**Cloud Consulage Internship – Development Log**

Name:Parth Satvekar  
Domain of Internship: Python

**Project Overview**

This project focuses on analyzing student data from the Cloud Consulage internship program. Using Python (pandas, seaborn, matplotlib), the objective is to explore patterns in academic performance, family income, technical skills, and event participation through descriptive statistics and visualizations.

**Initial Plan**

* Load and clean the dataset.
* Map categorical fields such as family income into numeric values.
* Compute key metrics including average CGPA, income levels, and expected salary.
* Visualize distributions and relationships between variables.
* Extract insights related to skills, college performance, and event participation.

**Development Steps**

* Imported required libraries (pandas, seaborn, matplotlib) and loaded the dataset.
* Mapped family income ranges to numeric values for analysis.
* Computed descriptive statistics: unique student count, mean CGPA, average family income.
* Generated visualizations: bar plots (graduation year, city-wise CGPA, top colleges), histograms (Python experience), boxplots (attendee status vs quantity), scatterplots (family income vs CGPA), and pairplots.
* Calculated correlations: family income vs CGPA, leadership skill vs expected salary.
* Identified specific groups: high CGPA + high Python experience, data science event participants, top colleges.
* Refined code by renaming variables, adjusting plot colors, and improving readability.

**Challenges & Solutions**

* Managed inconsistent income category labels by applying a manual mapping.
* Enhanced plot readability by rotating axis labels and choosing clearer color schemes.

**Future Enhancements**

* Integrate interactive visualizations using Plotly or similar libraries.
* Extend analysis with additional datasets or live data integration.
* Package the analysis as a web application for broader accessibility.