**Project Report**

**Introduction**  
This project was undertaken as part of a one-month internship with Null Class Pvt. Ltd. The core objective was to apply data analytics and machine learning approaches to real-life sales data, aiming to derive actionable business insights and forecast future performance.

**Background**  
The analysis revolved around historical sales data, targeting two main business use cases:

1. Forecasting total sales for Q2 of 2023 based on sales trends from 2022.
2. Examining how varying levels of discounts influence overall revenue.

**Learning Objectives**

* Gain hands-on experience in time series forecasting using historical datasets.
* Develop and assess regression models using Python.
* Analyze the relationship between discount tiers and revenue outcomes.
* Use Google Looker Studio to communicate insights through interactive visuals.

**Activities and Tasks**

* Performed data cleaning and preprocessing on key fields such as order\_date, qty\_ordered, before\_discount, and discount\_amount.
* Aggregated sales by quarterly periods for the year 2022.
* Built a linear regression model using Scikit-learn to predict Q2 2023 sales.
* Assessed model performance with metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).
* Categorized discounts into four brackets: None, Low, Medium, and High.
* Designed and implemented visualizations (bar charts, pie charts, scorecards) in Looker Studio.
* Developed interactive dashboards for stakeholders to monitor key performance indicators (KPIs).

**Skills and Competencies Acquired**

* Proficiency in Python libraries such as Pandas, NumPy, Matplotlib, and Seaborn.
* Regression modeling using Scikit-learn.
* Data wrangling and transformation techniques.
* Visualization and dashboard creation using Google Looker Studio.
* Analytical interpretation of sales data and discount patterns.

**Feedback and Evidence**

* The predicted sales for Q2 2023 were approximately **1802.90 units**.
* Looker Studio scorecards and visual KPIs effectively communicated Q2 performance.
* Discount-level analysis revealed that larger discounts often led to higher sales volumes before the discount was applied.

**Challenges and Solutions**

* **Challenge:** Inconsistent data formats.  
  **Solution:** Used Pandas for standardized date parsing and data cleaning.
* **Challenge:** Creating discount-based segments.  
  **Solution:** Implemented custom logic in Python and utilized calculated fields in Looker Studio for flexible binning.

**Outcomes and Impacts**

* Delivered a functional model for forecasting sales based on historical data trends.
* Uncovered meaningful insights into customer behavior under different discount conditions.
* Created a professional-level interactive dashboard to support data-driven business decisions.

**Conclusion**  
This internship project offered practical exposure to solving real-world data problems. It enhanced both technical proficiency and business analytics skills. Overall, the project demonstrates a full analytics pipeline—from raw data to strategic insights—supporting improved sales planning and decision-making.