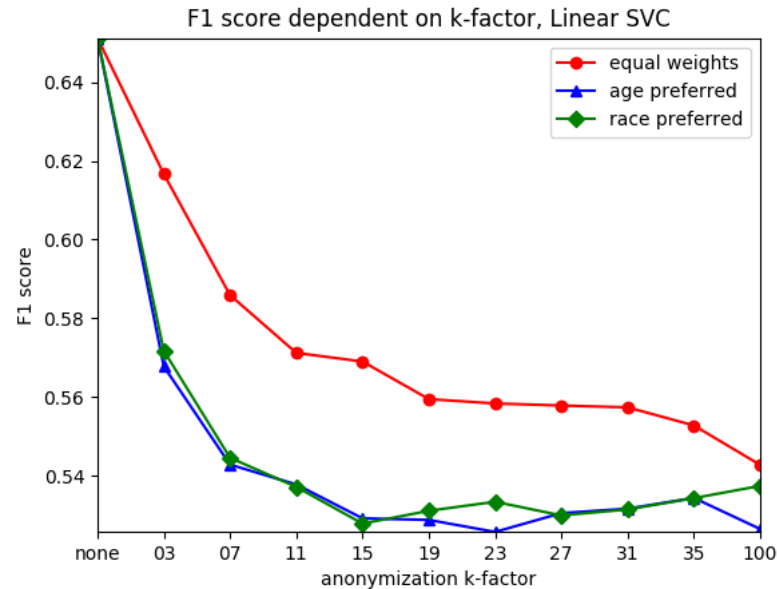
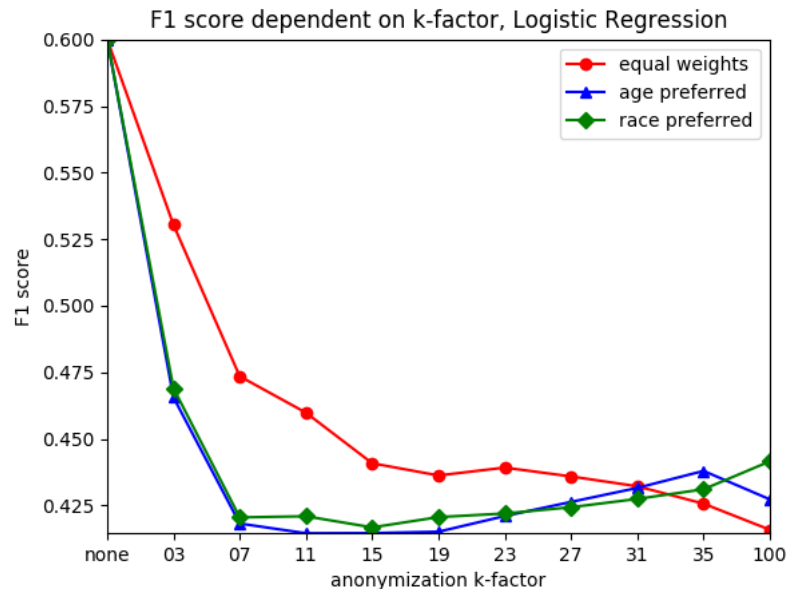
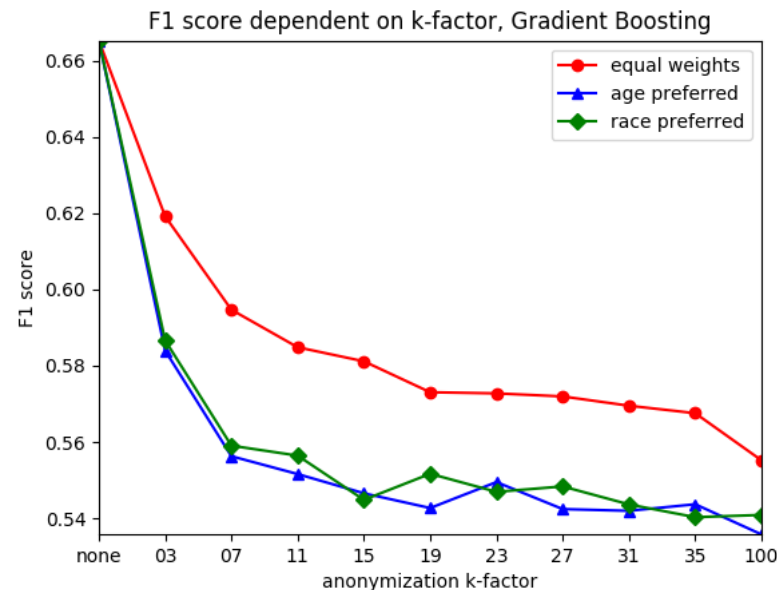
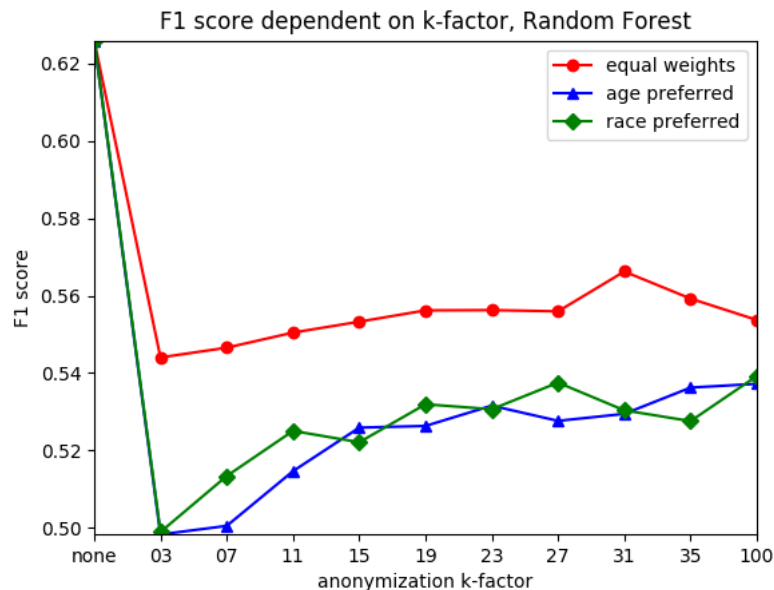


