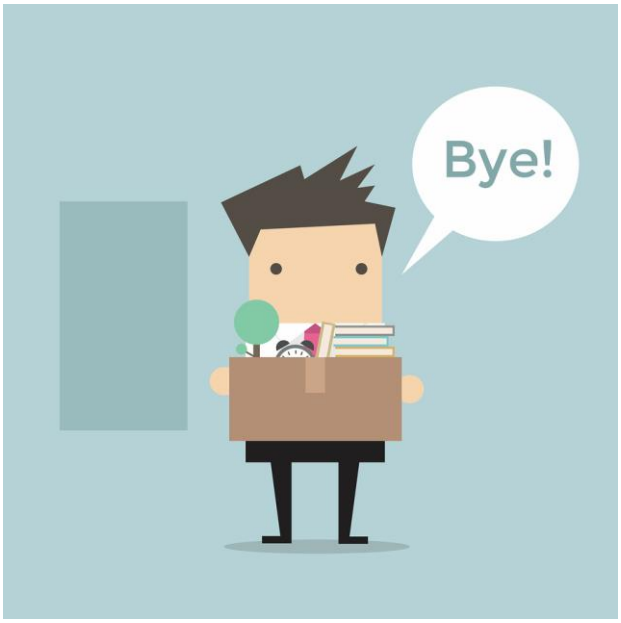
FINAL REPORT

Employee Attritions



The company wants to understand what factors contributed most to employee turnover and to create a model that can predict if a certain employee will leave the company or not. The goal is to create or improve different retention strategies on targeted employees. Overall, the implementation of this model will allow management to create better decision-making actions.

Summary:

This Project particularly describes 9the IBM HR Analytics Employee Attrition & Performance dataset for the further process in the project.
It explains what kind of steps were performed on this particular data set, how the missing values or the outliers handled

Solution:

1)Gathering Data: Identify various data source, collect data, Integrate the data obtained from different sources.

```
# Importing Libraries

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

2)Data Preparation: Data exploration A better understanding of data leads to an effective outcome. In this, we find Correlations, general trends, and outliers.

```
# Importing Data Set
df = pd.read_csv('C:\\Users\\pc\\Desktop\\Data Science\\MachineLearning\\Project\\Employee Attrition\\Employee-Attrition.csv')
df.head()
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	EmployeeNumber	...	RelationshipS
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences	1	1	...	
1	49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences	1	2	...	
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other	1	4	...	
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences	1	5	...	
4	27	No	Travel_Rarely	591	Research & Development	2	1	Medical	1	7	...	

