Data Analysis for Education Consultants

Qaif Fangari May 26th, 2020

Background

Education consultancy is one of the major industry which has been in operation for a quite a some time now, generally their success is measured based on the parameter that how well the student manage to earn after following their advice on the choice of college and the course.
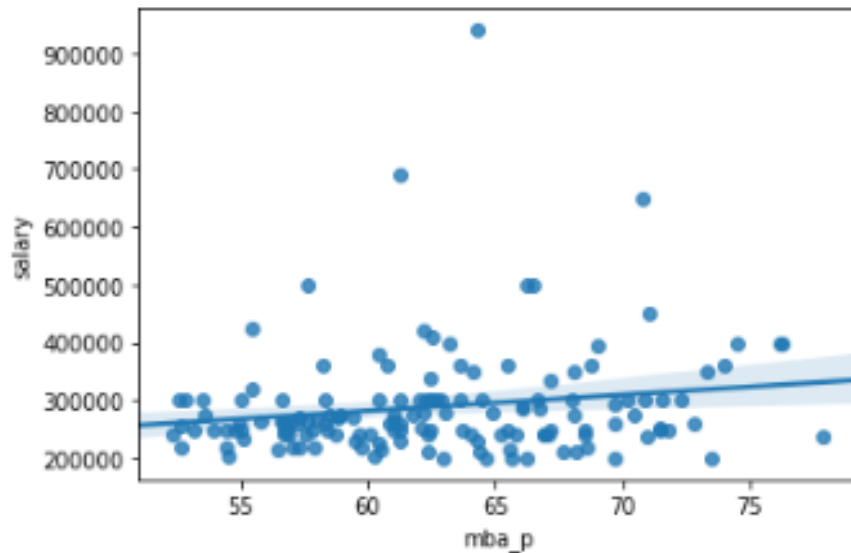
Exploratory Data Analysis
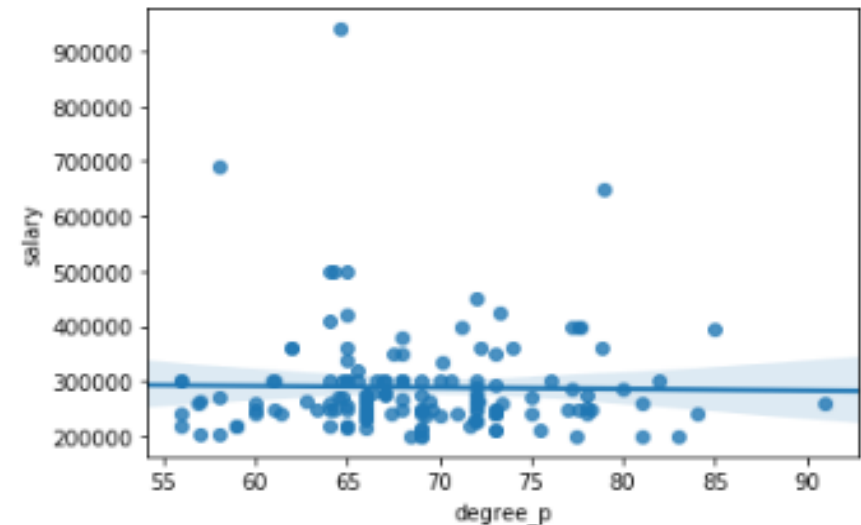
Correlation between non categorical factors

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
sl_no	1.000000	-0.093480	-0.218428	-0.102250	0.041467	-0.072432	0.063764
ssc_p	-0.093480	1.000000	0.293416	0.380657	0.317892	0.430560	0.035330
hsc_p	-0.218428	0.293416	1.000000	0.221307	0.284672	0.329983	0.076819
degree_p	-0.102250	0.380657	0.221307	1.000000	0.217683	0.494093	-0.019272
etest_p	0.041467	0.317892	0.284672	0.217683	1.000000	0.284143	0.178307
mba_p	-0.072432	0.430560	0.329983	0.494093	0.284143	1.000000	0.175013
salary	0.063764	0.035330	0.076819	-0.019272	0.178307	0.175013	1.000000

