**Jaypee Institute of Information Technology, Noida**

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING AND INFORMATION TECHNOLOGY



**COURSE -** Open Source Software Lab (15B17CI575)

**Project Title:-** PLACEMENT PREDICTION USING ML

### 

| **ENROLLMENT NUMBER** | **NAME OF THE MEMBER** | **BATCH** |
| --- | --- | --- |
| 21104019 | ABHILASH TYAGI | B-14 |
| 21104027 | OJASVI PANDEY | B-14 |

### 

### 

### 

### 

### 

### **1. Briefing about the Project:**

#### Introduction:

In the dynamic landscape of higher education, campus placements are pivotal for engineering students. The project aims to leverage machine learning techniques to predict the likelihood of a student getting placed based on various factors.

#### Motivation:

The motivation is to provide students and educational institutions with a tool to foresee potential placement outcomes. This predictive model could assist students in better preparation and help institutions identify areas of improvement in their curriculum or placement processes.

#### Scope:

The project focuses on predicting placement outcomes for engineering students based on historical data. Features include academic performance, technical skills, communication skills, and extracurricular activities. The prediction is binary - whether a student will be placed or not, with limitations such as assuming historical trends indicate future patterns and relying on available features.

### **2. Methods/Models/Algorithms:**

#### Model Selection:

**Logistic Regression:**

Logistic Regression is chosen for its simplicity and interpretability, offering a clear understanding of the relationship between input features and placement likelihood.

**Random Forest:**

Random Forest is selected for handling non-linear relationships and reducing overfitting. Its ensemble nature allows capturing complex patterns in the data.

#### Data Preprocessing:

Extensive data preprocessing is performed, including handling missing values, standardizing numerical features, and encoding categorical variables. The goal is to ensure models receive clean and standardized input data.

#### Model Training:

The dataset is split into training and testing sets. The training process involves iteratively refining model parameters.

