**Company that offered the internship:** The Sparks Foundation

**Nature and scope of the project:** During this internship, I worked on two main tasks:

1. I built a **regression model** to predict how many marks a student would score based on how many hours they studied.
2. I also analyzed **IPL (Indian Premier League)** cricket data to find patterns in player performances and factors that help teams win matches.

**Statistical techniques used:**

* **Linear Regression** to build the student marks prediction model.  
  **Exploratory Data Analysis (EDA)** to explore and summarize the IPL data.  
  Used **Python tools** like Pandas for data handling and Matplotlib for creating visualizations.

**Findings and recommendations:**

* The student marks prediction model had a high accuracy, with an **R² score of 94.7%**, showing a strong link between study time and marks. This model improved prediction accuracy by **15%** compared to the baseline.
* From the IPL data analysis, I uncovered trends such as which players perform best under certain conditions and what factors contribute most to winning a match. These insights can help teams make better decisions during player selection and game strategies.

### 

### **Apart from the internship, I have done some ML projects: 1. Facebook Marketplace Analysis**

**Nature and scope of the project:** I analyzed data from **7,050 Facebook Marketplace posts** to understand what kind of posts get more attention — like more likes, comments, and shares.

**Statistical techniques used:**

* **Exploratory Data Analysis (EDA)** to find patterns in the data.  
  **Correlation Analysis** to see which features (like number of comments or reactions) are related.
* **K-Means Clustering** (using the Elbow Method) to group posts into different segments based on their engagement levels.

**Findings and recommendations:** The project helped identify which types of posts perform better. This insight can help sellers or content creators optimize their posts for better engagement.

### **2. Advertising Impact on Sales**

**Nature and scope of the project:** I built a machine learning model to see how advertising on **TV, radio, and newspapers** affects product sales.

**Statistical techniques used:**

* **Linear Regression** to predict sales based on ad spending.
* **Correlation Analysis** to check which medium had the most impact.
* **Data Normalization** to improve model performance.

**Findings and recommendations:** The model showed how each type of advertising contributes to sales. This can help companies decide where to invest their advertising budget for the best results.