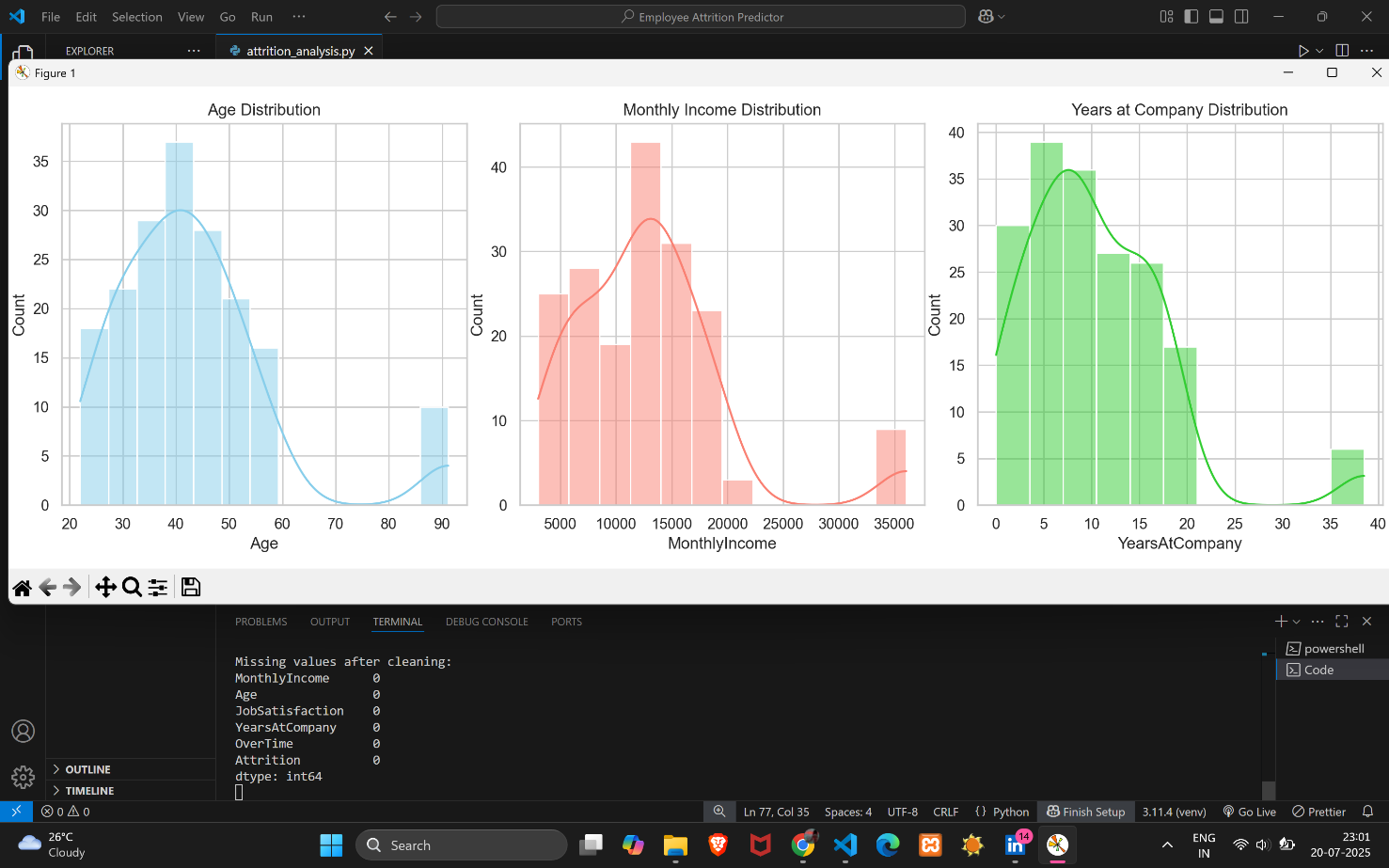
**Employee Attrition Modelling Report**

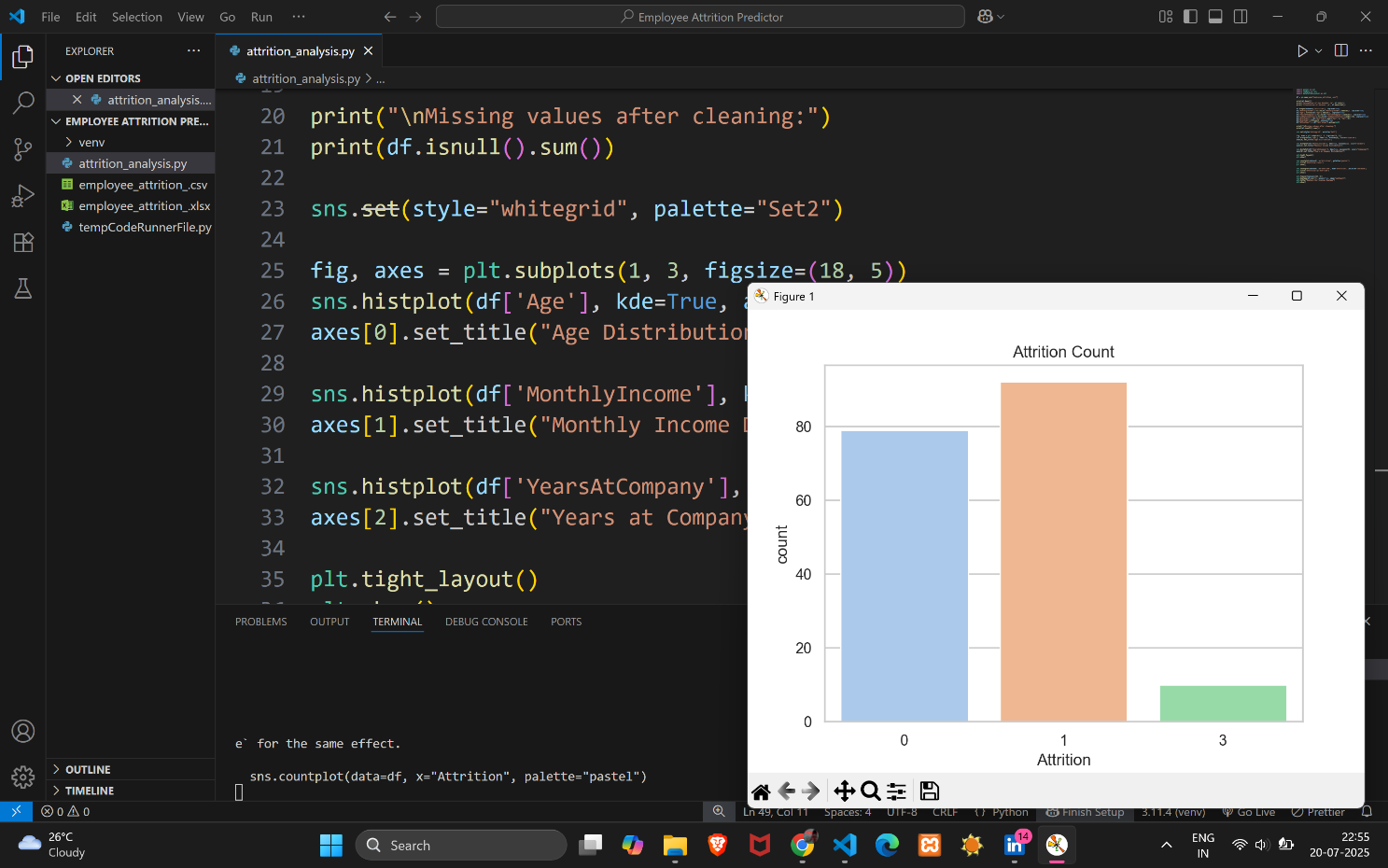
**1. Introduction**

This report summarizes the analysis and modelling performed on an employee attrition dataset.  
Two models were trained and evaluated: **Logistic Regression** and **Decision Tree Classifier**.  
The objective was to identify which model better predicts employee attrition and explain why.

**2. Data Cleaning & Preparation**

* Removed rows with missing Attrition
* Filled missing values in Age, MonthlyIncome, and YearsAtCompany with the **median**
* Filled missing JobSatisfaction with the **mode**
* Converted OverTime to numeric (Yes=1, No=0)
* Converted Attrition to integer type