MC on anonymized data - education num

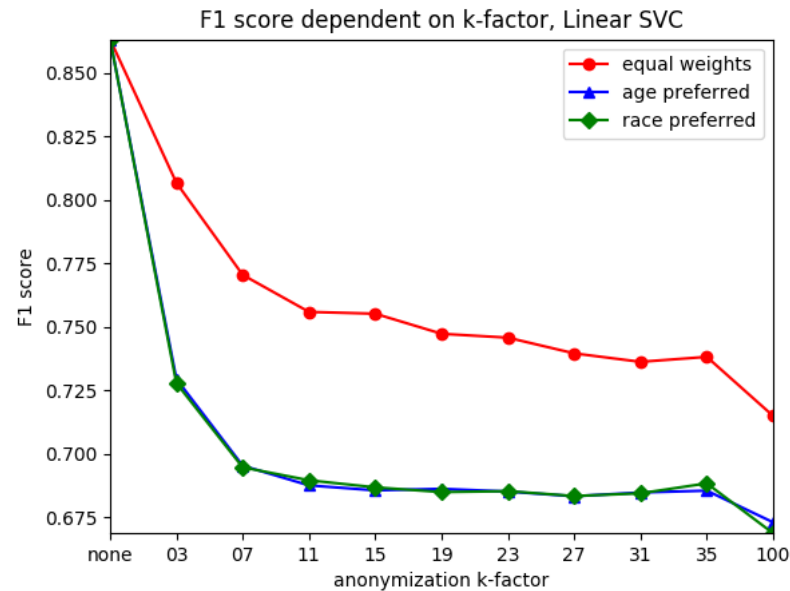
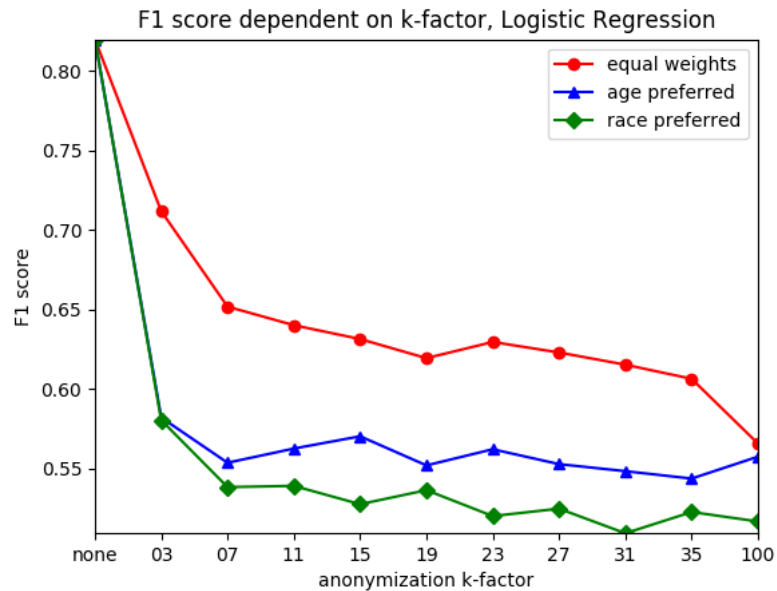