### 

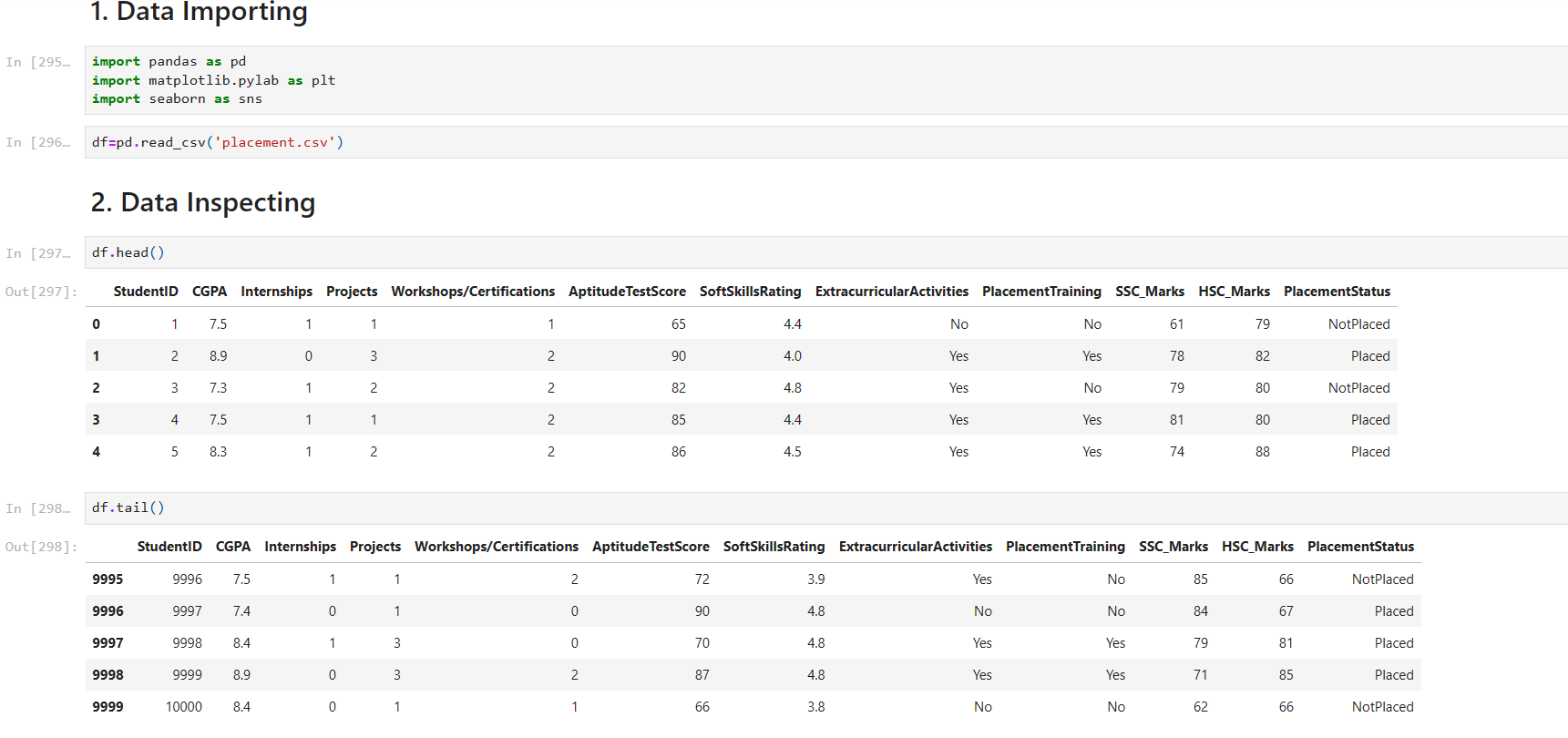
### 

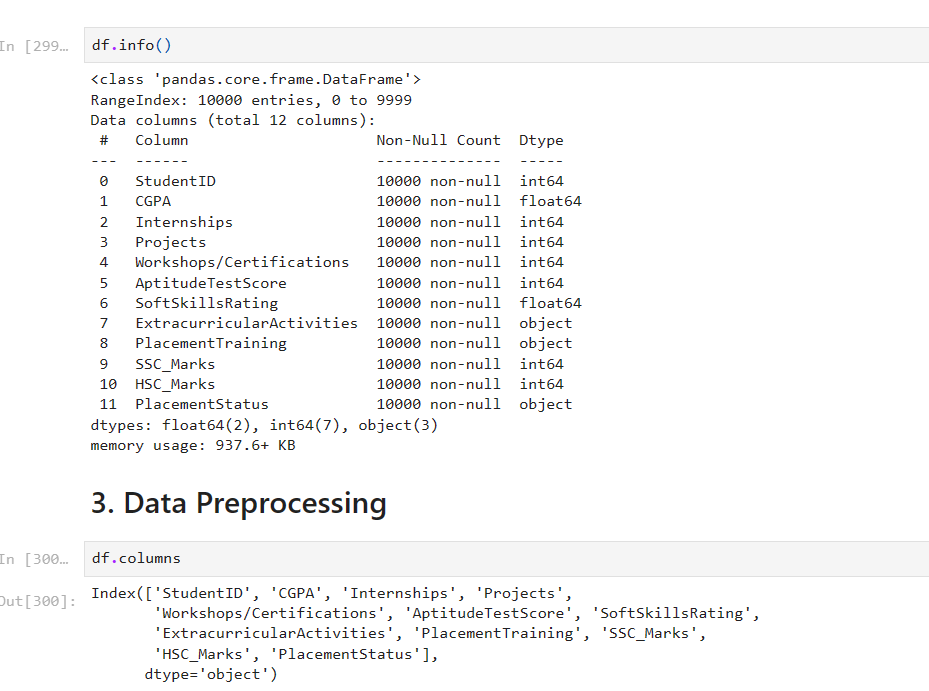
### **3. Results and Learnings:**

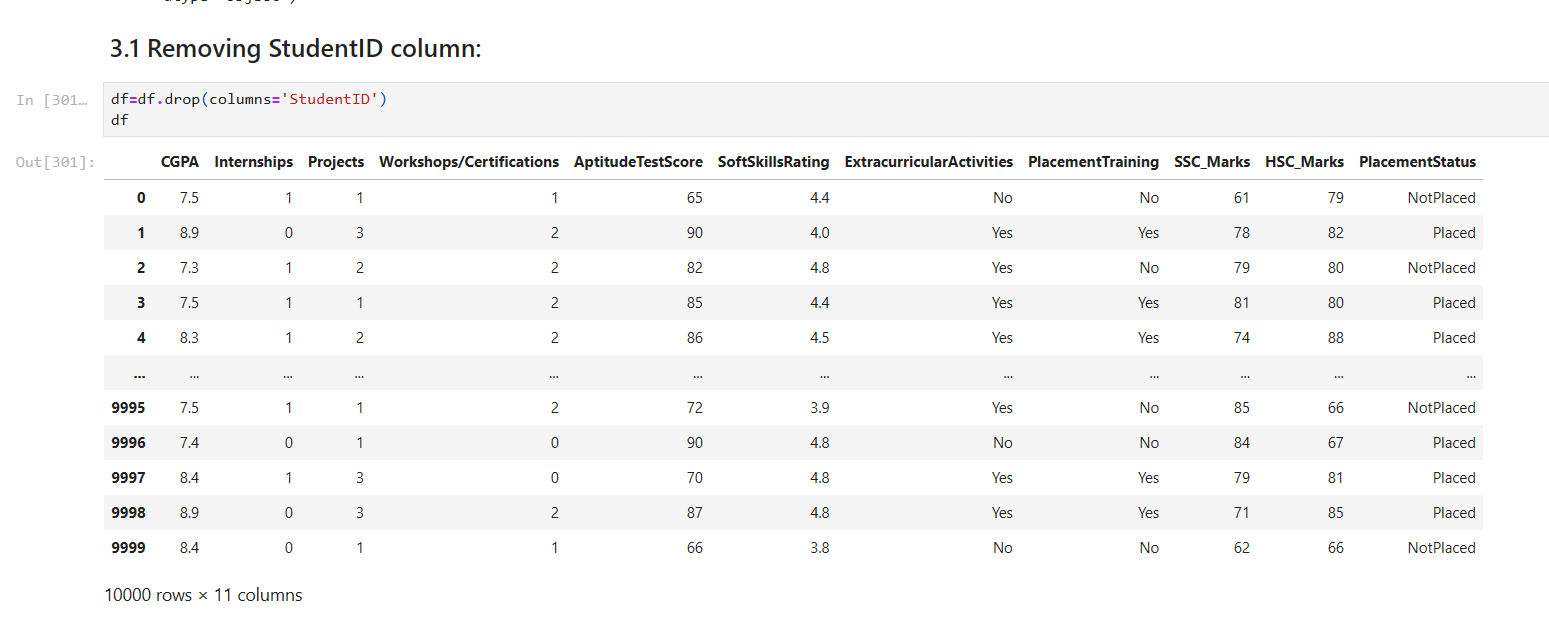