Correlation between Different factors and salary

Yhat= 114715.29018899679 + [2779.51111313] mba_p
The R-square is: 0.030629529410807547
The mean square error of price and predicted value is: 8409560162.271541

Yhat= 307650.38245922566 + [-276.32859597] degree_p
The R-square is: 0.0003714190255738048
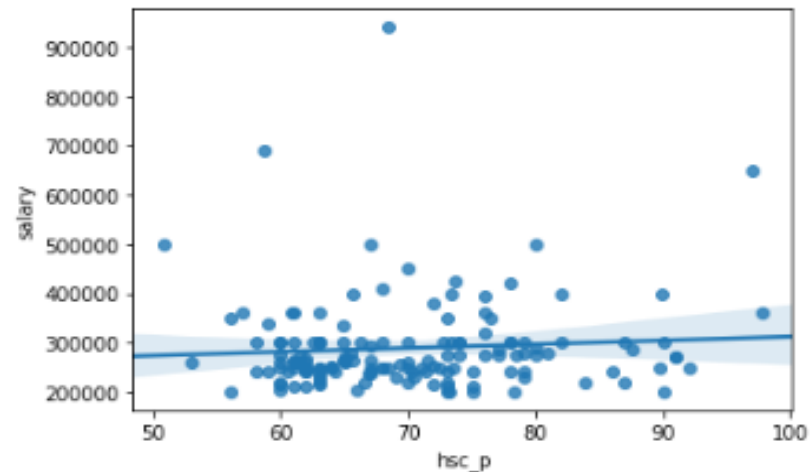The mean square error of price and predicted value is: 8672057739.205791

Correlation between Different factors and salary

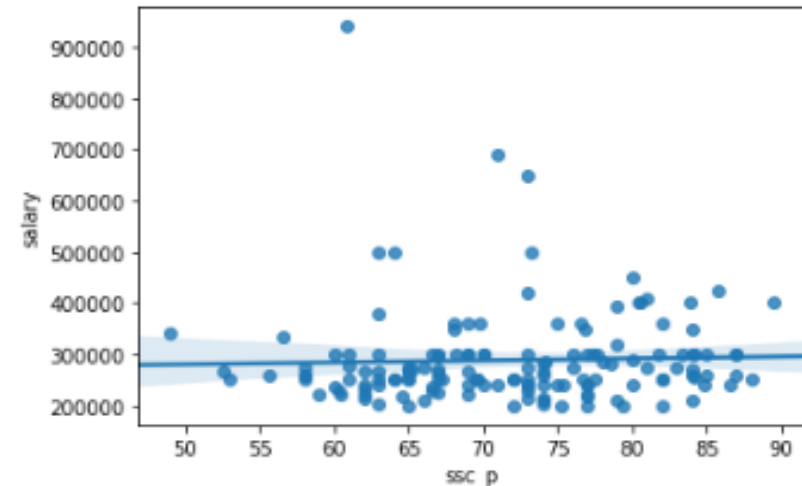Yhat= 234843.62750858403 + [769.54711447] hsc_p
The R-square is: 0.005901163822659261
The mean square error of price and predicted value is: 8624085655.297733