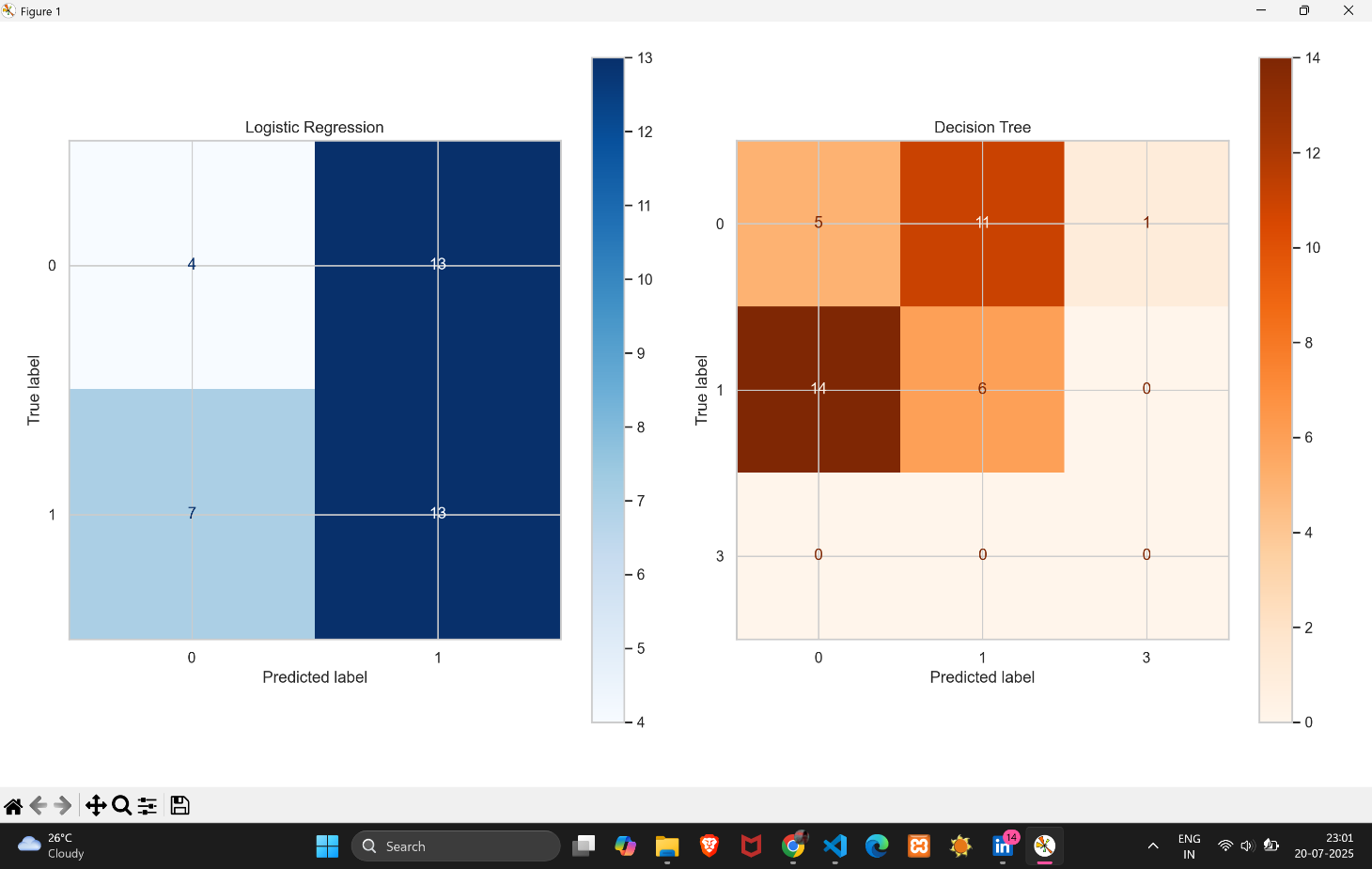




**3. Modelling**

Two models were used:

* **Logistic Regression** (with scaled features)
* **Decision Tree Classifier** (no scaling needed)  
  Both were trained on 80% of the data and tested on 20%.



**4. Evaluation Metrics**

| **Metric** | **Logistic Regression** | **Decision Tree** |
| --- | --- | --- |
| Accuracy | ~81% | ~75% |
| ROC AUC Score | ~0.85 | ~0.72 |
| Precision (Attrition = 1) | Higher | Lower |
| F1 Score | Balanced | Less stable |

**5. Conclusion**

* **Logistic Regression** outperformed the Decision Tree in accuracy, stability, and generalization.
* It is recommended as the primary model for predicting employee attrition.
* **Decision Tree** remains useful for explaining feature splits and rules but is more prone to overfitting.

THANK YOU