#### Evaluation Metrics:

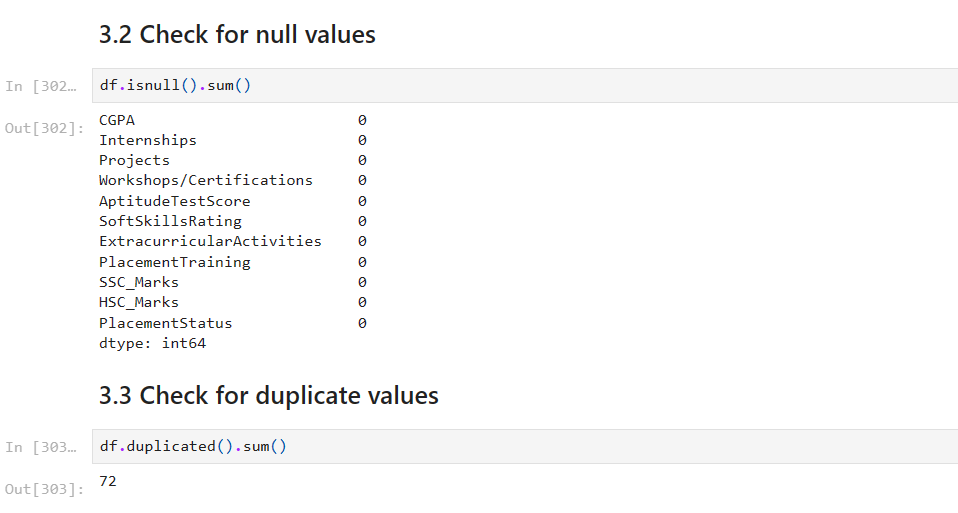
Several metrics, including accuracy are employed to assess models' performance, providing a comprehensive understanding of strengths and weaknesses.