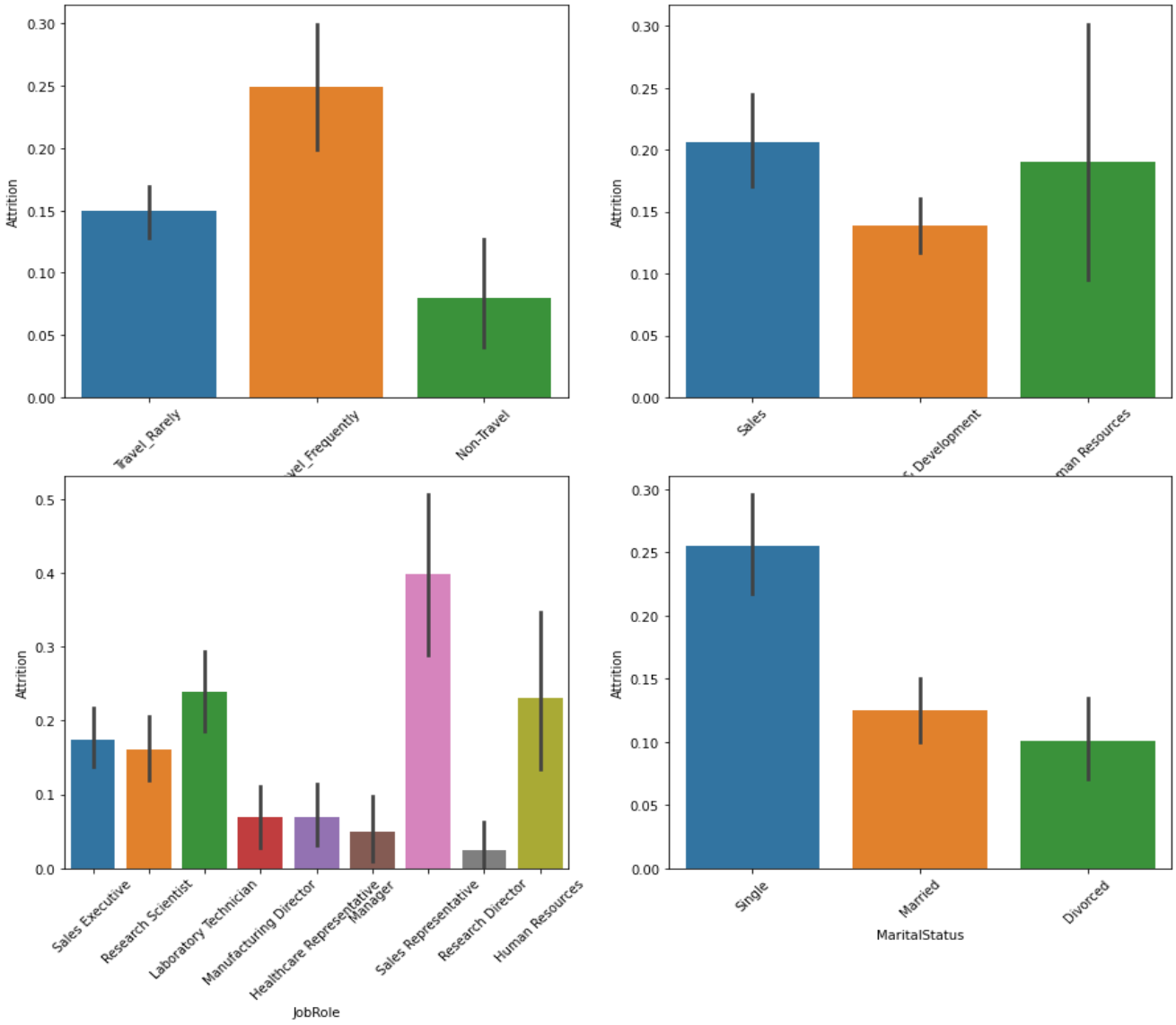
3)Data Wrangling: Missing Values, Duplicate Data, Invalid data, Noise.

```
# Checking Null Values
df.isnull().sum()
```