-) results almost as expected
-) same shape as with binary classification

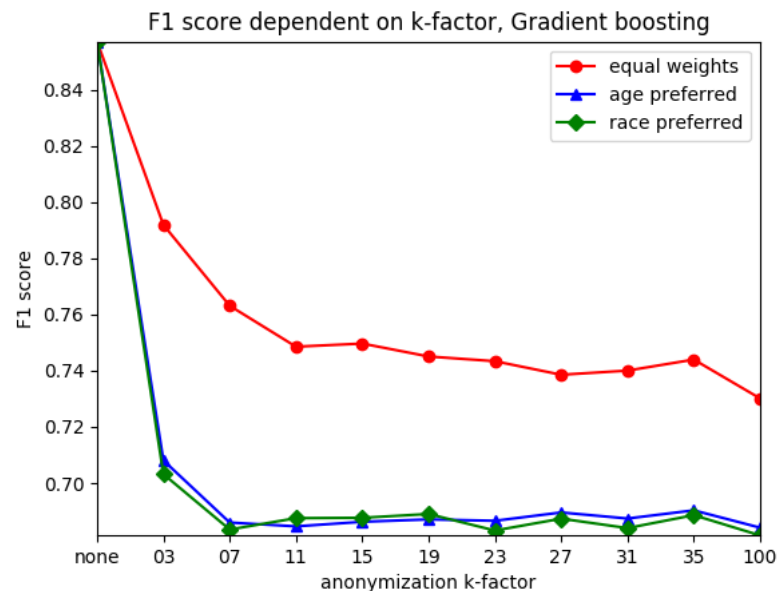
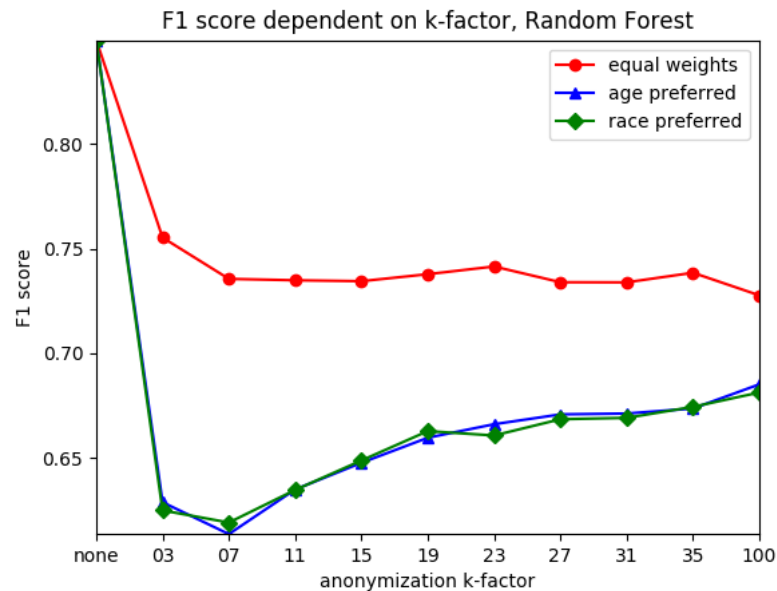


-) RF (and to slight degree LR & LinSVC) show improvement with increasing k... ?????

