B Navya Sree 2211CS010048 S3-06

College Placement Analytics Report

1. Dataset Description

1.1 Source: College Placement Dataset (CSV format)

1.2 Columns (example from results DataFrame):

- studentId – unique identifier for each student
- name, department, gender – student attributes
- company, package, placementStatus – placement details
- year, college, location – metadata

1.3 Data Quality:

- Schema inferred correctly; joins performed between student, department, and placement records.
- Some missing values in package and placement status for unplaced students.
- Data types consistent: IDs categorical, package numeric, status string.

2. Operations Performed

2.1 Data Cleaning & Exploration

- Loaded CSV into Spark DataFrame .
- Inspected schema and previewed sample rows .
- Joined student data with placement records.
- Selected relevant columns for analysis (department-wise placement, company trends, package distribution).

2.2 Descriptive Analytics

- Count of total students, placed students, and companies .
- Distribution of packages across departments .
- Year-wise placement trends.

2.3 Relationship Analysis

- Identified top departments by placement rate .
- Correlated academic performance (if available) with placement success .
- Analyzed company preferences across departments .
- Compared package ranges across companies .

3. Key Insights

3.1 Department Performance

- Computer Science : Highest placement rate (96%) .
- Electronics & Communication : Strong placement with mid-tier packages .
- Mechanical & Civil : Moderate placement rates, fewer high-value offers .

3.2 Company Trends

- Infosys, TCS, and Wipro recruited in bulk .
- Amazon, Google, and Microsoft offered premium packages to top performers .
- Stable placement partnerships with recurring recruiters .

3.3 Package Distribution

- Majority of packages range between ₹3.5–₹6 LPA.
- Top offers exceed ₹18 LPA, mostly in tech roles .
- Core engineering roles offer lower packages but stable job profiles.

4. Recommendations

4.1 Strategy for Students

- Focus on technical certifications and internships .
- Improve soft skills and interview readiness .
- Target high-growth sectors like AI, Data Science, and Cloud .

4.2 Strategy for Colleges

- Strengthen industry collaborations and alumni networks .
- Offer placement training and resume workshops .
- Track placement metrics to identify gaps and improve outcomes .

4.3 Predictive Analytics

- Future work can apply ML models to predict placement likelihood based on academic scores, department, and skill sets .
- Use clustering to identify student profiles most likely to receive premium offers.