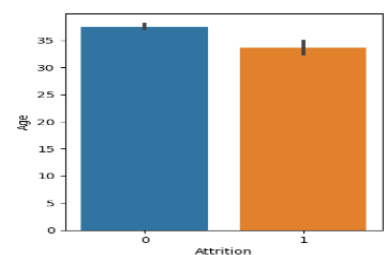
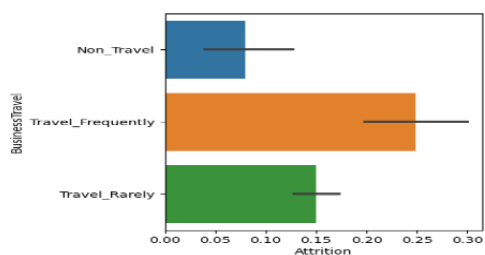
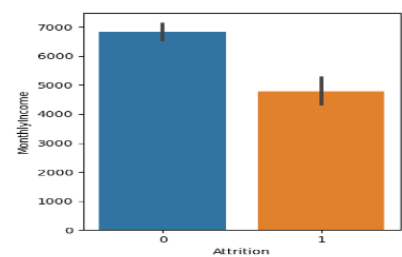
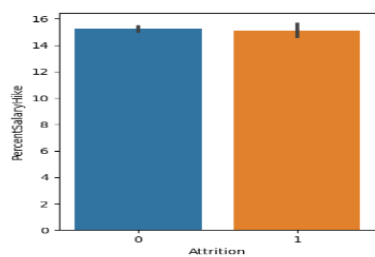
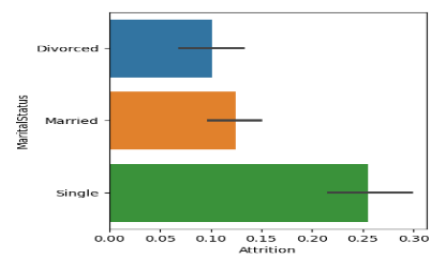
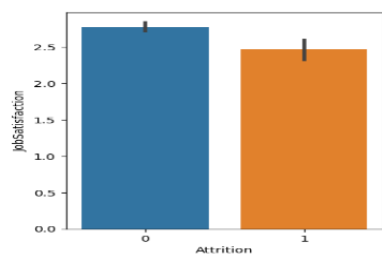
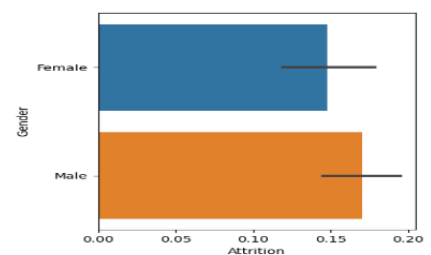
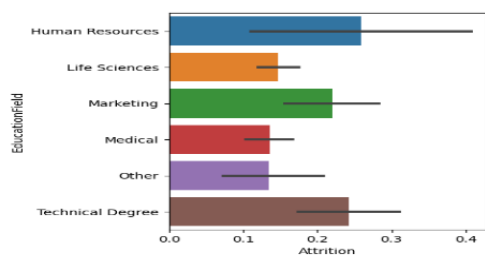
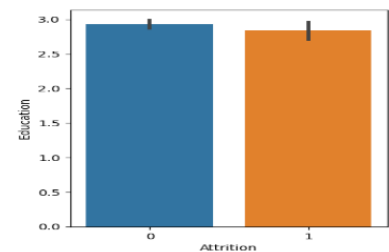
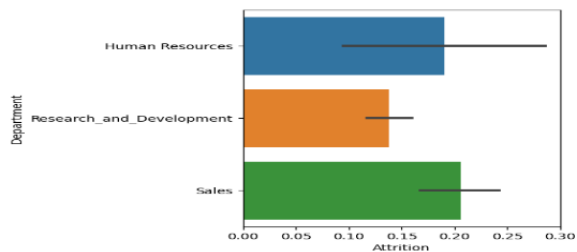
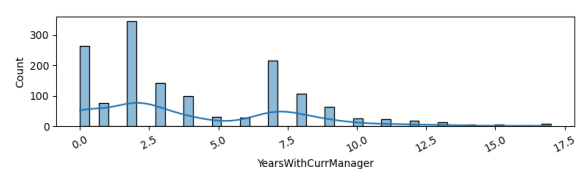