MC on anonymized data - marital status

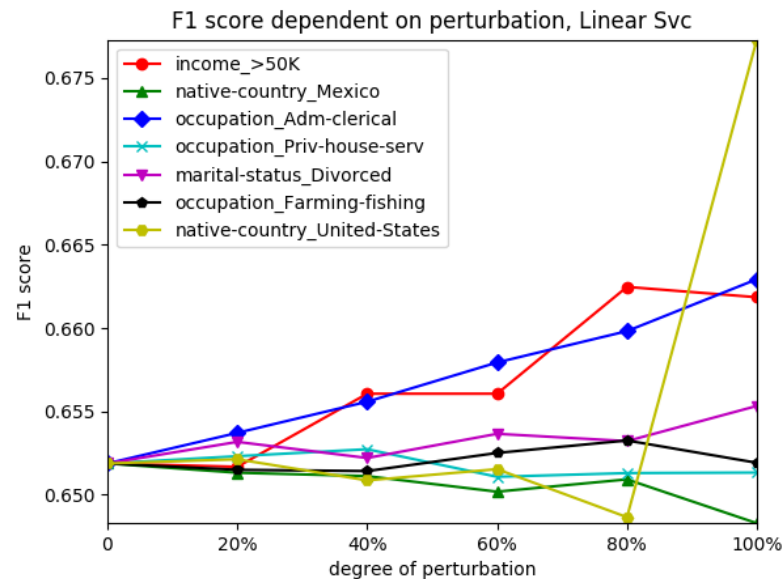
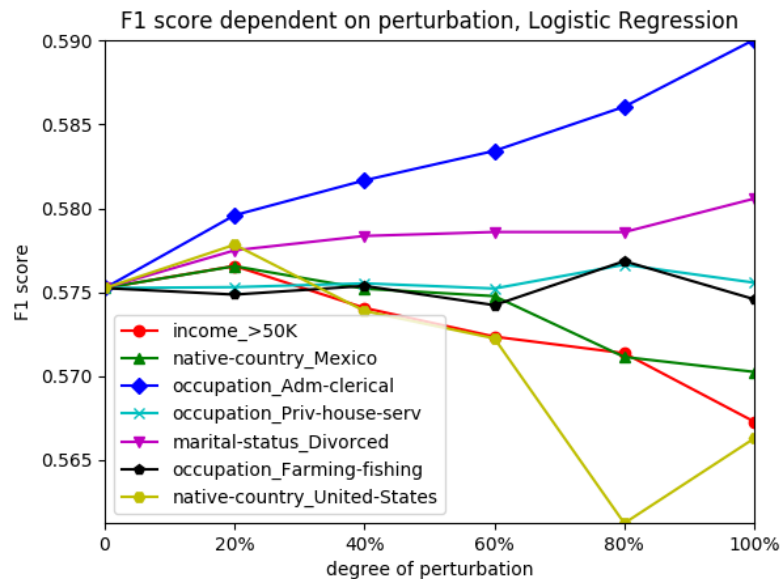


