

R Markdown assignment 4#

Abhishek Vijaykumar

December 2nd, 2018

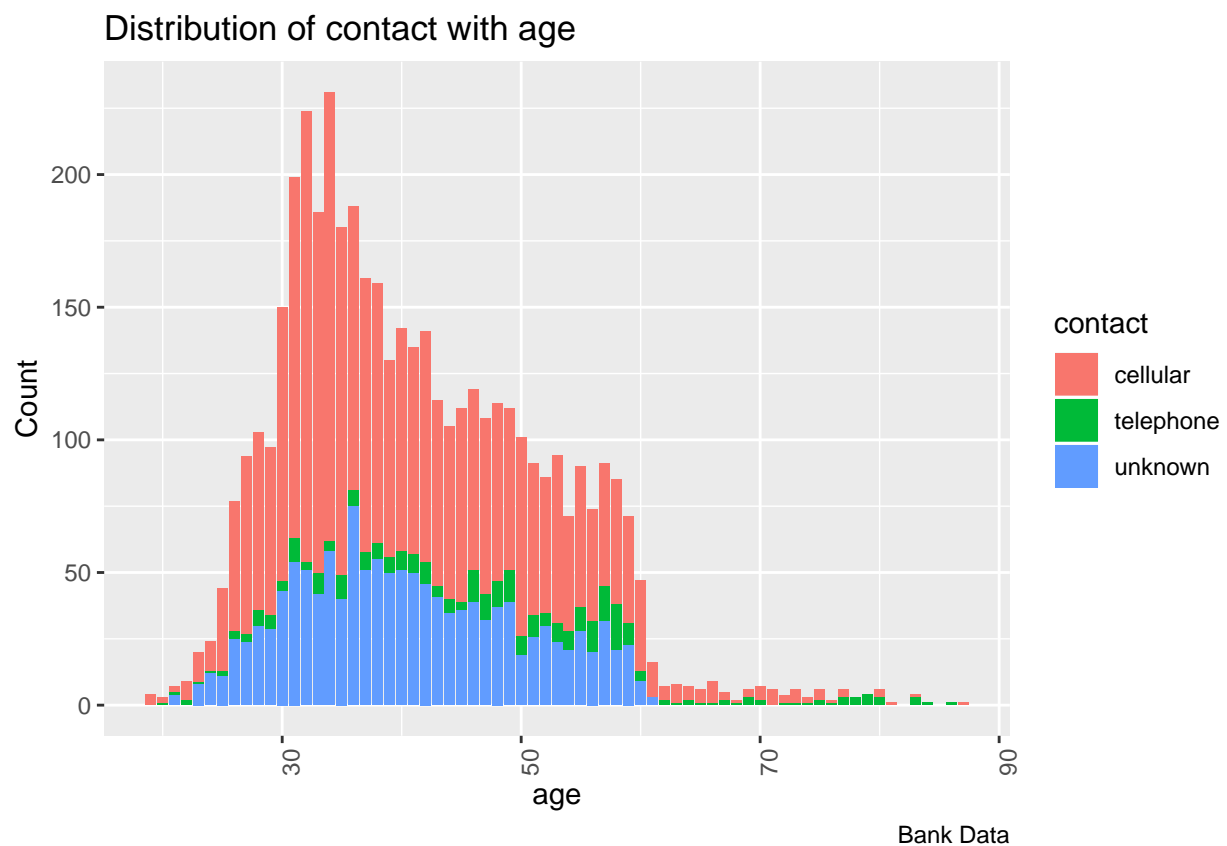
R Markdown

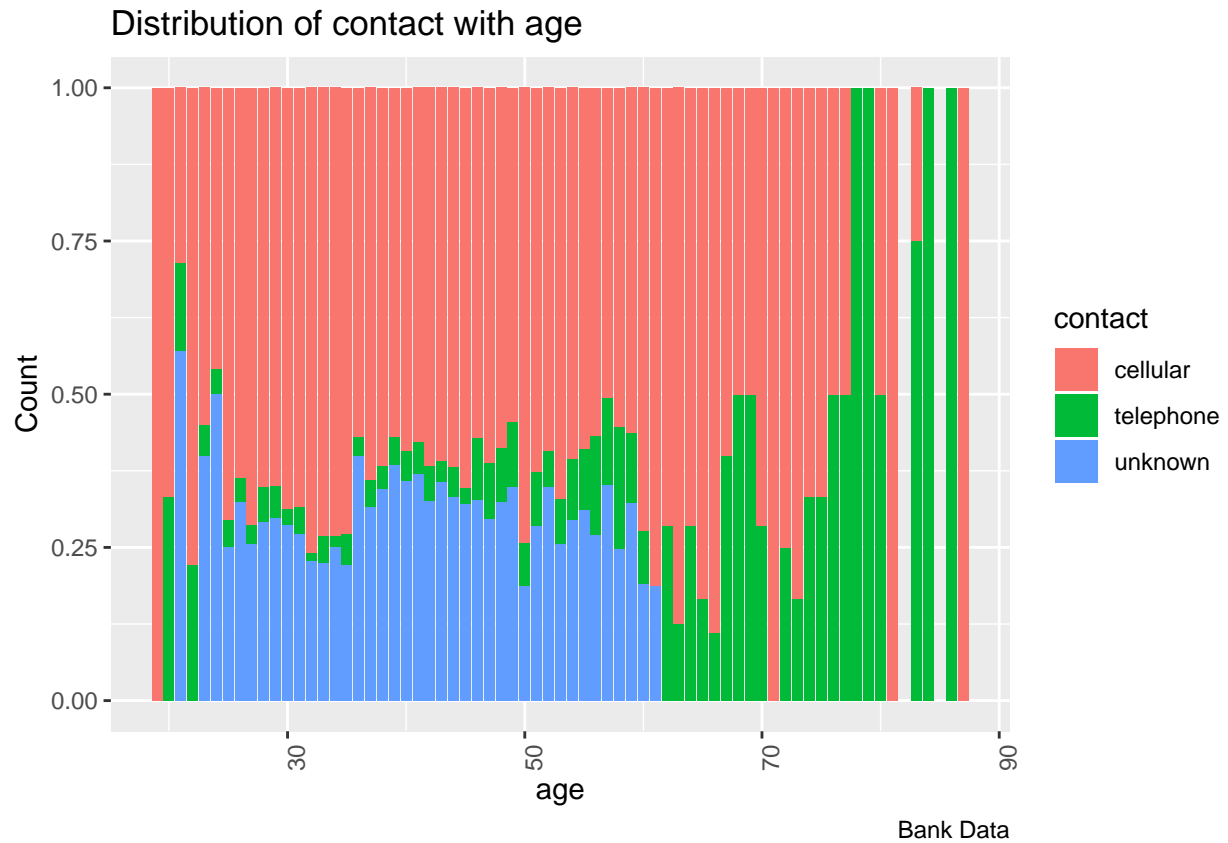
Here are the ten insights into the Banking Data using GGPILOT