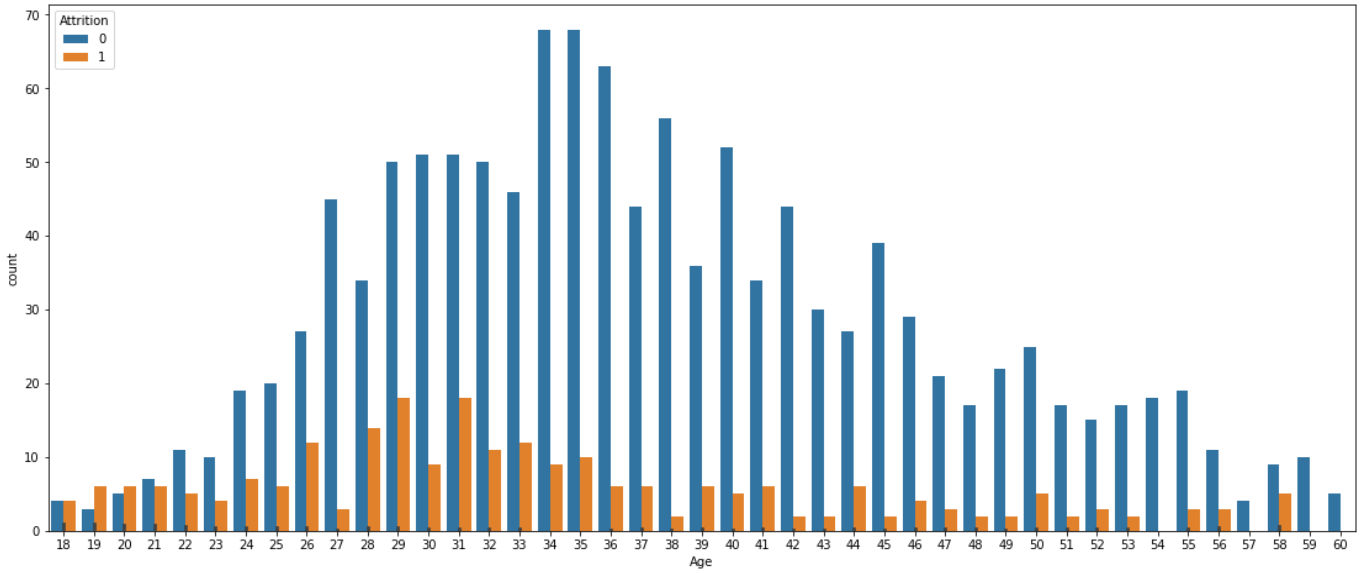
```
Age                         0
Attrition                   0
BusinessTravel              0
DailyRate                   0
Department                  0
DistanceFromHome            0
Education                   0
EducationField              0
EmployeeCount               0
EmployeeNumber              0
EnvironmentSatisfaction     0
Gender                      0
HourlyRate                  0
JobInvolvement              0
JobLevel                    0
JobRole                     0
JobSatisfaction             0
MaritalStatus               0
MonthlyIncome               0
MonthlyRate                 0
NumCompaniesWorked          0
Over18                      0
OverTime                    0
PercentSalaryHike           0
PerformanceRating           0
RelationshipSatisfaction    0
StandardHours               0
StockOptionLevel            0
TotalWorkingYears           0
TrainingTimesLastYear       0
WorkLifeBalance             0
YearsAtCompany              0
YearsInCurrentRole          0
YearsSinceLastPromotion     0
YearsWithCurrManager        0
dtype: int64
```

4)EDA

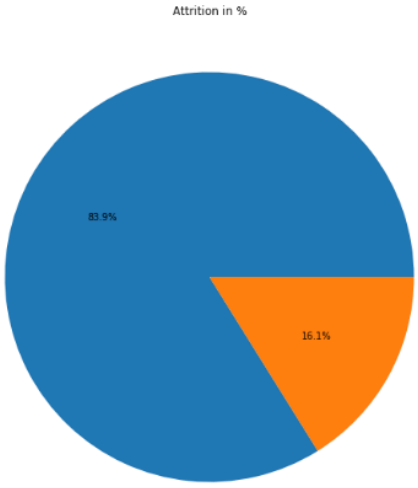
Job Roles by Various Categories



Attrition by Age



4)Data Analysis: Selection of analytical techniques, building models, Review the Result.



5)Train Model: train the model using various machine learning algorithms. Training a model is required so that it can understand the various patterns, rules, and, features.

1. **Logistic regression**
2. **Decision Tree Classification**
3. **Random Forest Classification**
4. **Ada Boost Classifier**
5. **Gradient Boosting Classifier**

6)Test Model: Testing the model determines the percentage accuracy of the model as per the requirement of project or problem.

1. Accuracy with LogisticRegression: 83.06
2. Accuracy with DecisionTreeClassifier: 75.92
3. Accuracy with Random Forest Classifier: 84.29
4. Accuracy Score with AdaBoostClassifier: 85.5

7)Model Selection

Ada Boost Classifier performed well on this data set

Conclusion:

Top Reasons why Employees leave the Organization:

No Overtime:

Employees who don't have overtime are most likely to leave the organization. This could be that employees would like to have a higher amount of income or employees could feel that they are underused.

Monthly Income:

As expected, Income is a huge factor as why employees leave the organization in search for a better salary.

Age:

This could also be expected, since people who are aiming to retire will leave the organization.
Knowing the most likely reasons why employees leave the organization, can help the organization take action and reduce the level of Attrition inside the organization.