-) results almost as expected
-) same shape as with binary classification

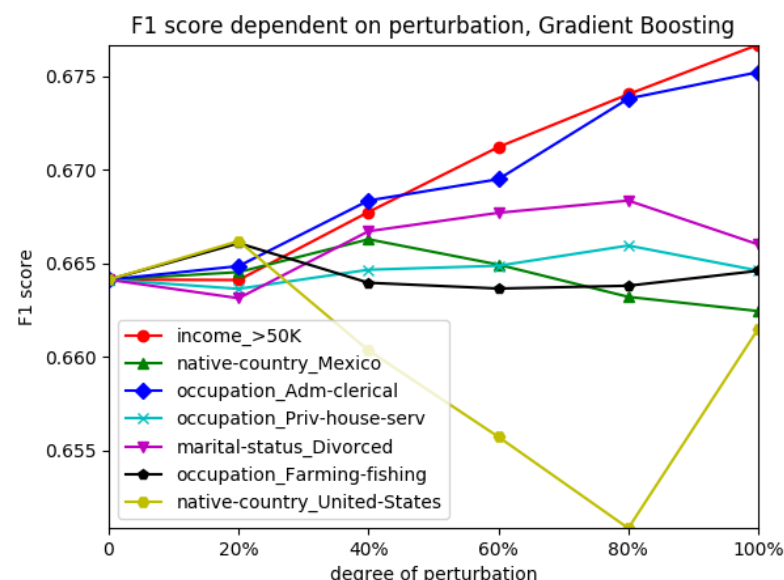
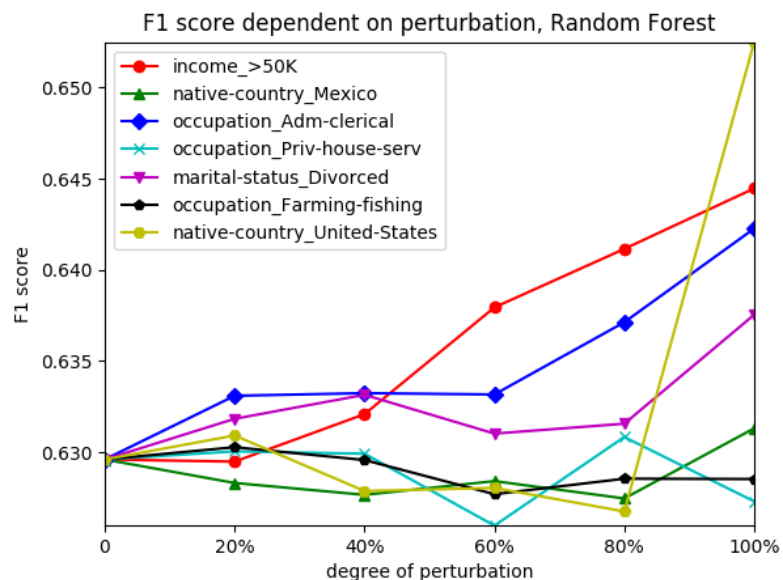


-) RF shows improvement with increasing k...?
-) GB stays practically stable?

MC on perturbed data - education_num

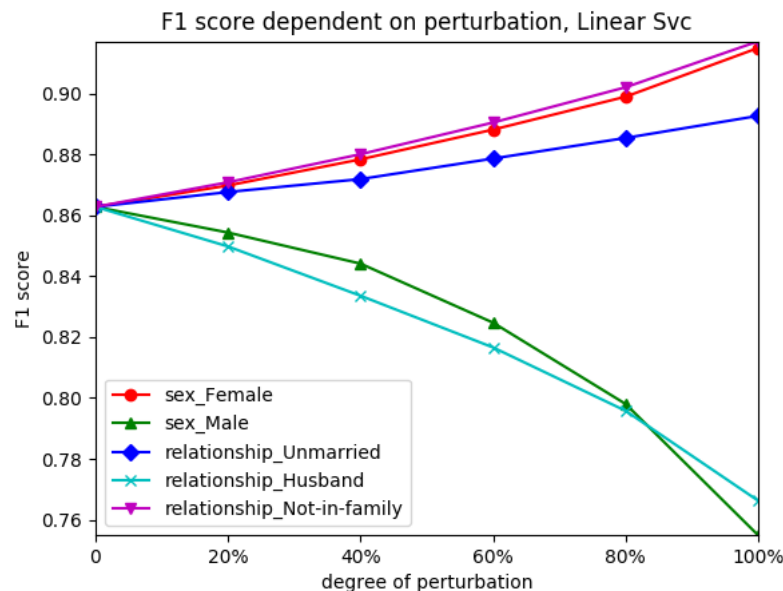
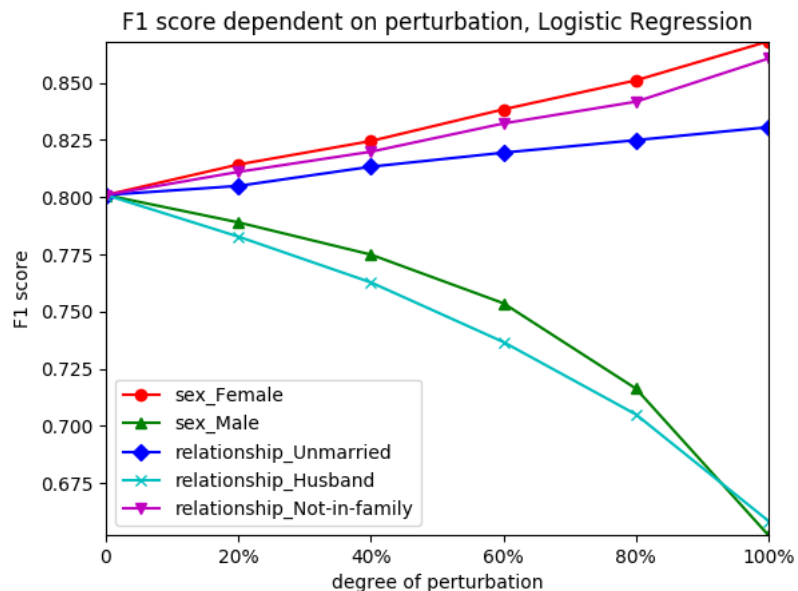


