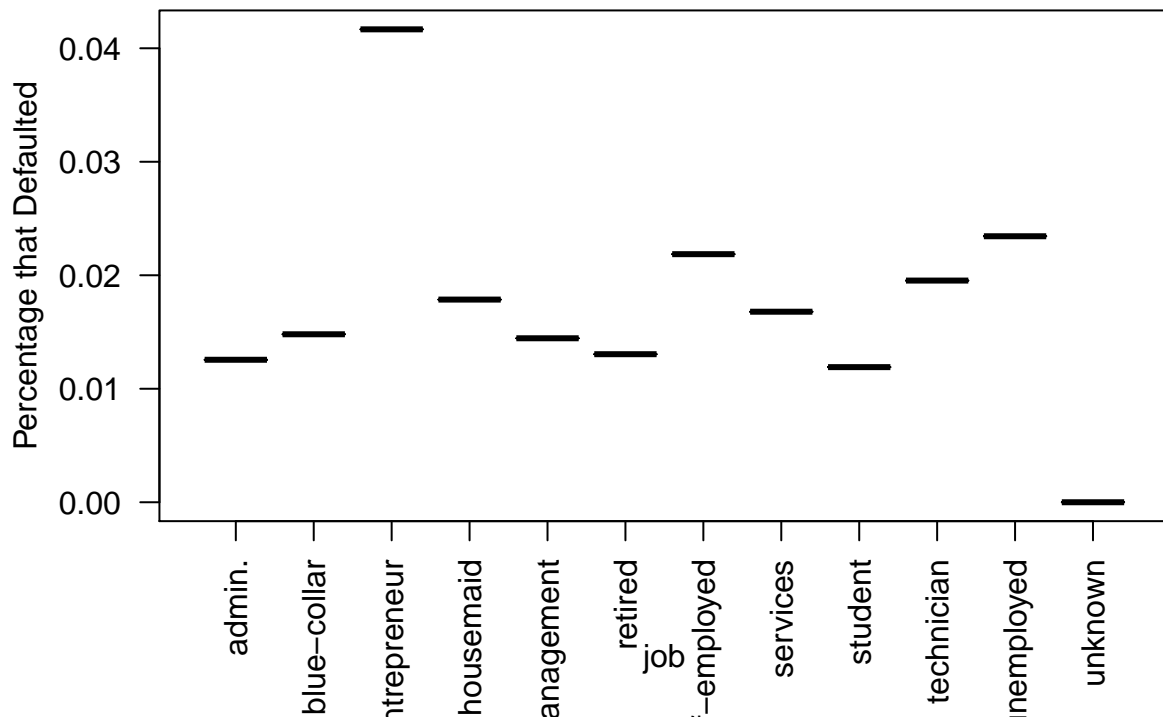
10) Does having a higher education make you richer, in each profession

We see that this is not necessarily true! professions such as the technician and management defy this. We also see this in retired people, indicating that by the time you get that old, your education wont make much difference. Not something I believe in but this is what the data says. Education does seem to matter if you're unemployed though!

job	education	avg_bal
admin.	primary	390.5882
admin.	secondary	1269.6794
admin.	tertiary	1053.2941
blue-collar	primary	1072.2141
blue-collar	secondary	1068.5916
blue-collar	tertiary	2385.5000
entrepreneur	primary	383.9231
entrepreneur	secondary	1276.1724
entrepreneur	tertiary	2585.9041
housemaid	primary	1807.1053
housemaid	secondary	2011.8929
housemaid	tertiary	2392.5455
management	primary	2685.4103
management	secondary	1250.1034
management	tertiary	1776.3431
retired	primary	2744.6000
retired	secondary	2089.0952
retired	tertiary	2476.7419
self-employed	primary	1471.7333
self-employed	secondary	1164.5526
self-employed	tertiary	1615.9659
services	primary	1107.3200
services	secondary	998.8843
services	tertiary	1894.8750
student	primary	1787.5000
student	secondary	1610.4255
student	tertiary	1175.6842
technician	primary	2593.0000
technician	secondary	1153.6077
technician	tertiary	1631.6303
unemployed	primary	873.1923
unemployed	secondary	1025.1618
unemployed	tertiary	1224.7812
unknown	primary	360.2857
unknown	secondary	1229.0000
unknown	tertiary	2497.7500