🎯 Student Performance Analysis System

📦 Libraries Used & Purpose

Below are the main libraries you should use (and already partially used).

🐼 1. Pandas

✅ Purpose:

Used for **data loading, cleaning, and manipulation**.

✓ What it does in your project:

- Read CSV dataset
- Handle missing values
- Filter student records
- Create new columns (average, grade, risk level)

💻 Example:

```
import pandas as pd


df = pd.read_csv("student_performance.csv")

df["Average"] = (df["Math"] + df["Science"] + df["English"]) / 3
```

📌 Why important:

Pandas is the backbone of any Data Science project.

📊 2. Matplotlib

✅ Purpose:

Used for **data visualization** (graphs).

✓ What it does:

- Bar charts for marks
- Line chart for progress
- Histogram for distribution

💻 **Example:**

```
import matplotlib.pyplot as plt


plt.bar(df["Name"], df["Average"])
plt.show()
```

📌 Why important:

Helps to visually understand student performance.

📈 **3. Seaborn (Optional but powerful)**

✅ **Purpose:**

Advanced beautiful charts.

✓ **Used for:**

- Correlation heatmap
- Boxplots
- Performance trends

🤖 **4. Scikit-learn (Machine Learning)**

✅ **Purpose:**

Build prediction model.

✓ **What it does:**

- Encode categorical data
- Scale values
- Train model
- Predict performance

Libraries inside sklearn:

Tool	Use
LabelEncoder	Convert text → numbers

Tool	Use
StandardScaler	Normalize data
train_test_split	Split dataset
LinearRegression	Prediction model
RandomForest	Advanced ML

💻 **Example:**

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

📌 Why important:

Used to predict future student scores.

🌐 **5. Streamlit**

✅ **Purpose:**

Build **web app UI** without frontend coding.

✓ **What it does:**

- Upload CSV file
- Display tables
- Show charts
- Accept inputs
- Show predictions

💻 **Example:**

```
import streamlit as st

st.title("Student Performance Dashboard")
st.dataframe(df)
```

📌 Why important:

Converts your ML project into a real application.

🔢 6. NumPy

✅ **Purpose:**

Mathematical operations and arrays.

🏗️ How Your Project Works (Flow)

CSV Dataset

↓

Pandas cleans data

↓

Visualization (Matplotlib / Seaborn)

↓

ML Model Training (Sklearn)

↓

Prediction Output

↓

Streamlit UI Display

🎯 What You Can Say in Interview

"I built a Student Performance Analysis System using Python.
Pandas is used for data preprocessing, Matplotlib and Seaborn for visualization, Scikit-learn for machine learning model training and prediction, and Streamlit for creating a web dashboard interface."