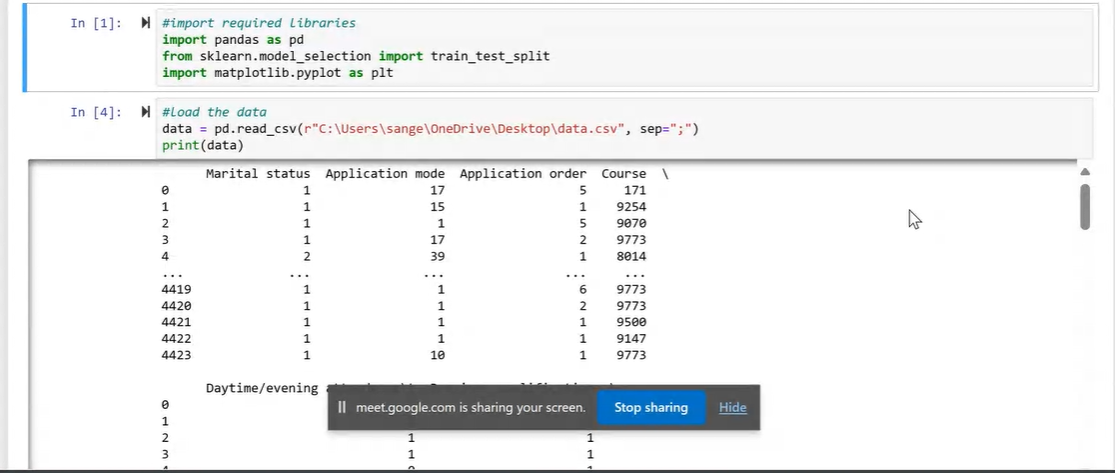
Train – 90% test -10% most efficiency use train\_test\_split

