

**Report
On
Project 1**

**“COFFEE SALES PREDICTION USING
MACHINE LEARNING”**

ABSTRACT

Coffee sales from vending machines vary with time, day, and consumer preferences. This project utilizes machine learning techniques to analyze historical transaction data and predict coffee sales trends. The aim is to support inventory optimization, pricing strategies, and improved customer satisfaction. Key insights into peak sales periods and top-selling products provide valuable business intelligence for decision-making.

Table Of Contents

1. Introduction

1.1 About the Project

1.2 Objective and Deliverables

2. Methodology

2.1 Tools and Technologies Used

2.2 Data Preparation

3. Implementation

3.1 Model Building

3.2 Model evaluation

4. Exploratory Data Analysis

4.1 Graphs

4.2 Key Insights

5. Business Applications

6. Conclusion

7. Future Scope

1. INTRODUCTION

1.1 About the Project

This capstone project focuses on analyzing coffee sales from vending machine logs using predictive modeling. With rising demand for automated services, understanding sales behavior becomes vital for efficient operations.

1.2 Objective and Deliverables

- ❖ Predict future coffee sales using machine learning.
- ❖ Identify peak sales hours and days.
- ❖ Evaluate most and least selling coffee variants.
- ❖ Recommend business strategies based on data insights.

2. METHODOLOGY

2.1 Tools and Technologies Used

- ❖ Programming Language: Python
- ❖ Libraries: Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn
- ❖ Model Used: Linear Regression
- ❖ Data Visualization: Matplotlib & Seaborn charts

2.2 Data Preparation

- ❖ Extracted features: Month, Day, Hour, Coffee Type, Cash/Card
- ❖ Cleaned missing and inconsistent data
- ❖ Split data: 80% for training, 20% for testing

3. IMPLEMENTATION

3.1 Model Building

- ❖ Applied Linear Regression to predict sales.
- ❖ Trained model on historical data with sales volume as the target.
- ❖ Evaluated using MSE and R^2 score.

3.2 Model Evaluation

- ❖ Mean Squared Error (MSE): **0.79** (low)
- ❖ R^2 Score: **0.96** (high accuracy)
- ❖ Important Predictors: Time of purchase, coffee type, transaction mode

```
PS E:\AICETE & Edunet> & C:/Users/hp/AppData/Local/Programs/Python/Python313/python.exe "e:/AICETE & Edunet/coffee_sales.py"
Model Evaluation:
Mean Squared Error: 0.79
R2 Score: 0.96

Model Coefficients:

```

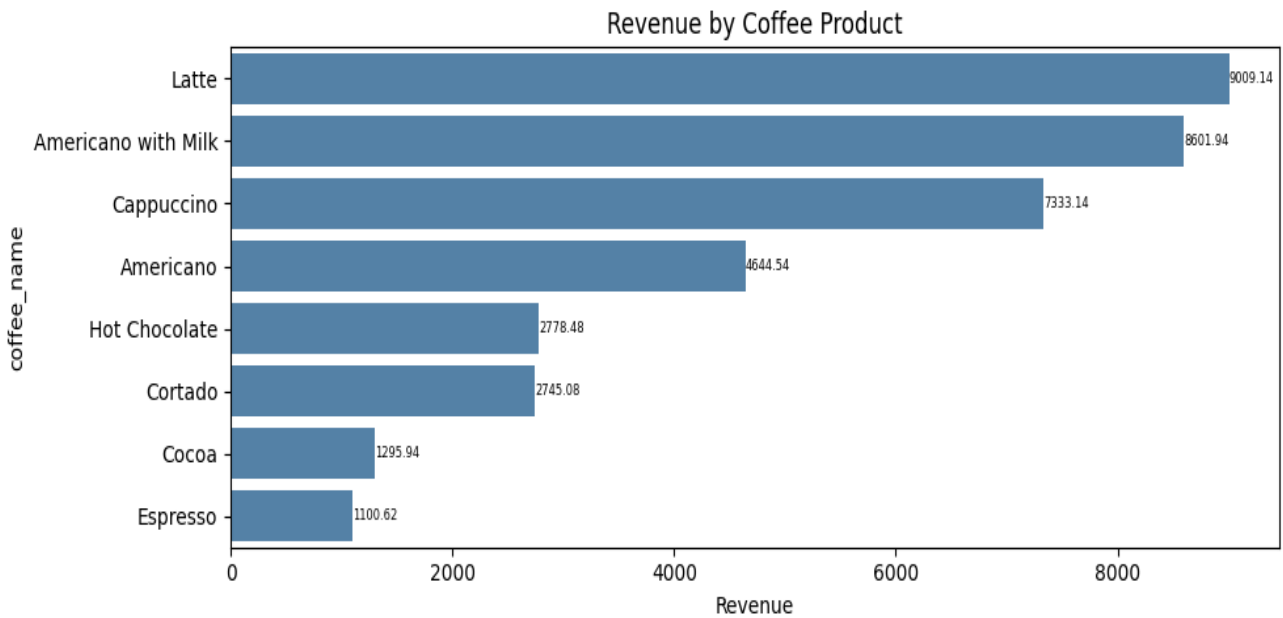
	Feature	Coefficient
28	coffee_name_Cocoa	9.827475
27	coffee_name_Cappuccino	9.812987
32	coffee_name_Latte	9.781796
31	coffee_name_Hot Chocolate	9.753614
26	coffee_name_Americano with Milk	4.935424
25	cash_type_cash	1.323496
18	hour_16	0.570958
24	hour_22	0.455914
20	hour_18	0.347354
12	hour_10	0.346046
14	hour_12	0.290758
15	hour_13	0.288361
8	day_5	0.274694
6	day_3	0.199089
21	hour_19	0.167856
19	hour_17	0.164874
17	hour_15	0.158001
16	hour_14	0.138231
5	day_2	0.095135
7	day_4	0.089492
23	hour_21	0.074238
13	hour_11	0.062855
9	day_6	0.053696
22	hour_20	0.053190
4	day_1	0.045242
10	hour_08	-0.028731
29	coffee_name_Cortado	-0.090195
11	hour_09	-0.146980
0	month_2024-04	-0.317911
2	month_2024-06	-0.893711
1	month_2024-05	-0.924665
30	coffee_name_Espresso	-4.719472
3	month_2024-07	-5.275066

```
PS E:\AICETE & Edunet>
```