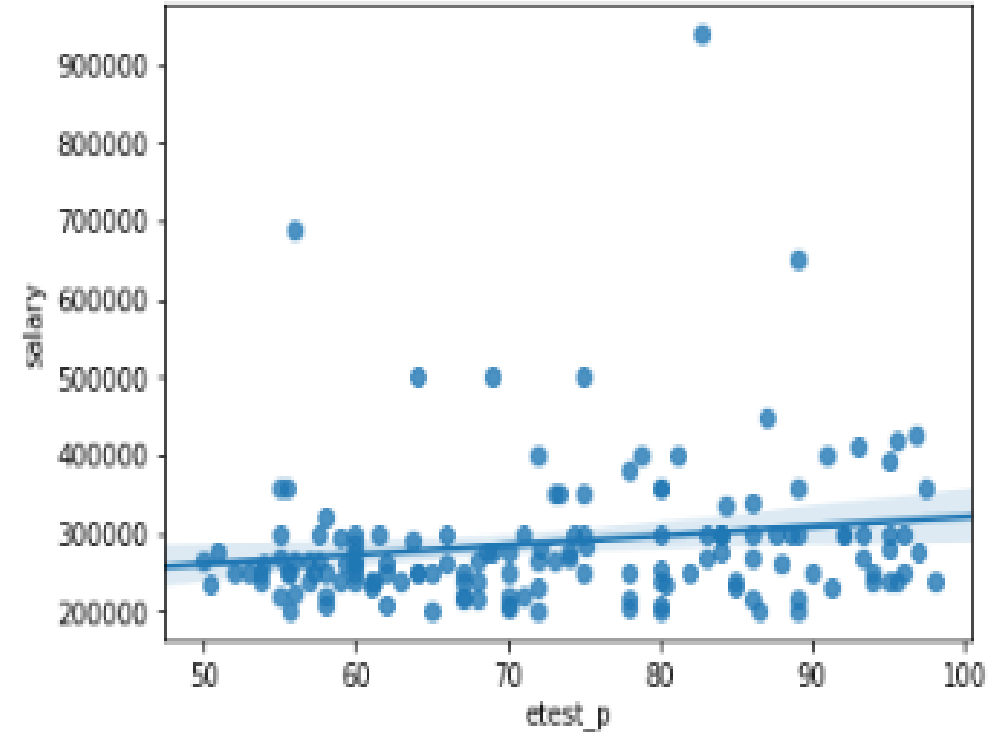Yhat= 261483.4135696075 + [378.85427599] ssc_p
The R-square is: 0.0012482330170948686
The mean square error of price and predicted value is: 8664451132.406292

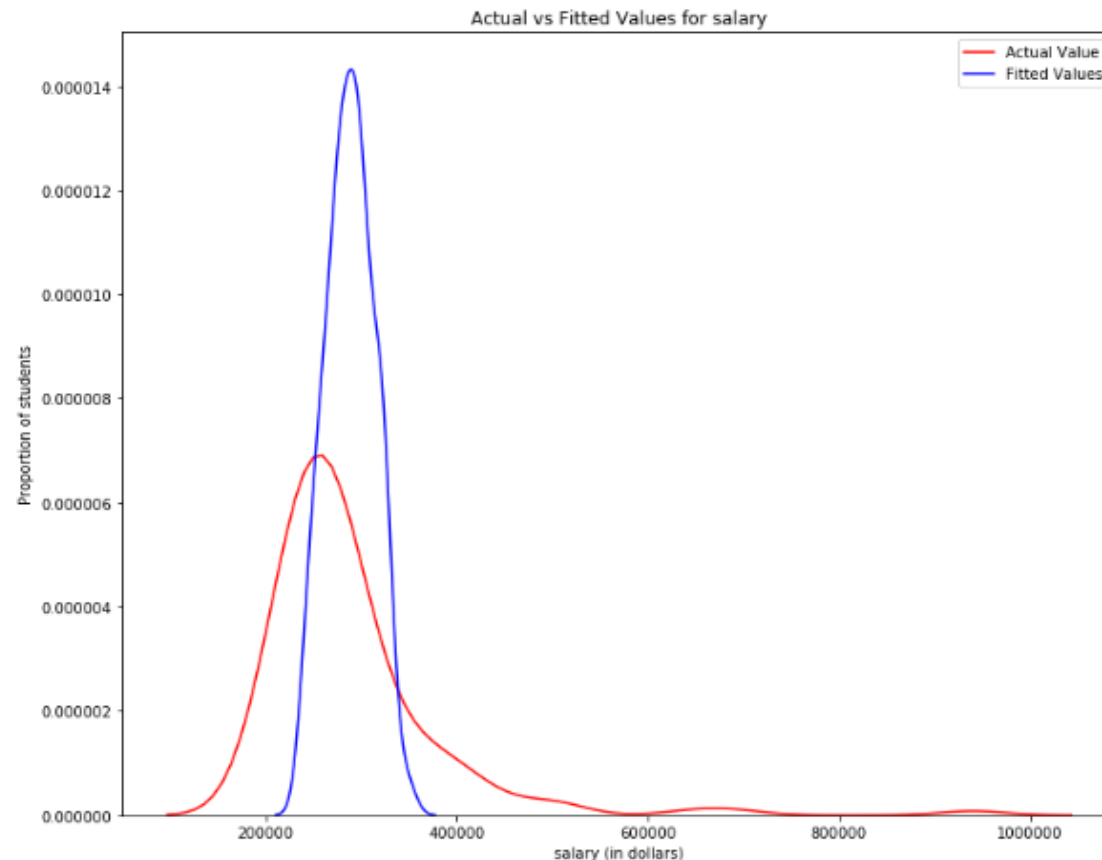# Correlation between Different factors and salary