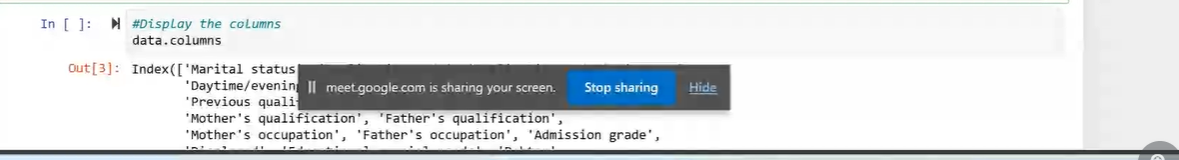
matplot.lib for visualize in grp

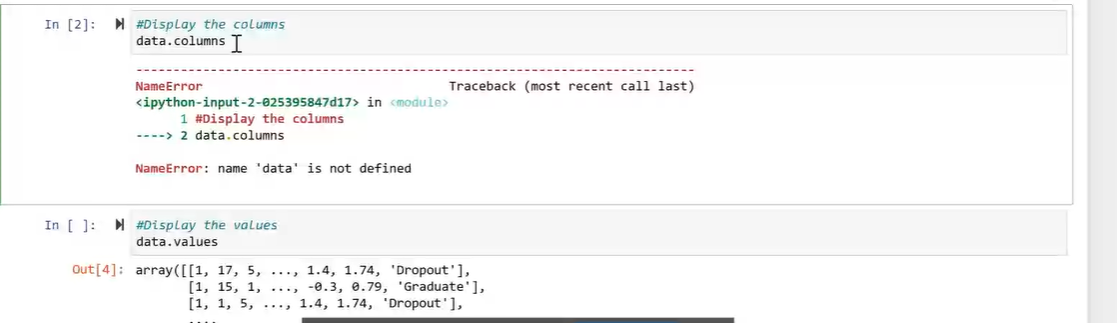
csv file – read\_csv

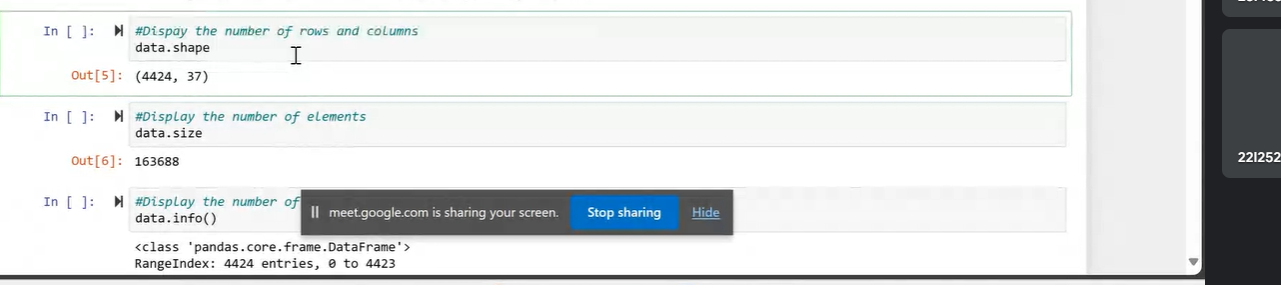
xml – read\_xml

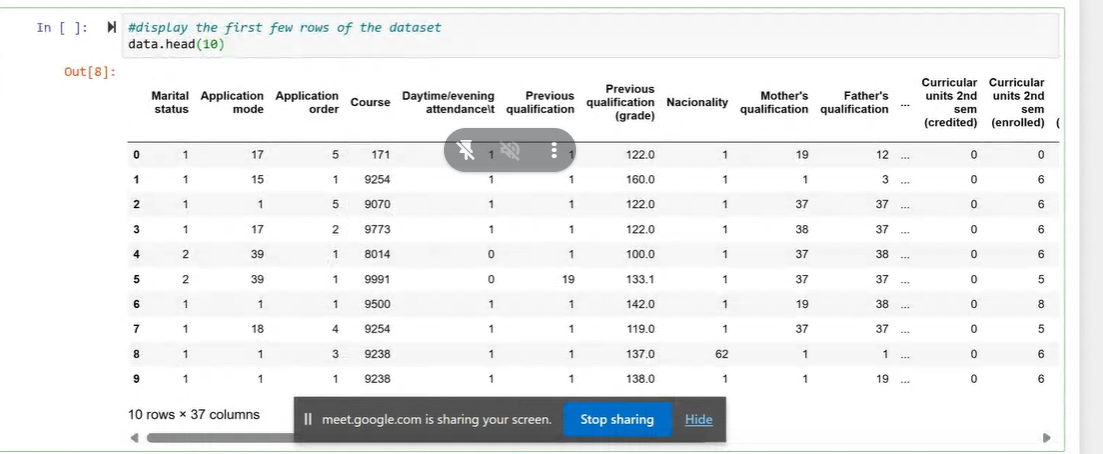
json – read\_json

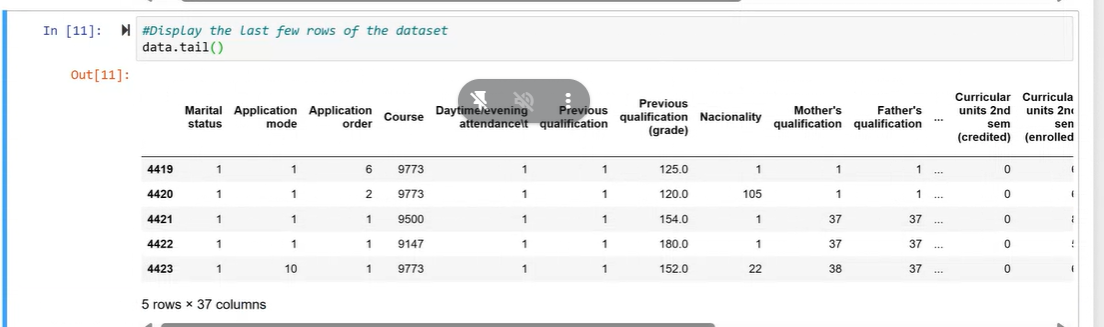


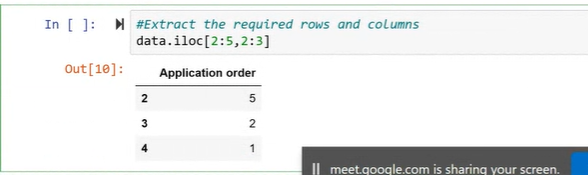


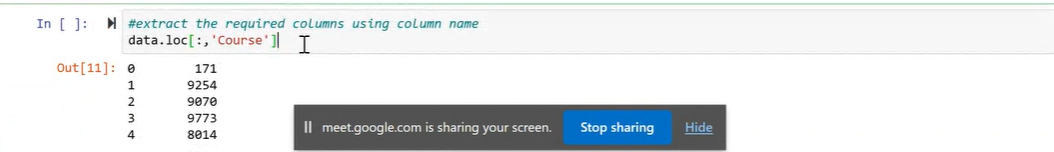


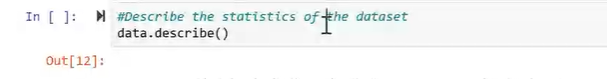


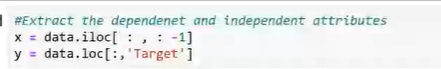








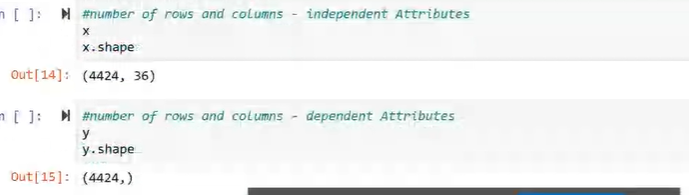




[::,:-1] EMPTY MEANS ALL

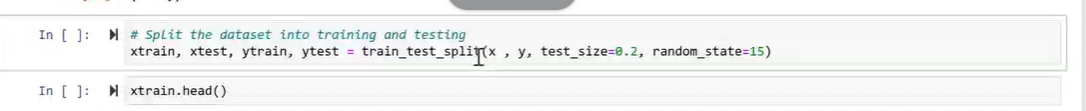
-1 means total -1

Target is the col name



y.shape

(value, ) means



