**ANALYSIS REPORT**

**INTRODUCTION:**

It is a fact as depicted for MBA programs that they promise graduates from their college would get higher salaries.These MBA colleges boast on their quality of faculty and other factors like amenities they provide.

Financial Times uses rating of MBA programs uses graduates salaries as large components of its rating system.

**BIG QUESTION*: What are the factors that affect the starting salary of graduate students?***

Marie Daer collected data set from anonymous MBA College. She considered various factors such as satisfaction with MBA program, graduates age, sex, GMAT scores, work experience, quartile ranking and other factors.

The main factors she wanted to consider as how starting salary may be dependent on Gender/Age, The likeness of program, GMAT score

Some of the people didn’t wanted to reveal their salary data hence it was **assumed that they were all placed for a second regression model.**

The following analysis was found and these were the key takeaways:-

1. The maximum age was of 48 yrs old but it didn’t affect the mean age as it was 27 yrs.
2. For consideration of GMAT score the highest was 790 while average score is 620 score.
3. The work experience was quite varied with majority of people in the range of 3-5 yrs of work experience.
4. The age showed a positively skewed graph.
5. The no. of male students enrolled for the program was higher than compared to female students.The ratio was approx. 3:1.
6. The GMAT score range showed a farer scoring as it was normal distribution score depicting the college considered students for a particular range.
7. While depicting the Comparison between the genders and the salaries although there was an outlier for females the males still dominated in terms of salary.
8. Although it the mean is approximately same but it might not be true as there is an outlier in the female salary that can affect the mean.
9. On the basis of age it is clear that with age the salaries do increase.Further analysis should be done in this area.
10. There is a fairly linear relationship in the comparison between the Salary and work experience.
11. A negative linear relationship was established between the percentile and Salary.
12. The histogram clearly depicts the quarter ranking that most of the unplaced students are from 1-2 range.
13. The relationship between salary and age are linear relationship.
14. Same is for the work experience and salary. However due to strong linear relationship between work experience and age chances are correlated.
15. Corrplot provides some very interesting insights of the correlation  
     a)work experience and age have value of 0.88 that is almost close to 1  
     b)It is evident that gmat\_tot and gmat\_tpc would be correlated  
     c) Quarter percentile has a negative correlation of 0.84  
     d) corrgram depicts the same plots to understand correlationship better in graphical form.
16. Three models are created.Details are given as following:-  
     i)lm.fit\_full:This model has all factors for the model  
     Residuals are high and Residual Standard error is too high(15430)  
     r-squared:0.3422 and adjusted r-squared:0.2545  
     ii)lm.fit:This model has factors as frstlang+age+sex+work\_yrs+work\_yrs:gmat\_tot  
     Residuals are high and similarly for RSE(15570)  
     r-squared:0.2702 and adjusted r-squared:0.2409  
     iii)lm.fit2:This factors considers only age as an element  
     Residuals are high and so is the RSE(15550)  
     r-squared:0.2496 and adjusted r-squared:0.2422   
     vif stands for variance inflation factor  
     'visreg' to realise visualisations of regression made  
     anova() was used as a comparison method
17. Linear model for people having job and satisfaction from college  
     i)lm.fit3:For Y as a job parameter it is evident that salary would be a factor and so is satisfaction from college   
     r^2=95.29,adjusted r^2=95.23  
     RSE=0.106  
     ii)lm.fit4:Not considering salary as one of the factors  
     r^2=0.1898 and adjusted r^2=0.176  
     RSE=0.4407
18. Hypothesis is accepted and Male have greater salary than woman p>0.05 and it is within the confidence interval
19. For the chi-square test between age and work years these are dependent as x-squared is quite high.

**RESULT: The model created is quite unreliable as it wouldn’t create a good fit when prediction is made. The major factor is though the Work Experience.**

**Secondly, the salaries received by men is more than female.Satisfaction of program is quite high only when when there is job otherwise it’s a negative response.GMAT score with the range of 550-700 is good enough to get selected in the college but it may not be a factor for deciding of salary.**