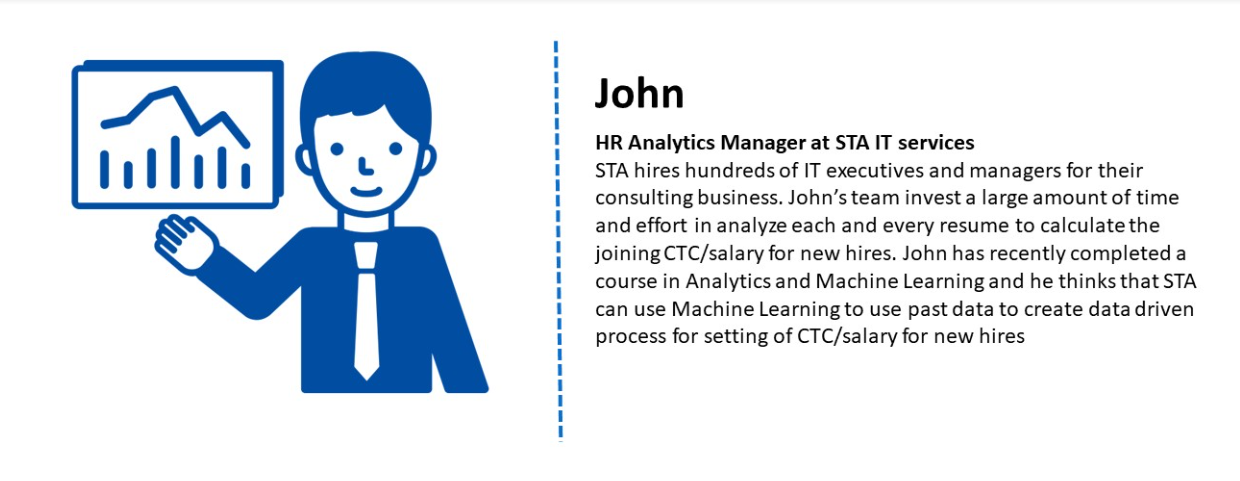
***Task 1 Problem***

Your task is to help John from STA IT services to create a linear regression model in Python or R to predict the CTC/Salary of new hires from the data provided. Here is a snapshot of the case



Your task is to

1. Understand the data
2. Transform categorical variables to the numerical variables
3. Apply linear regression to predict the CTC
4. Evaluate the results of the analysis

Excel data files are provided to you to help you with the assignment.

[Data\_file.xlsx](https://docs.google.com/spreadsheets/d/1bcOXd9dFhgxsmonqC0TtFLQrAKfzxQNL/edit?usp=sharing&ouid=106115340208586563833&rtpof=true&sd=true)

Data\_file is a dataset that you will use to train your model.

Test Data files:

[Test\_data\_file.xlsx](https://docs.google.com/spreadsheets/d/1rsyHvyjVzl8YoIzRgyqNTxJIVUoNhivL/edit?usp=sharing&ouid=106115340208586563833&rtpof=true&sd=true)

Test\_Data\_file is a dataset that you will use to test your model and find out the Mean Square Value.