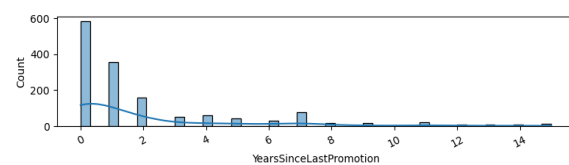
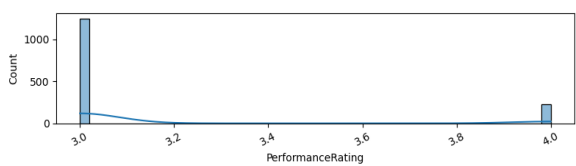
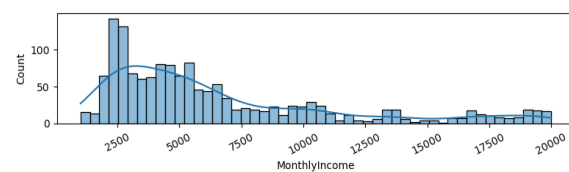
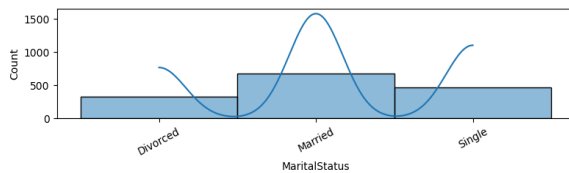
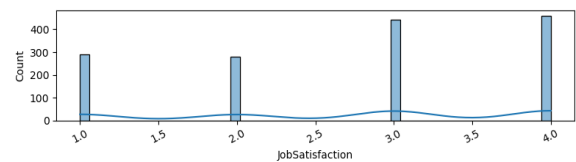
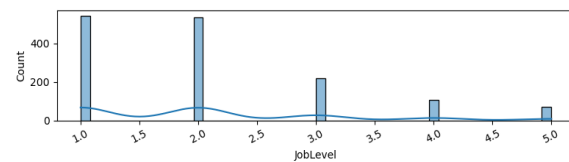
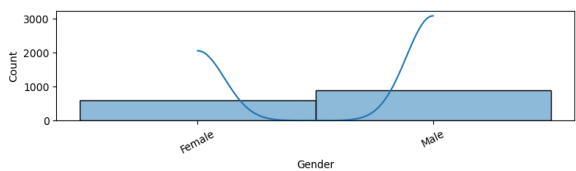
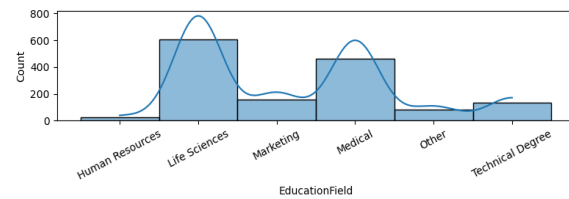
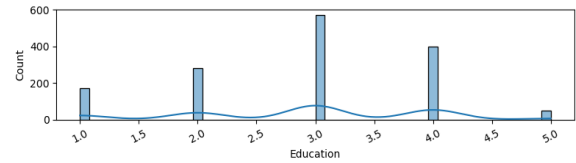
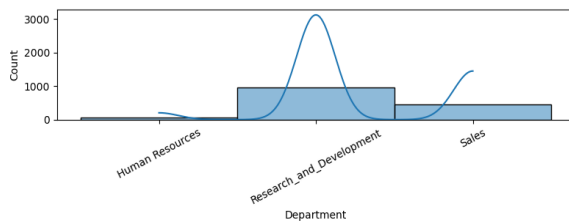