Including Plots

1) Distribution of contact method across age

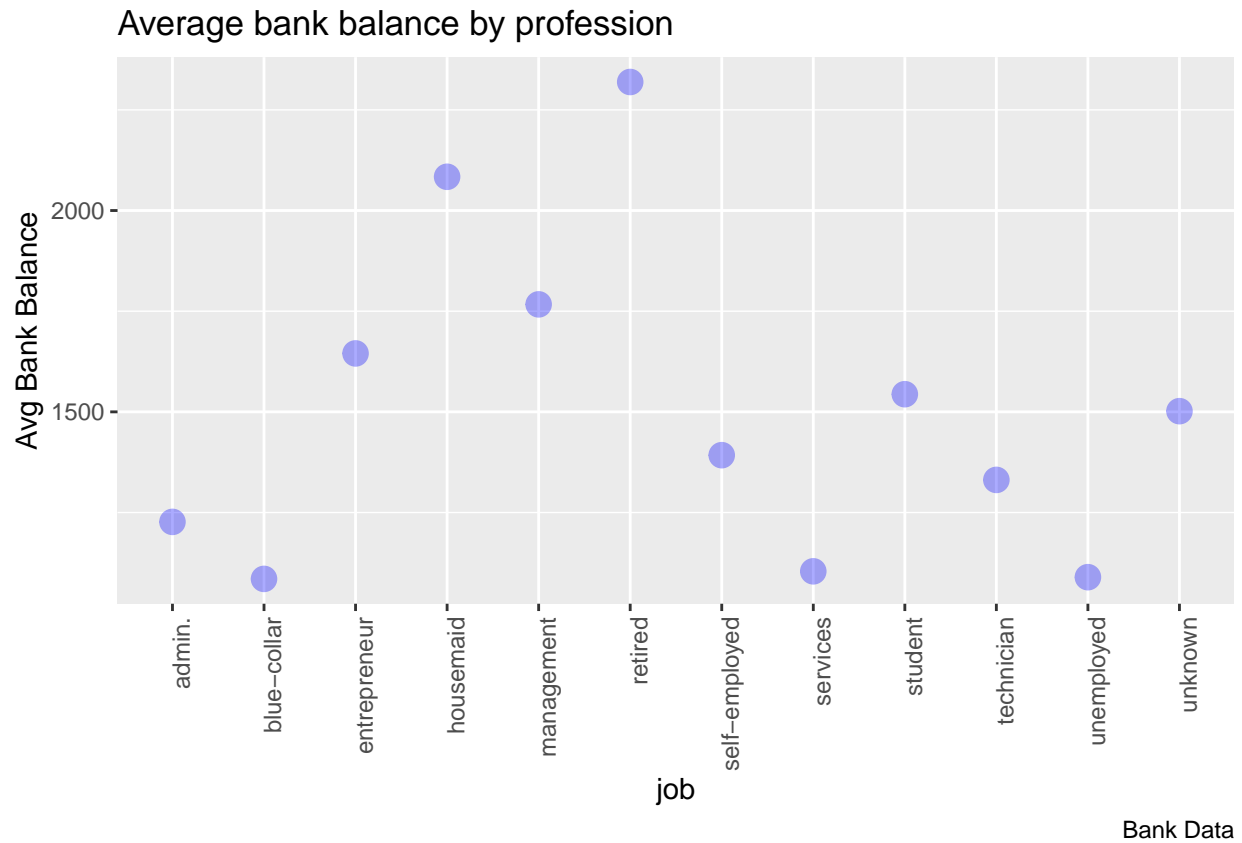
Observation - Mostly people below the age of 60 have cellphones, the older generations mostly still use telephones