-) results as expected
-) explanation gives insight into the adult DS topology

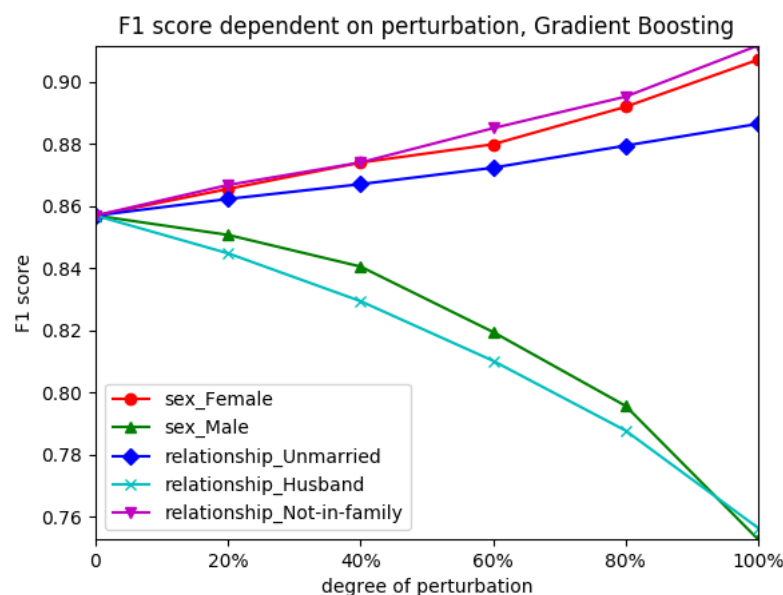
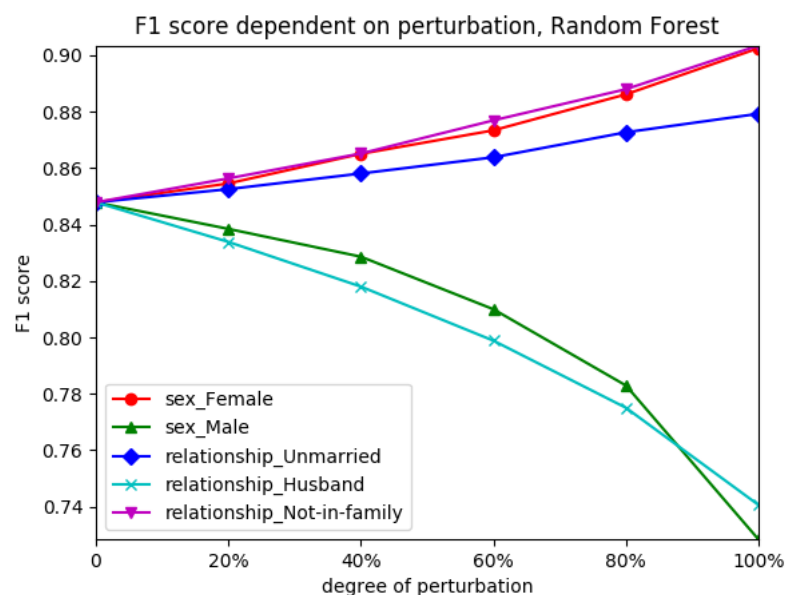


-) relevant for a ML paper (seems to be more data science)

