**JOB-A-THON PROJECT ANALYSIS**

**Agenda:**

The main objective is to determine whether the employee will leave the company or not. We can achieve this by using multiple algorithms/Statistical Techniques.

To determine whether the employee will leave the company or not, we have to analyse the respective predictors or supporting features like Age, City, Salary etc.

**Data Pre-processing:**

* In the current dataset, the response variable is not given directly. Instead of that they had given the last working date. By using python code we created a new variable called ‘Target’ and assigned values by using last working date variable
* I analysed each feature to know the Insights of that feature. I achieved this by using different plots like Histogram, Boxplot, countplot etc
* I used the **Standard Scaler** to standardize the numerical features, it transformed the numerical features values between the range -1 to 1
* The categorical features are labelled into the numerical values using Label Encoder
* In the current dataset, the values of Target variable are unbalanced. To balance the data, we used **SMOTE** method to balance the data
* Grouping is performed on Employee id by averaging the Total Business Value as there are repeated records with same Employee Id with only Total Business Value changing.

**Splitting the Data:**

We will split the data into training and testing. Here, we are splitting the train data into 75% and test data into 25%. We achieved this by using train\_test\_split in the sklearn library.

**Statistical Techniques/Algorithms:**

* For the current dataset, we used multiple machine learning algorithms like DecisionTree Classifier, RandomForest Classifier, etc.
* But, svm/Support Vector Classifier gives us maximum accuracy(73.07%) with respect to public test data.
* So, I am using the svm algorithm to predict whether the employee will leave the company or not.

**Further Improvements:**

We can further improve our model by applying different hyperparameter tuning techniques for respective algorithms.