2) Average Balance of people based on profession

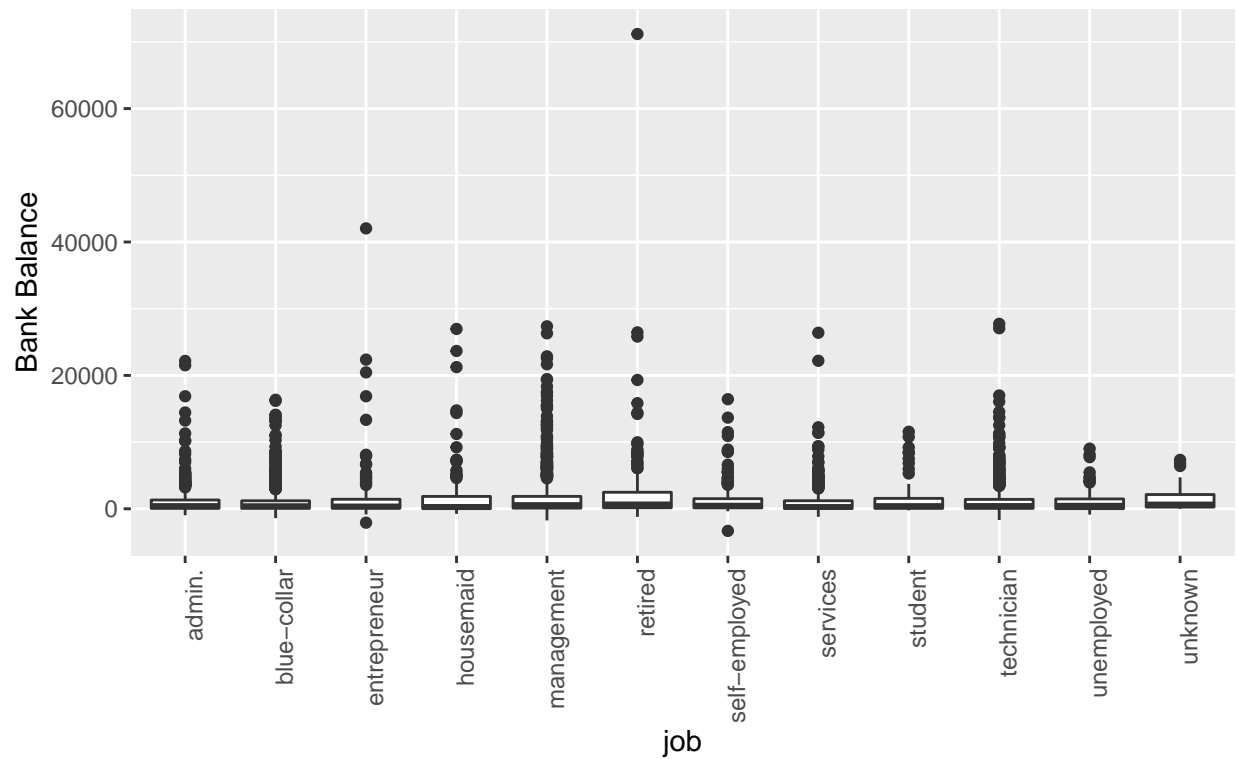
Observation - Retired people have the highest balance, believable, but housemaids have the second highest? The next Query will try to findout which profession has data points that deviate from the median



3) Looking for Outliers in Bank balance in each profession by examining the boxplot for the data

The two outliers present under “retired” and “entrepreneur” make it difficult to look at the box plot