test\_size or train\_size can define any one for the partition of test and train

random selection – random\_state

divide into x (ind) , y(dep) , test\_size , train\_size



axis 1 col or axis is 0 row value

how- any none then delete

how – all only if all none then delete

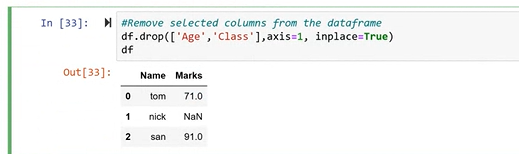
inplace = true – del data completely even when we have slized and kept in another variable

insert:

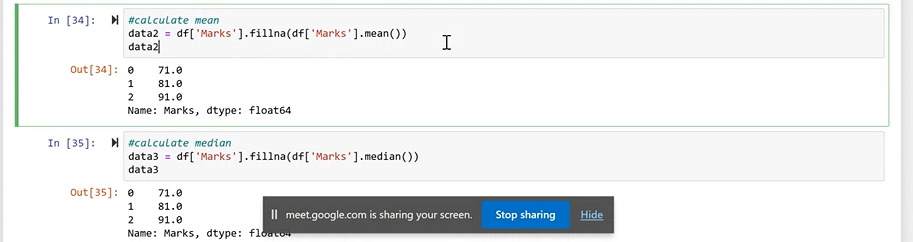
(index,name of col, values for all rows, tell if duplication is possible or not{if true multiple values possible})







If the null values with mode,mean and median





To determine how to fill the null values