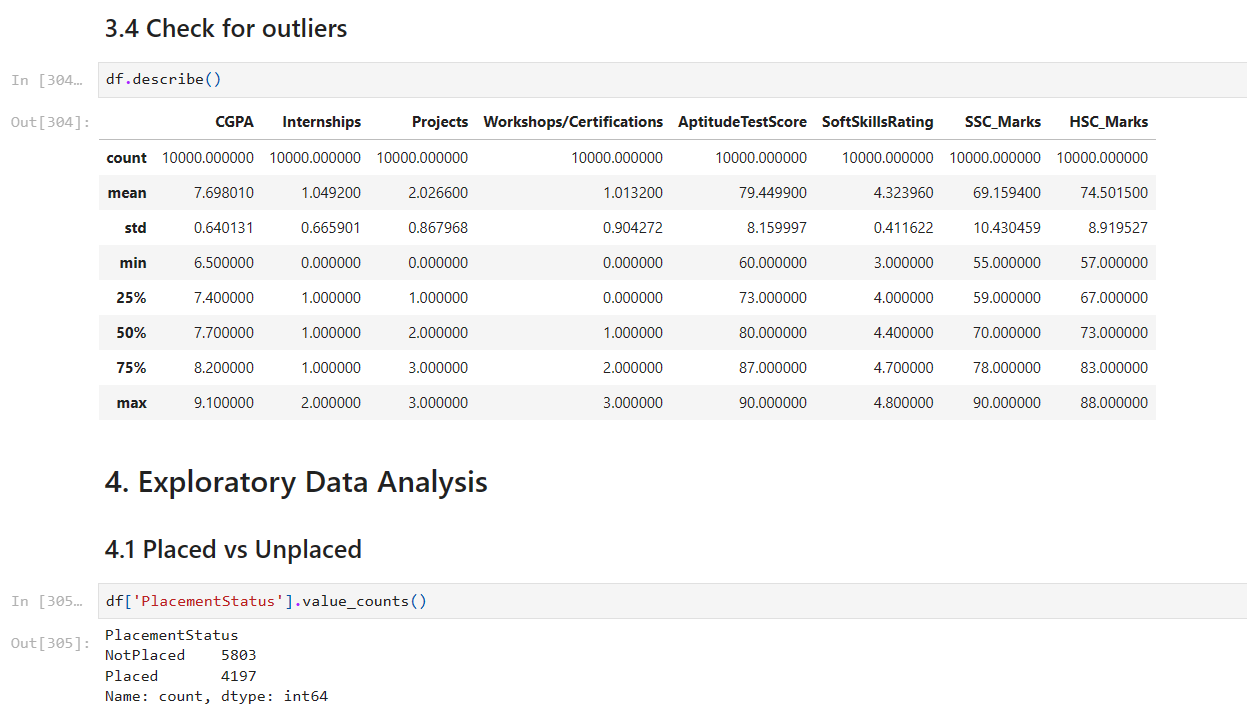
**SNIPPETS OF THE OUTPUT:-**

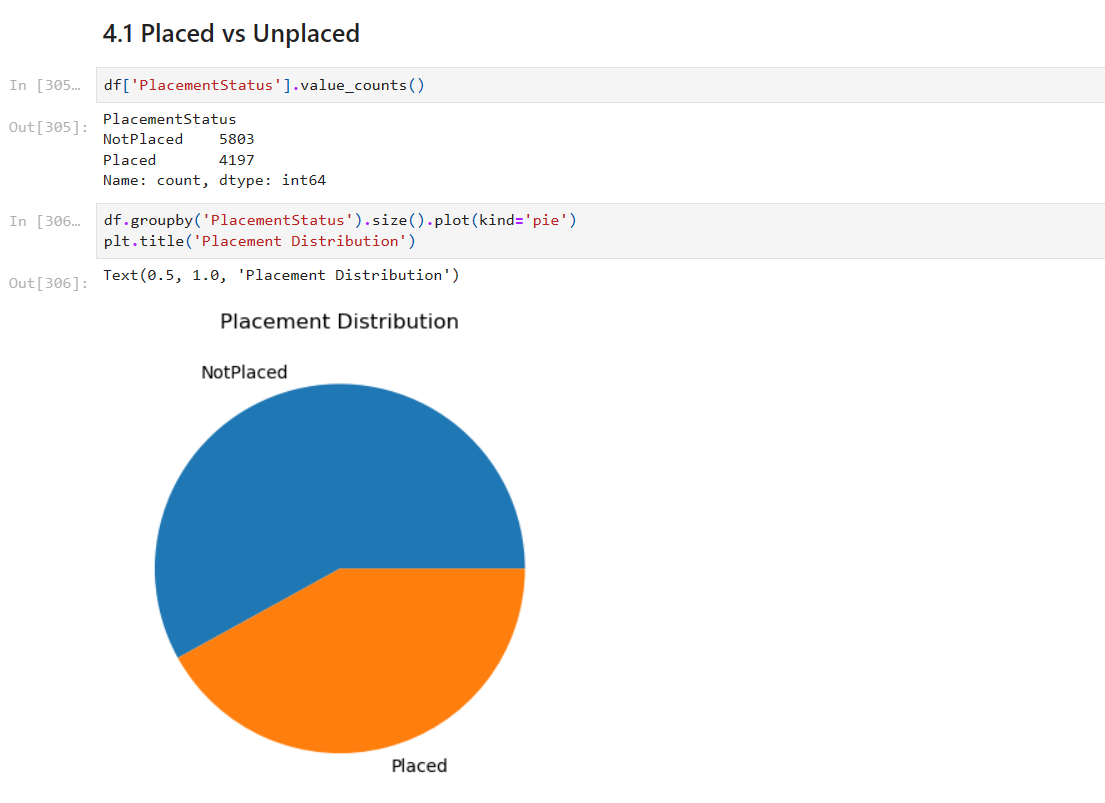
****

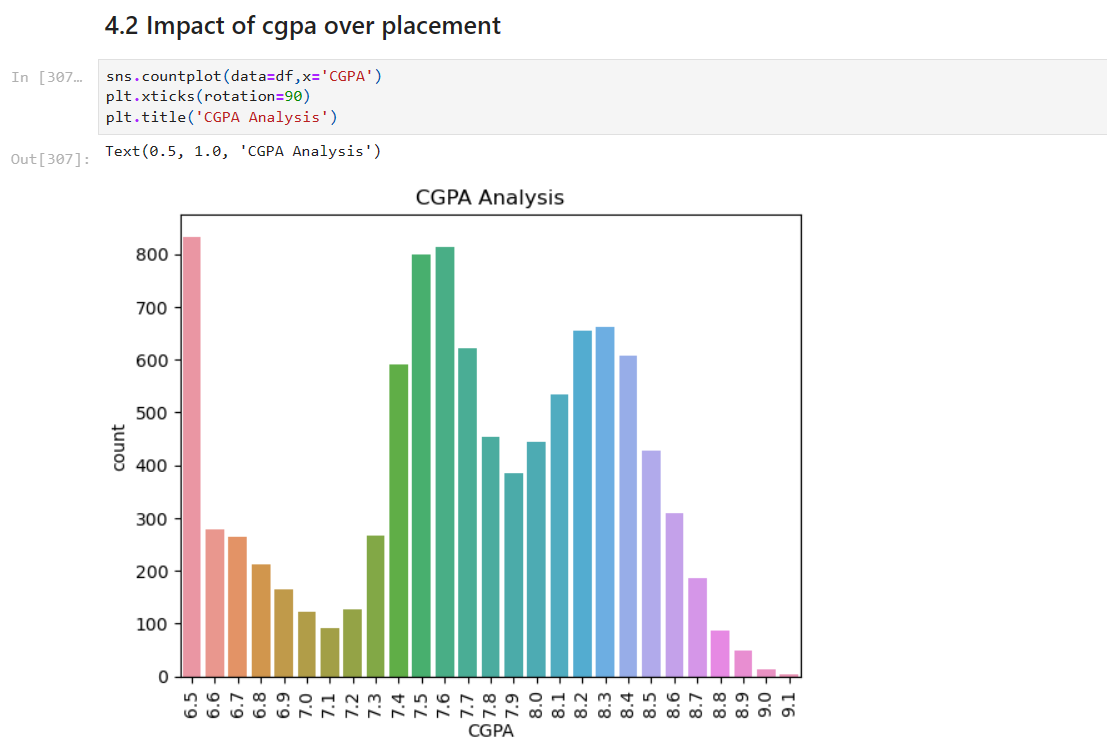
****

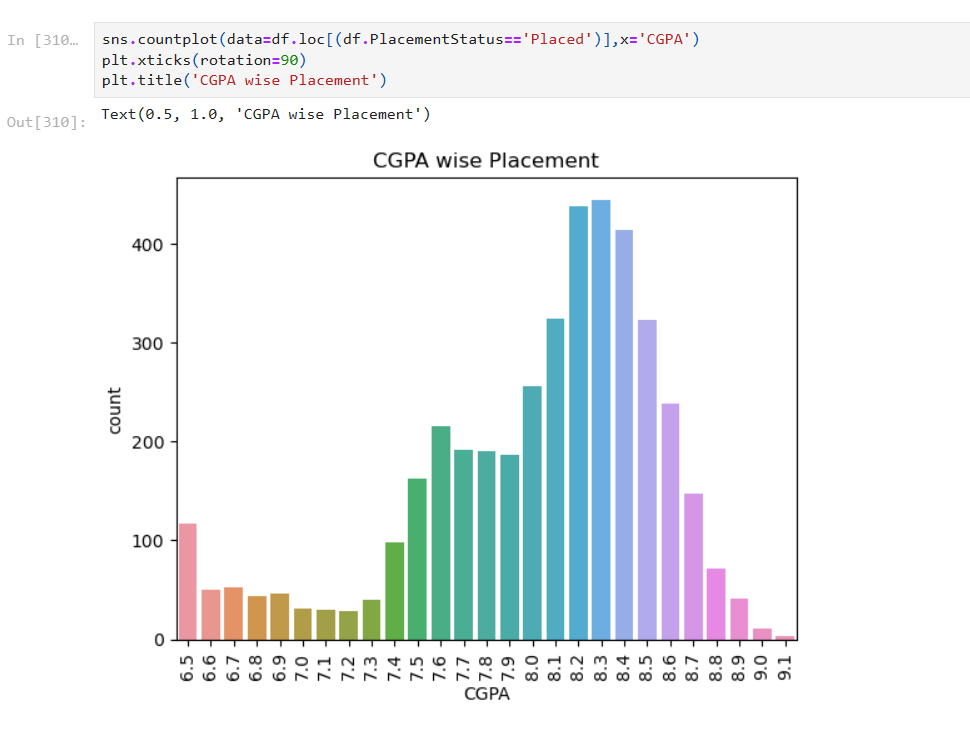
****

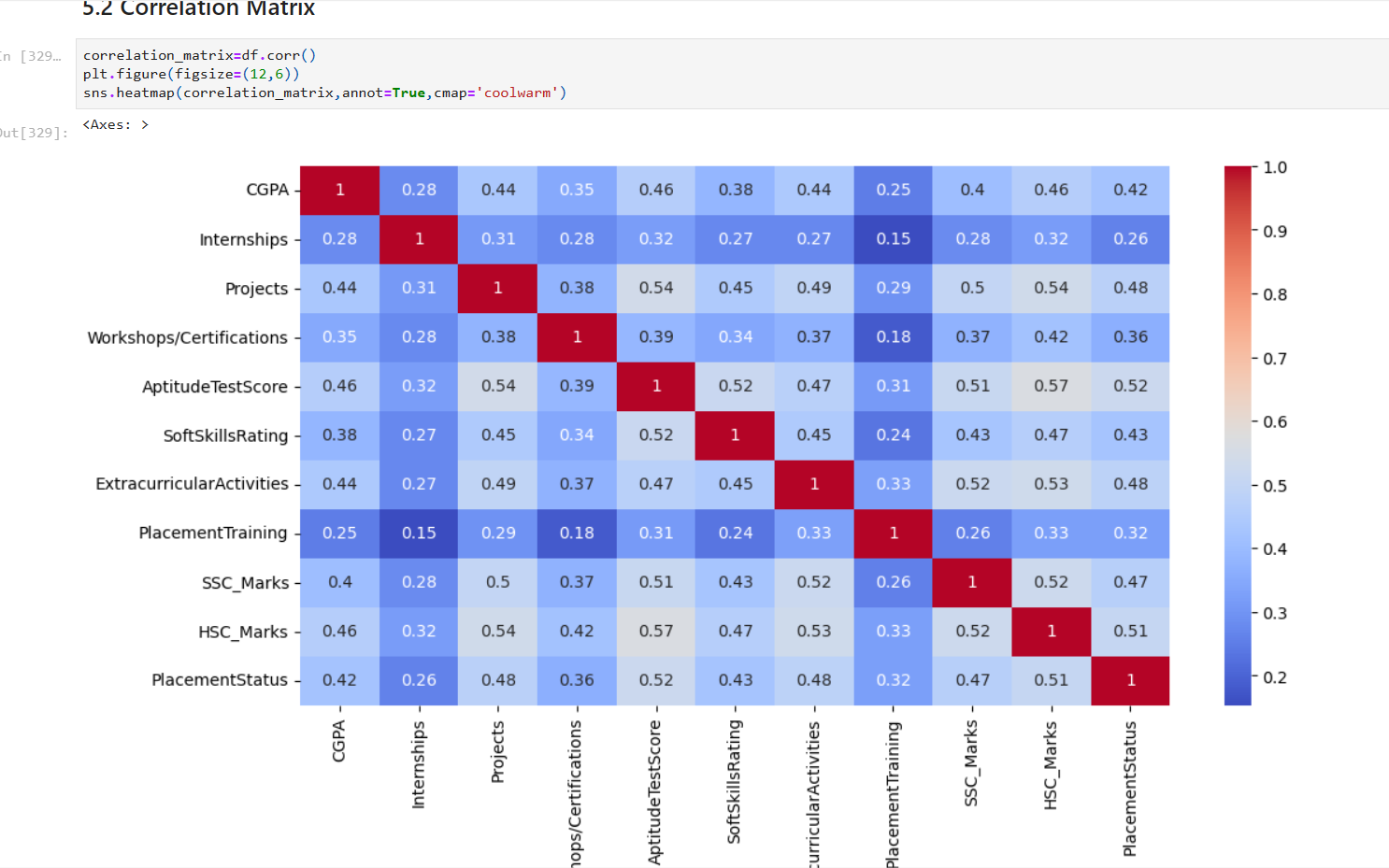
****

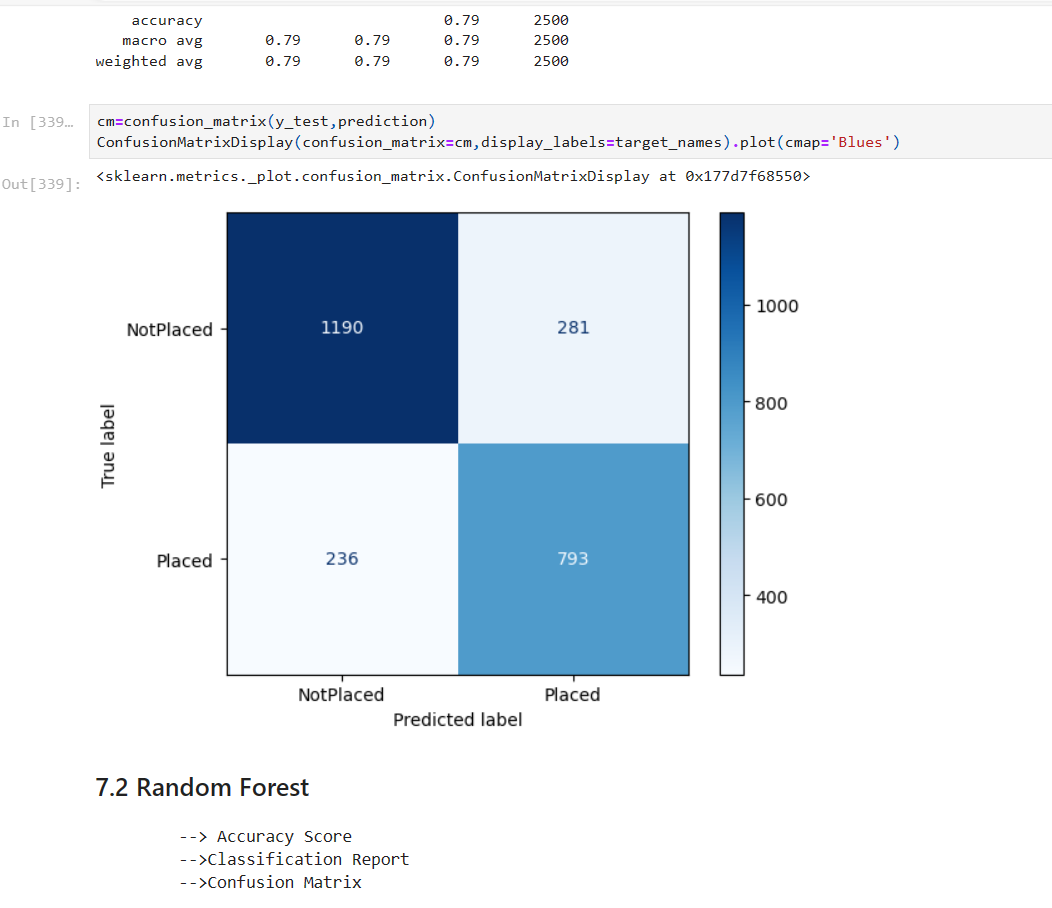
****

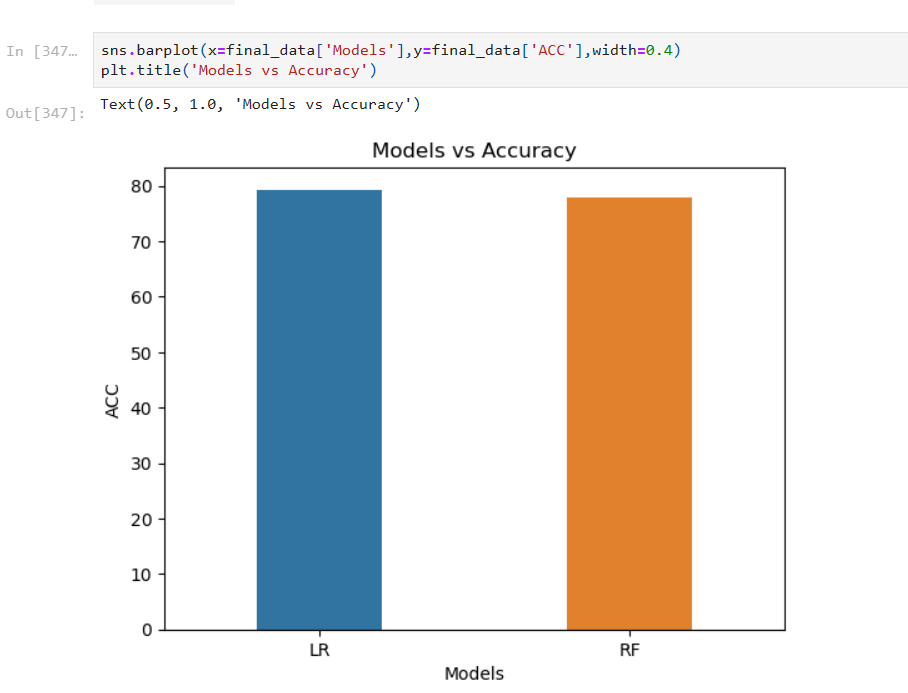
****

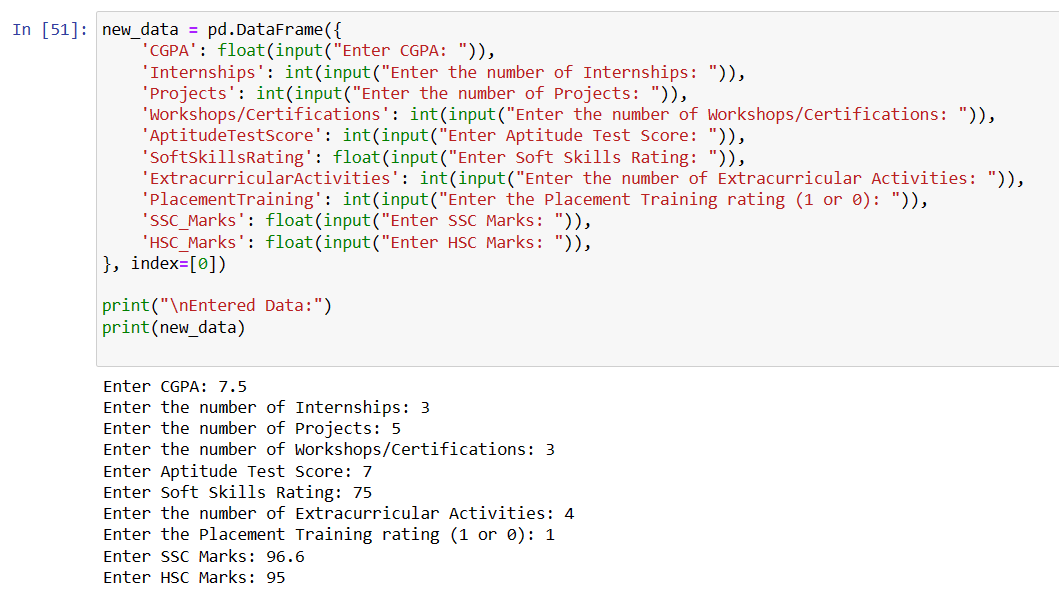
****

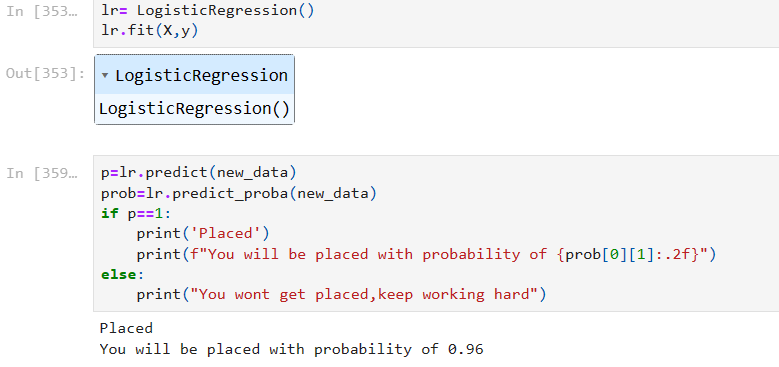
****

****

****

****

****

****

#### **Results**:

Both Logistic Regression and Random Forest show commendable performance, with Random Forest outperforming Logistic Regression in predictive accuracy. Precision and recall metrics are analyzed to understand the trade-off between correctly identifying positive cases and minimizing false positives.

#### **Learnings:**

The project provides insights into the significance of different features in predicting placements. Academic performance is a strong predictor, but extracurricular activities also play a non-negligible role. Challenges include handling imbalanced data and optimizing model parameters for better generalization.

### 

### **4. Usefulness as Subject Skills:**

#### Application and Relevance:

The project enhances understanding of applying machine learning to real-world scenarios. Skills in data preprocessing, model selection, and evaluation are directly applicable to industries where predictive analytics is valuable.

#### Future Directions:

To improve the model, future directions could involve exploring advanced algorithms, incorporating additional features, and considering temporal aspects. The project emphasizes the need for continuous data collection and model retraining to adapt to evolving trends.

In conclusion, the project not only provides a predictive tool for campus placements but also serves as a valuable learning experience, showcasing the practical application of machine learning in the education domain. The insights gained and the skills developed are transferable to various domains where predictive modeling is essential for decision-making.