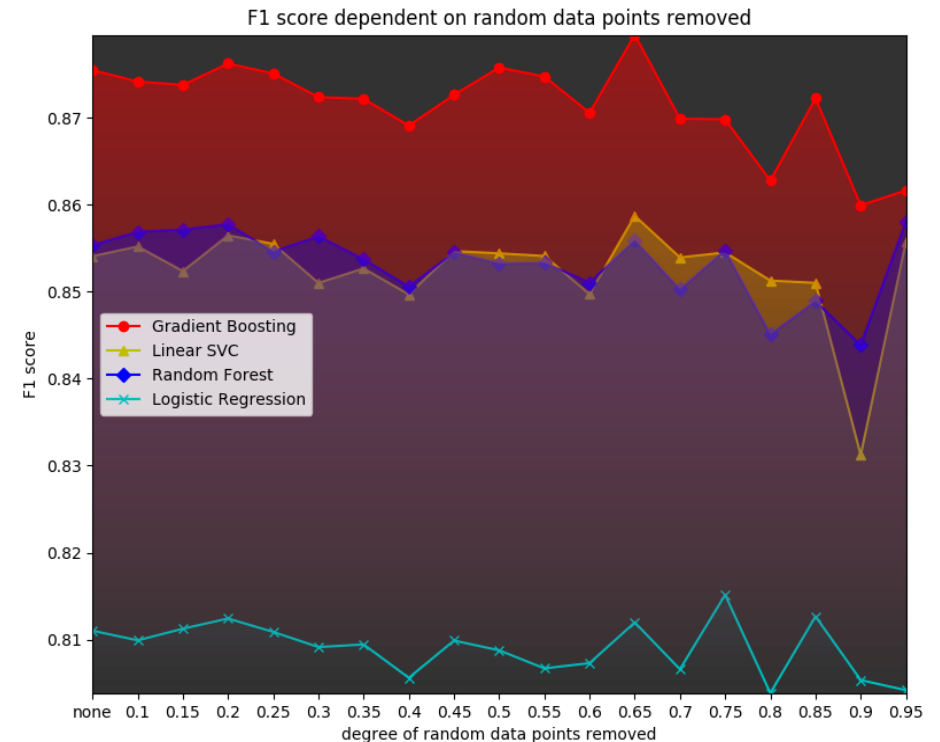
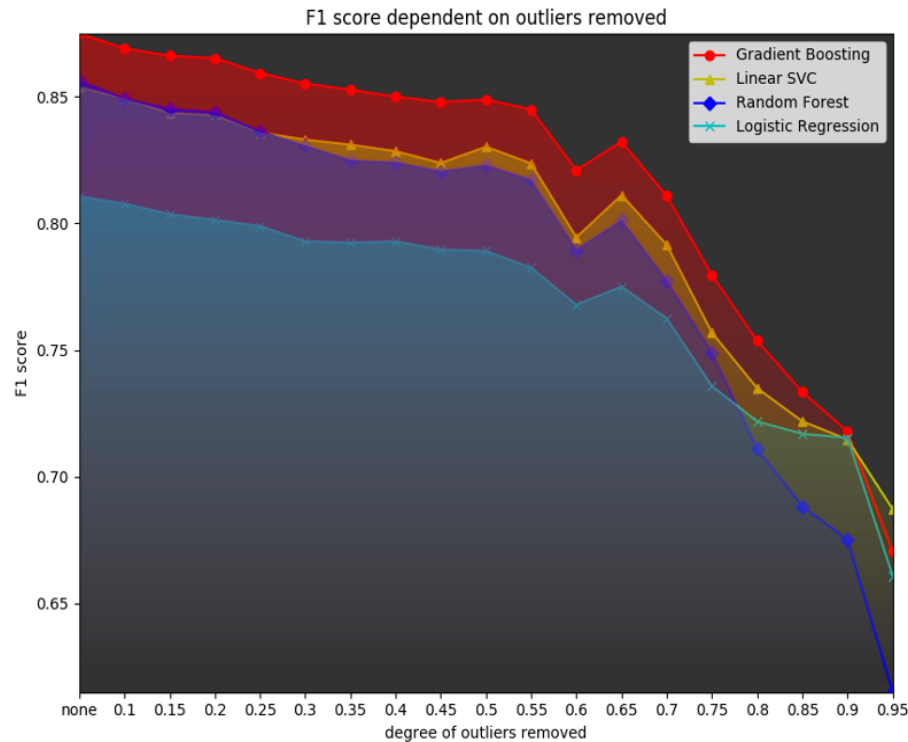
MC on perturbed data - marital_status