BoxPlot for average bank balance by profession

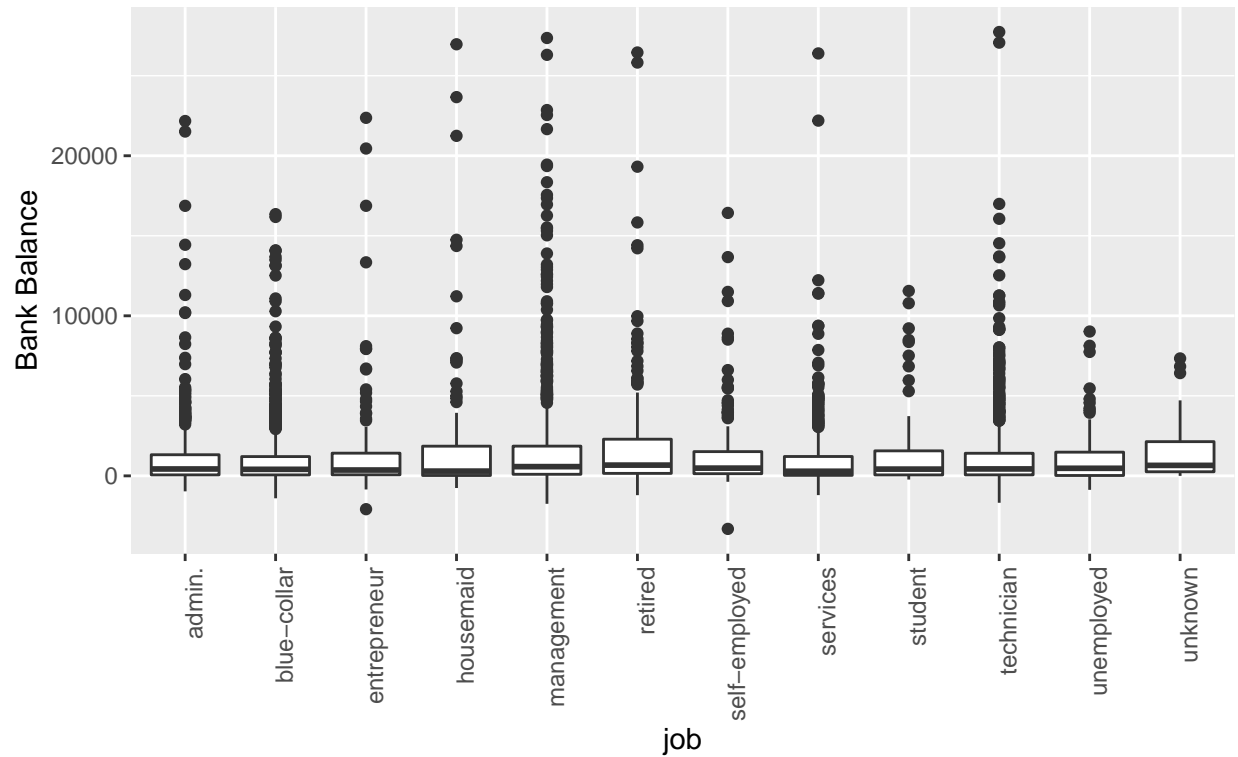


Bank Data

Once these are removed, we are able to clearly see how the “housemaid” and “retired” profession category is being impacted by the presence of outliers

Warning: package 'bindrcpp' was built under R version 3.5.1

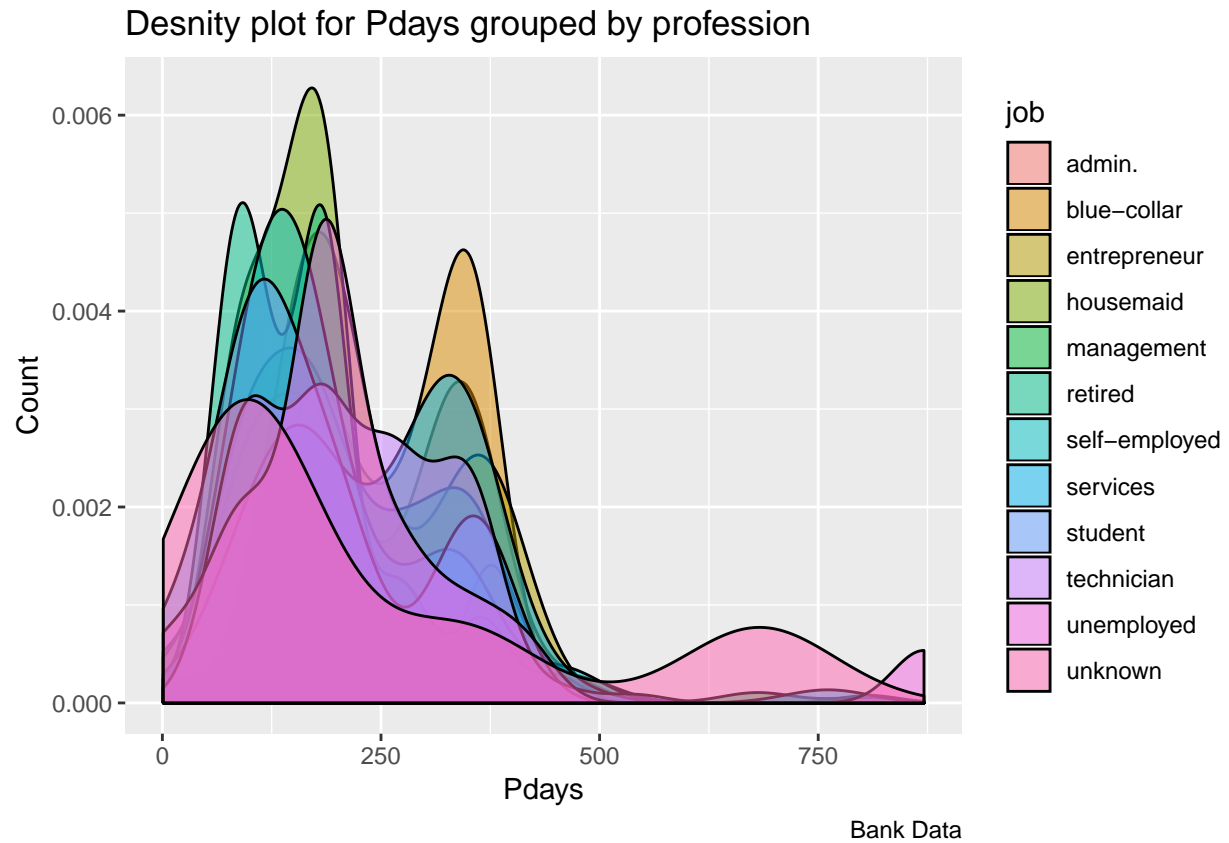
BoxPlot for average bank balance by profession