Build a predictive model for employee attrition on Python

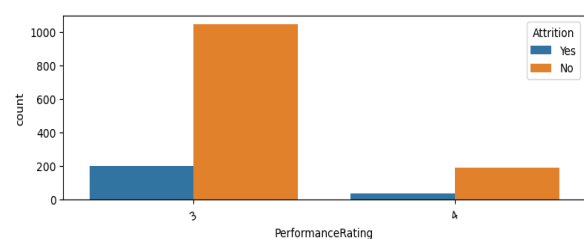
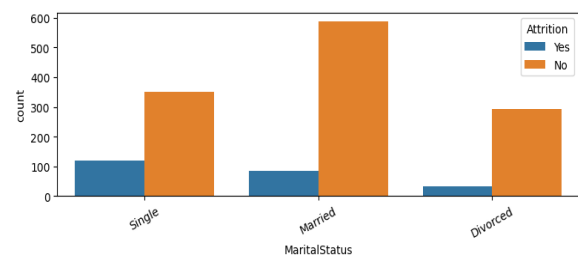
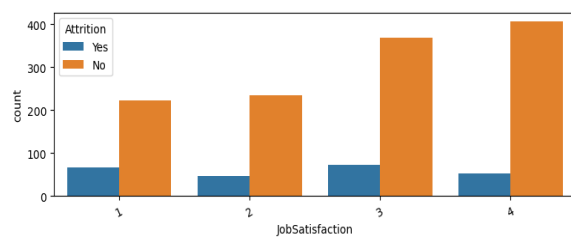
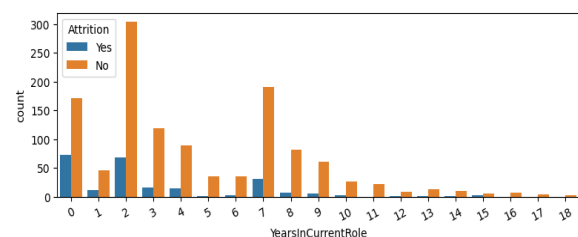
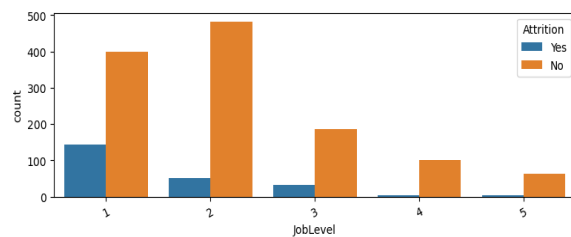
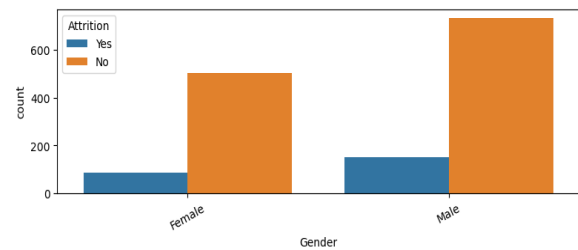
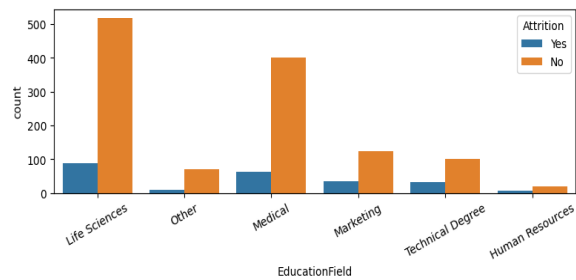
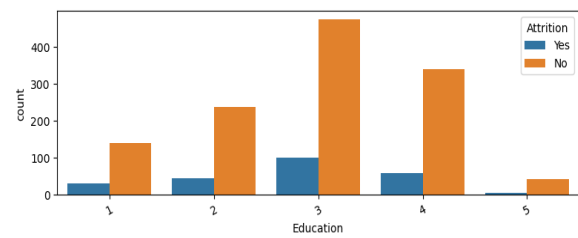
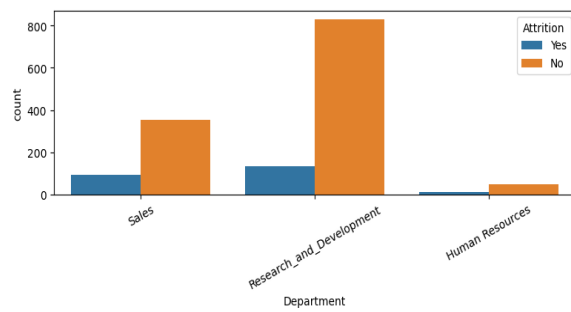
Drawn insights



Build a predictive model for employee attrition on Python



Build a predictive model for employee attrition on Python



Build a predictive model for employee attrition on Python

- Sales Department has highest Attrition rate
- The employee who Travel Rarely has highest Attrition rate
- Human Resources has most Attrition Education Field
- Generally male employee has highest attrition rate
- The one who has Work Life Balance is 1 , having greater attrition rate
- Lower level designation has highest attrition rate
- Number of employee who got performance rating 3
- At younger age their is highest attrition rate

Model Accuracy:

1. Logistic Regression=67%
2. Decision tree=67%
3. Random Forest=66%

ROC Curve for Logistic Regression

Conclusion: 1) The more that the ROC curve hugs the top left corner of the plot, the better the model does at classifying the data into categories. 2)To quantify this, we can calculate the AUC (area under the curve) which tells us how much of the plot is located under the curve. 3)The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. 4) The Model which has the highest AUC, which indicates that it has the highest area under the curve and is the best model at correctly classifying observations into categories. 5)Here the AUC value=0.768 ,so that this model has been Acceptable discrimination.

ROC Curve for Logistic Regression using SMOTE

Conclusion The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. Here the value of AUC is 0.711 ,means this model is Acceptable discrimination for classification

ROC Curve for Decision Tree classifier

Conclusion=The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. Here the value of AUC is 0.591, which is closer to 0.5 ,means this model is not good for classification.

ROC Curve for Decision Tree classifier with SMOTE

Build a predictive model for employee attrition on Python

Conclusion : The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. Here the value of AUC is around 0.940 ,means this model is good for classification.

ROC Curve for Random Forest classifier

Conclusion=The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. Here the value of AUC is around 0.830, which closer to 1, means this model has Excellent discrimination.

ROC Curve Random Forest classifier SMOTE

Conclusion=The closer AUC is to 1, the better the model. A model with an AUC equal to 0.5 would be a perfectly diagonal line and it would represent a model that is no better than a model that makes random classifications. Here the value of AUC is 0.985, which is closer to 1 ,means this model is better for classification.