-) wonderful !!
-) results can be explained perfectly
-) results might be very relevant for later use
-) explains what to look for in a dataset when perturbing by removing significant columns



MC on outlier-filtered vs. randomly deleted data - income



-) clear behavior consistent across all classifiers

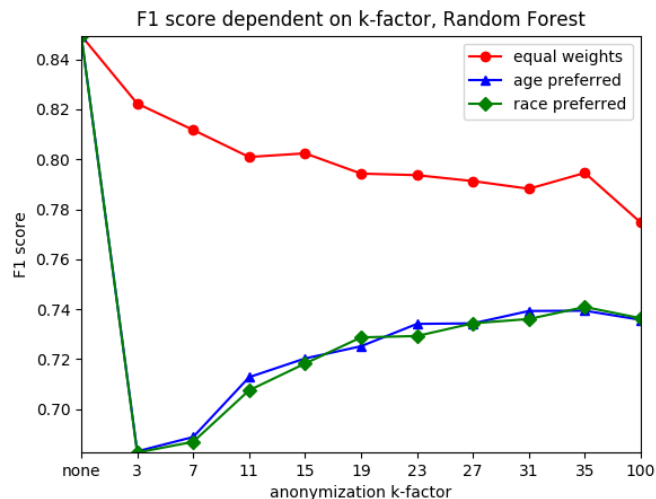
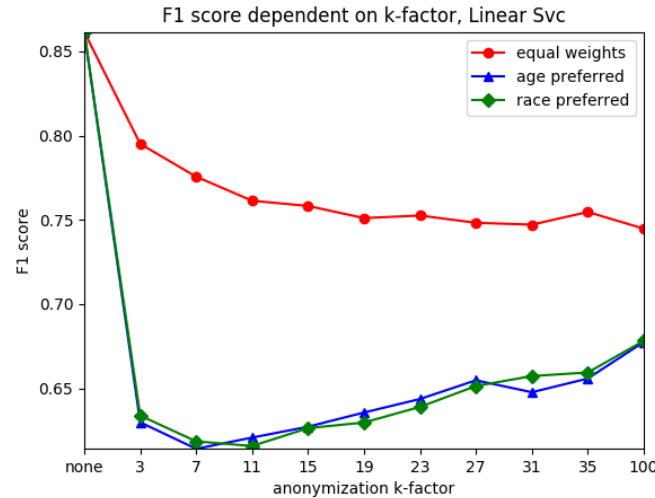
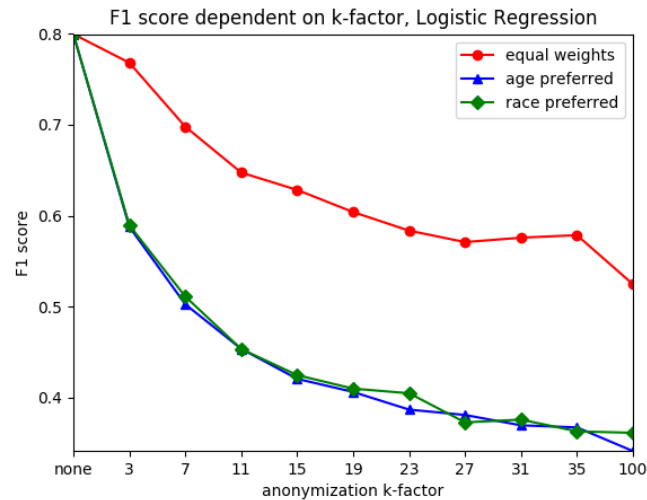
-) maybe good explanation



-) should performance not **increase** with less variance ???

-) maybe not a good explanation

MC on anonymized data - (30% outliers removed) - marital_status



gradient
boosting not
finished
computing...



-) results as
expected with
only slight
deviance from
'normal'
anonymization



-) same
problem with
RF
-) equal seems
to get slightly
better results,
others worse
(except RF)