**IBM HR Analytics Employee Attrition Modelling**

DESCRIPTION

IBM is an American MNC operating in around 170 countries with major business vertical as computing, software, and hardware.  
Attrition is a major risk to service-providing organizations where trained and experienced people are the assets of the company. The organization would like to identify the factors which influence the attrition of employees.

****Data Dictionary****

* Age: Age of employee
* Attrition: Employee attrition status
* Department: Department of work
* DistanceFromHome
* Education: 1-Below College; 2- College; 3-Bachelor; 4-Master; 5-Doctor;
* EducationField
* EnvironmentSatisfaction: 1-Low; 2-Medium; 3-High; 4-Very High;
* JobSatisfaction: 1-Low; 2-Medium; 3-High; 4-Very High;
* MaritalStatus
* MonthlyIncome
* NumCompaniesWorked: Number of companies worked prior to IBM
* WorkLifeBalance: 1-Bad; 2-Good; 3-Better; 4-Best;
* YearsAtCompany: Current years of service in IBM

Analysis Task:  
- Import attrition dataset and import libraries such as pandas, matplotlib.pyplot, numpy, and seaborn.  
- Exploratory data analysis

* Find the age distribution of employees in IBM
* Explore attrition by age
* Explore data for Left employees
* Find out the distribution of employees by the education field
* Give a bar chart for the number of married and unmarried employees

- Build up a logistic regression model to predict which employees are likely to attrite.