Bank Data

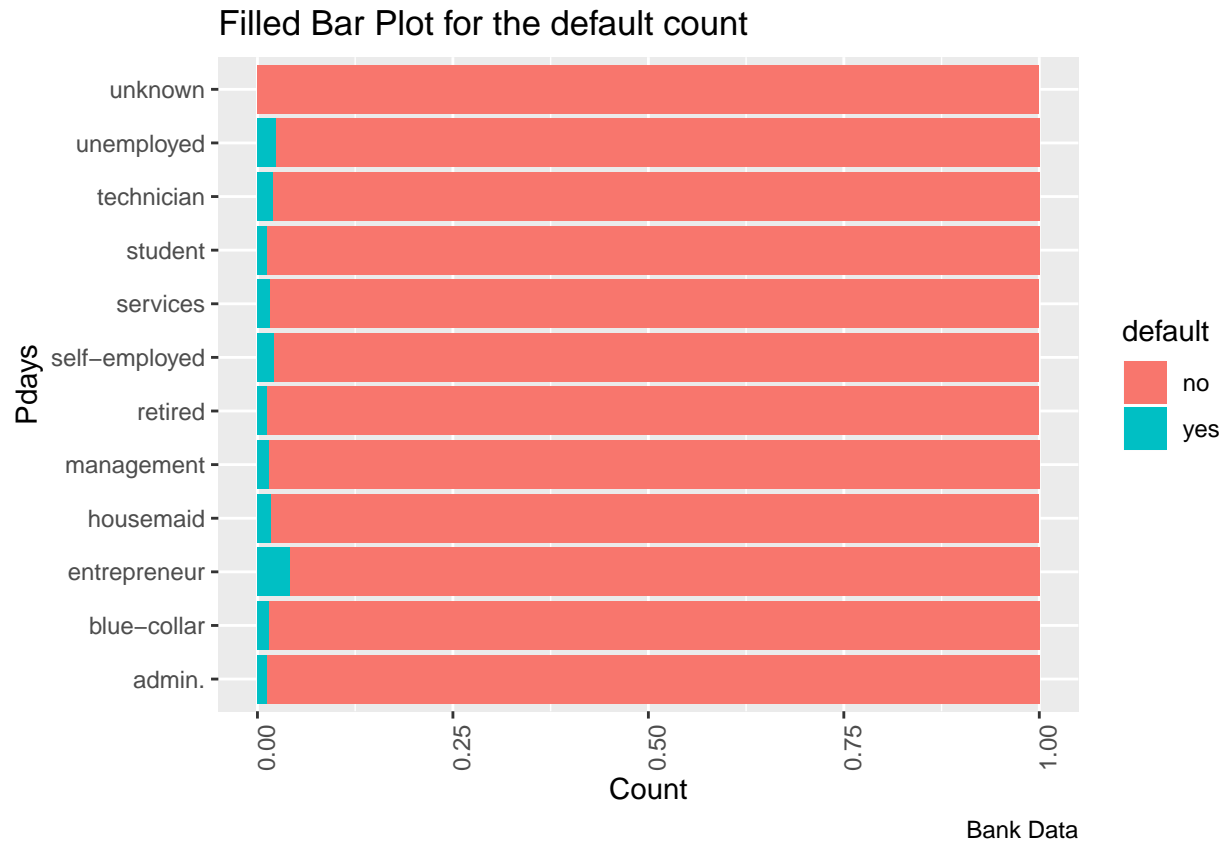
4) Which employment category is most frequently contacted by bank employees

“Pdays” is the days since the last time the person was contacted, the more left skewed the distribution, the more number of times that particular profession group was contacted, from this we see that entrepreneur’s are amongst the most frequently while unemployed people are amongst the least



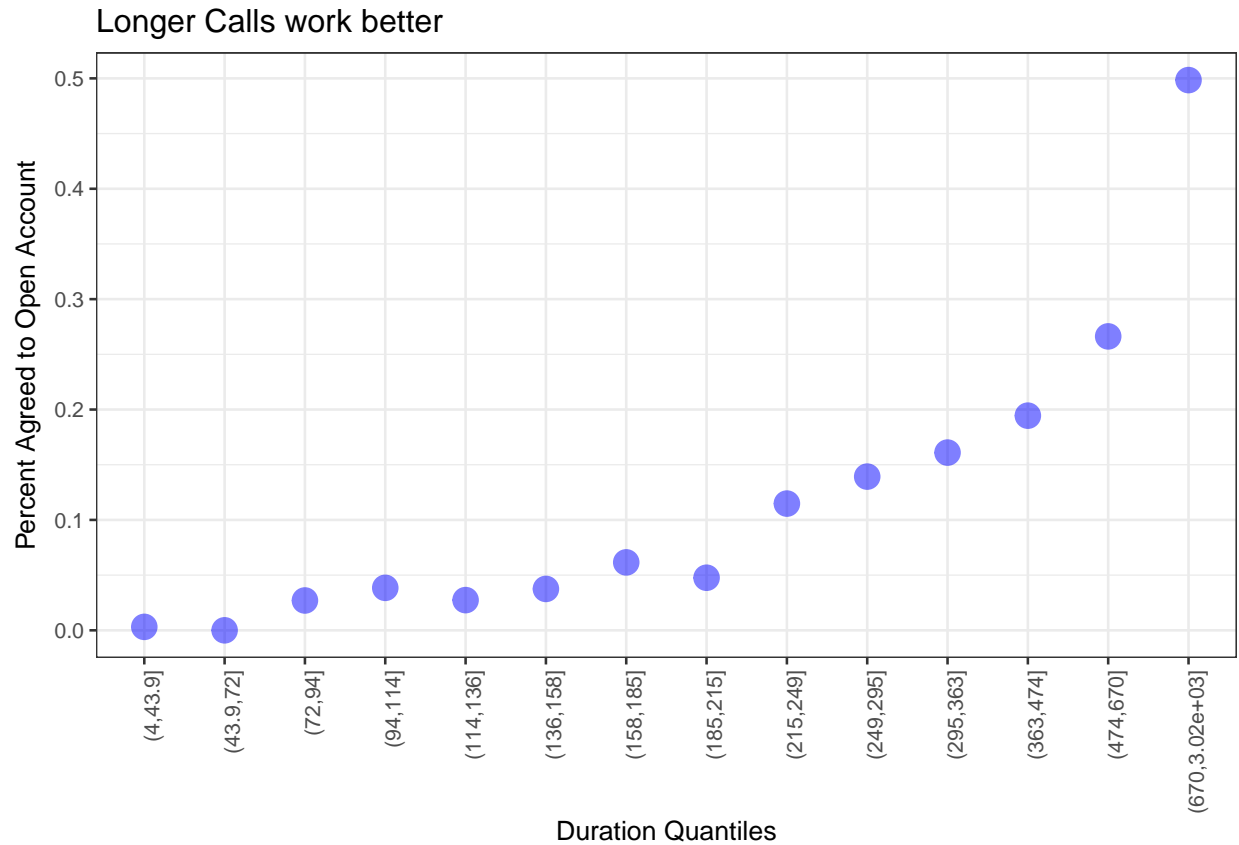
9) Proportion of defaults according to profession

