**Data Science Salary Prediction**

**Steps Involved in Building up the model:**

1. Removed unnamed 0 columns as having all unique.

2. After visualizing the rating and founded column got to know they are having -1 value, so replaced it with mean and median.

3. The Next columns data i have analysed is Job Title which has many unique values after cleaning up the data, I have bought it into 7classes.

4. Based up on the Job Title I have created a new variable called job seniority it has 3 levels they are sr, jr and other.

5. Cleaned up the Salary variable as it is in string with the combination of $ symbol and K making it into per annum.

6. Visualized Top 15 companies and their locations.

7. Cleaning Ownership, Revenue, Industry, Sector, Competitors, column.

8. Adding columns like job in headquarters and python, excel, sql, tableau jobs. and drop job location, job description.

9. Trimming up the columns like Sector which are having a greater number of classes.

10. Dropping Salary Estimate, Company name, Headquarters, Industry, job location.

11. performing labelEncoding, Get dummies and Standardscaler.

12. Model selection Random Forest performed better when we compare with all other algorithms.