Yhat= 199761.83197581192 + [1213.76231223] etest_p
The R-square is: 0.03179349643438667
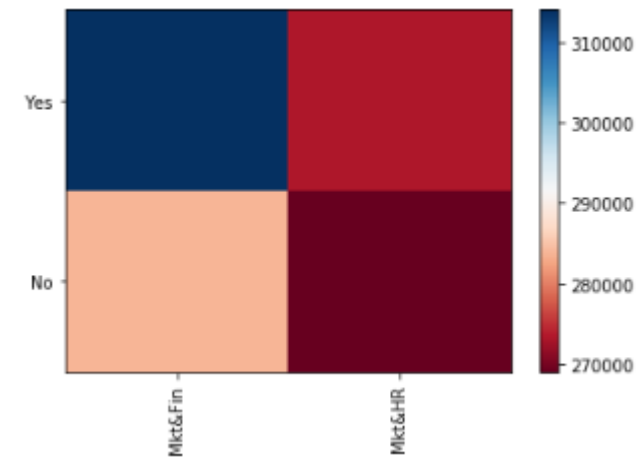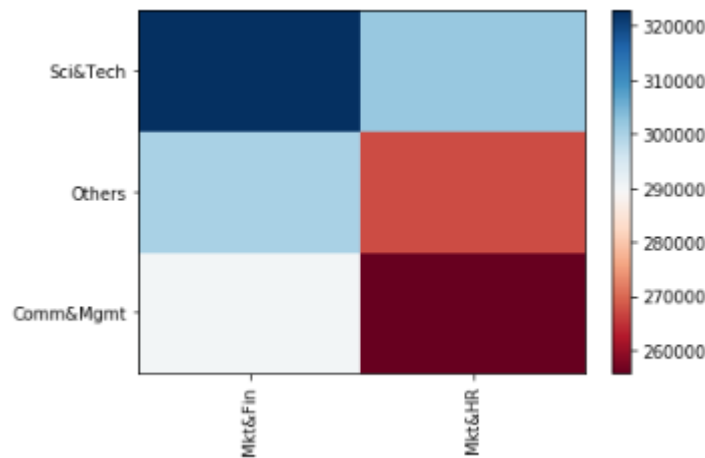The mean square error of price and predicted value is: 8399462422.543881

MODEL FIT

Actual vs Fitted Values for salary

Independent variable used in Multiple linear regression are MBA, Undergrad, Higher secondary, Secondary and aptitude test percentage. The predicted value curve is fitted on actual value curve to know the efficiency of model. If we remove outliers from our dataset then efficiency of model will improve drastically.

Categorical factors

Following matrix helps us summarize the impact of different choices on salary offered to students after post-graduation. From the below two matrix, it can be safely concluded that student with Science and Technology background, having some amount of work experience and Completed MBA with Marketing and Finance specialization are generally offered the highest salaries.

CONCLUSION

It can be safely concluded that student with Science and Technology background, having some amount of work experience and Completed MBA with Marketing and Finance specialization are generally offered the highest salaries.

THANK YOU