Here we see that the profession group that defaults the most is entrepreneurs



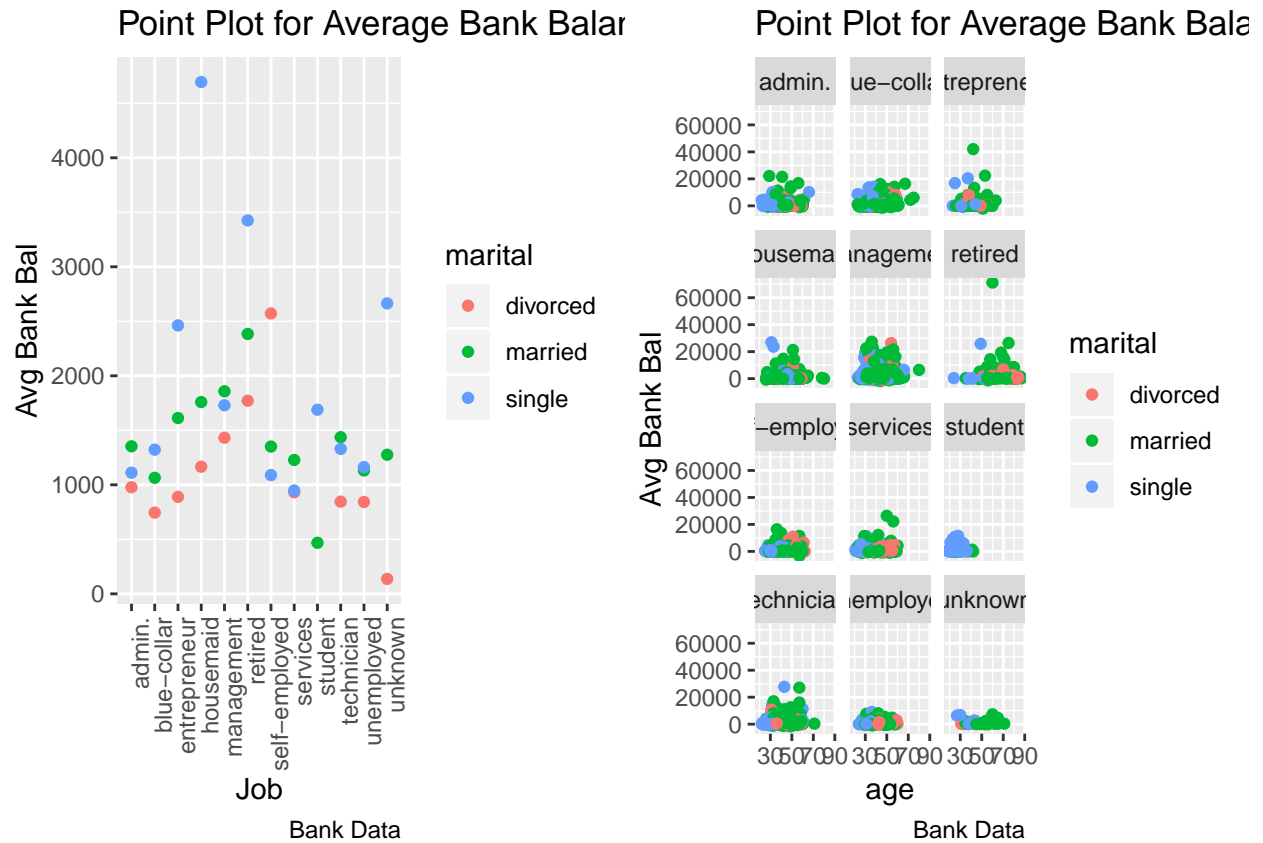
6) Which bucket of Call duration has the highest probability of getting a customer

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”)



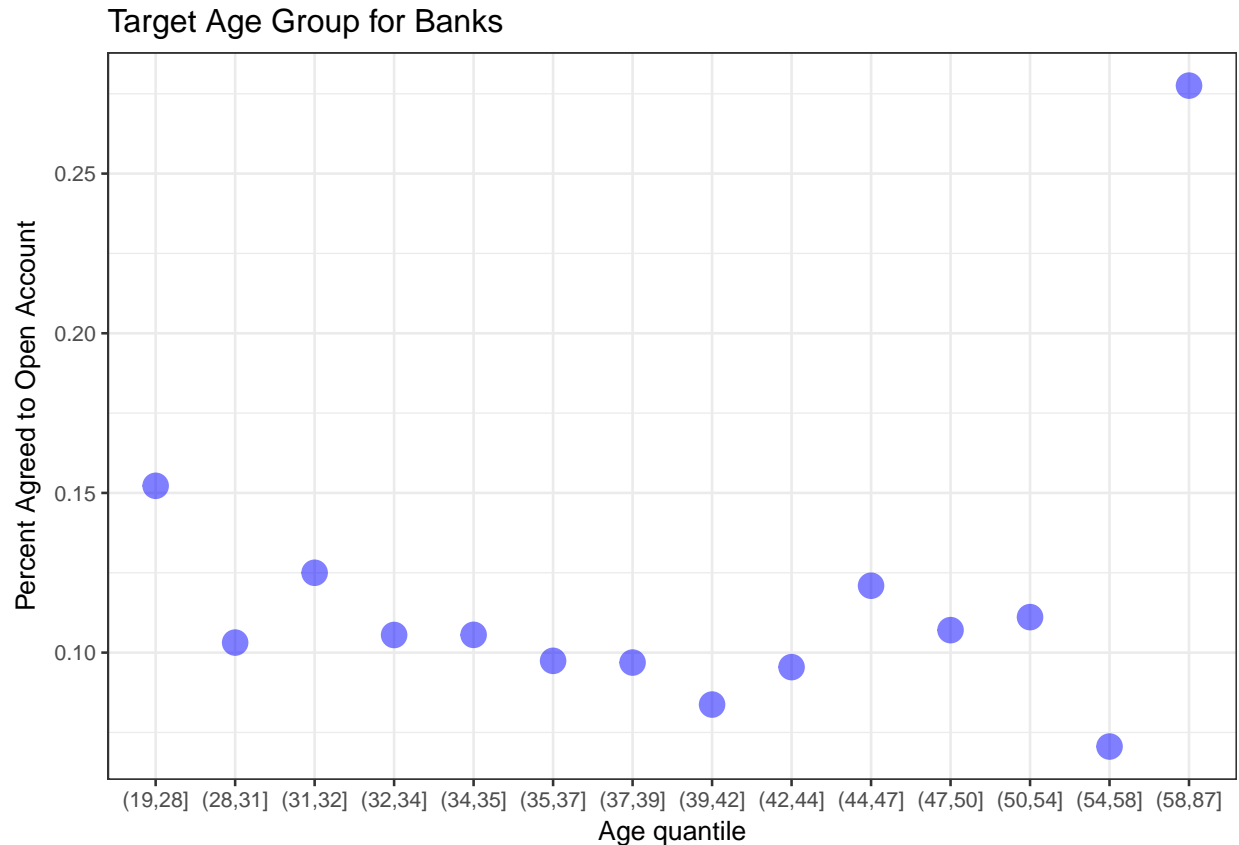
7) Cost of Marriage!

Here we see the effect of bank balance on married people grouped by their job. The plot on the left suggests that marriage is quite expensive for most professions, divorce even more so! The plot on the right however, does not show any change in bank balance induced by marital status



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

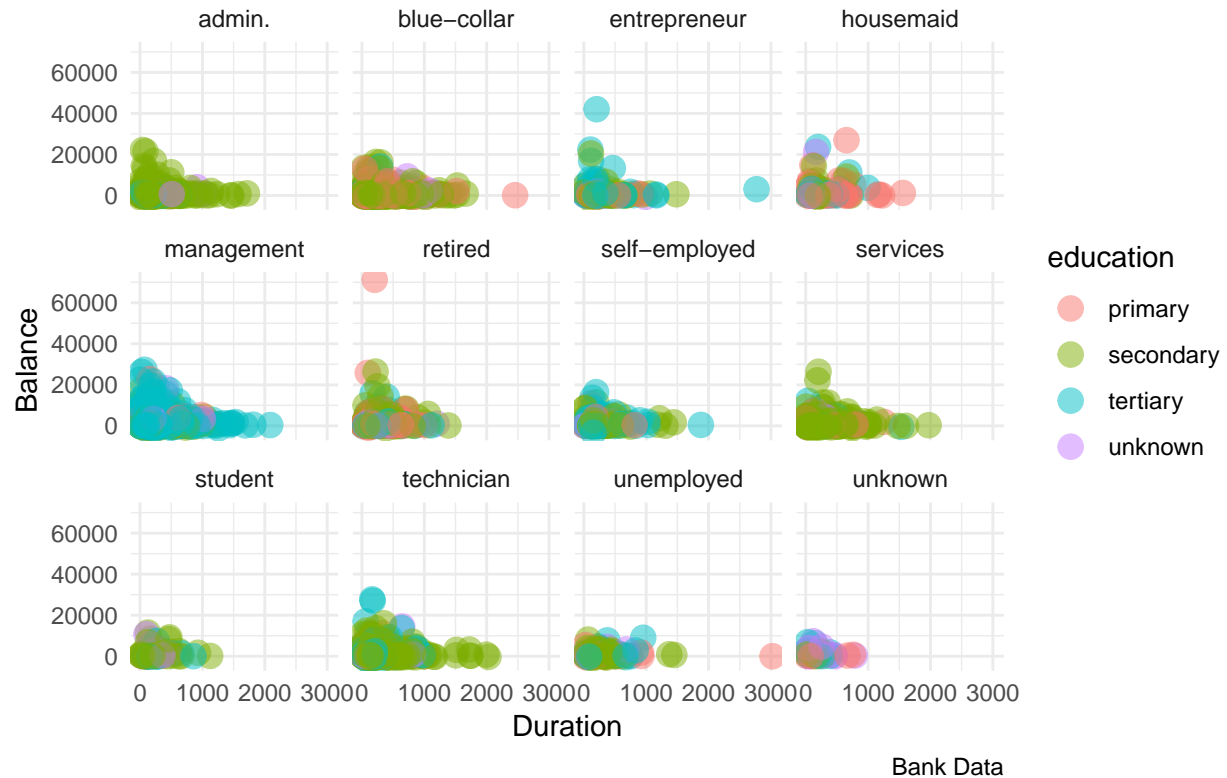
Here we see that the target group should be people over the age bucket of 58-87 are most likely to open a term deposit account with the bank



9) Does having a higher education make you richer, in each profession ?

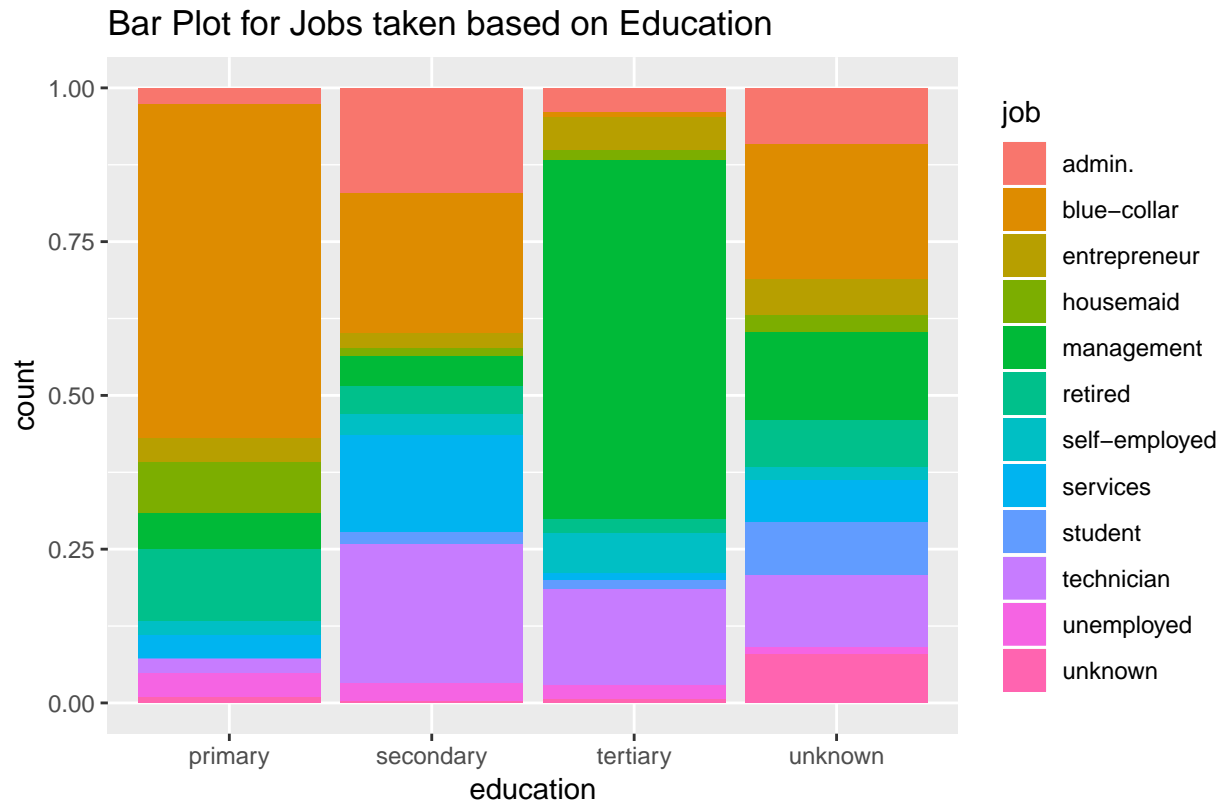
We do not see many blue spots rising above the others in many cases, we do see a few in the entrepreneur profession, but not too many. We can say that education does not make you richer (according to the data!). We also have a chance to see how the bank balance varies by profession, no particular group seems to be dominating the roster here. (note- turns out the person who spoke to the bank employee for 3000 seconds was infact unemployed, not surprising!)

Point Plot for Education against Bank Balance



10) How does Education affect the job you end up taking?

We plot two graphs, both for the same end purpose. We see clearly that people with a higher level education end up in managerial roles, the profession of the secondary education group seems to be distributed uniformly, and the people with a primary level education end up in the blue collar roles



Bank Data

Polar Coordinate Plot for Jobs taken based on Education