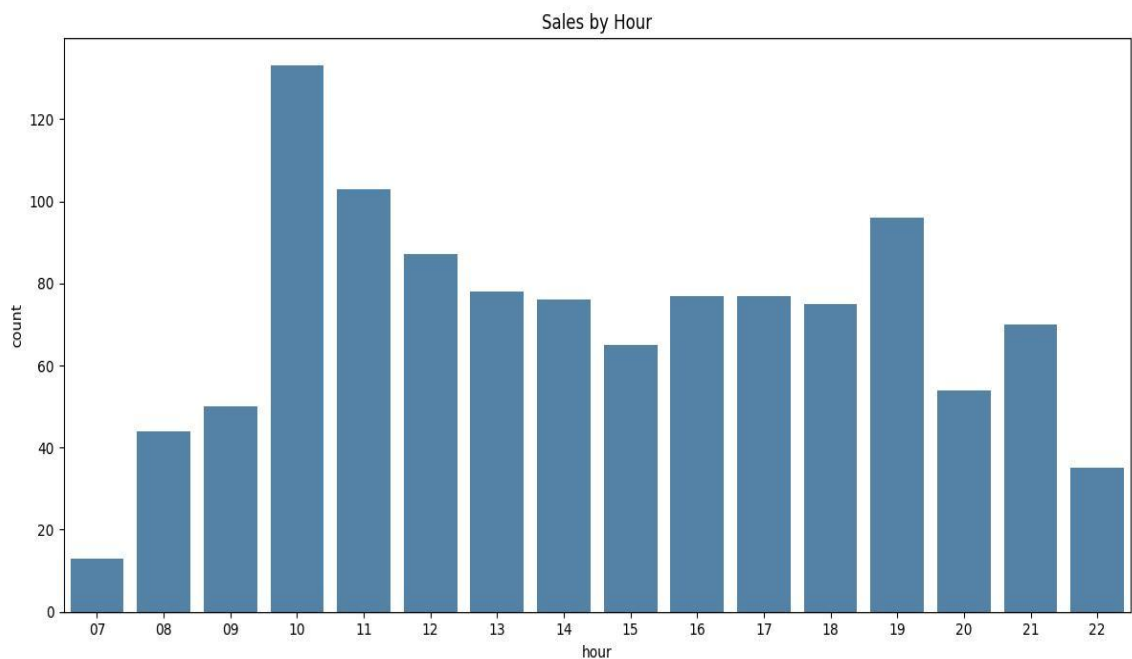
4. EXPLORATORY DATA ANALYSIS

4.1 Graphs

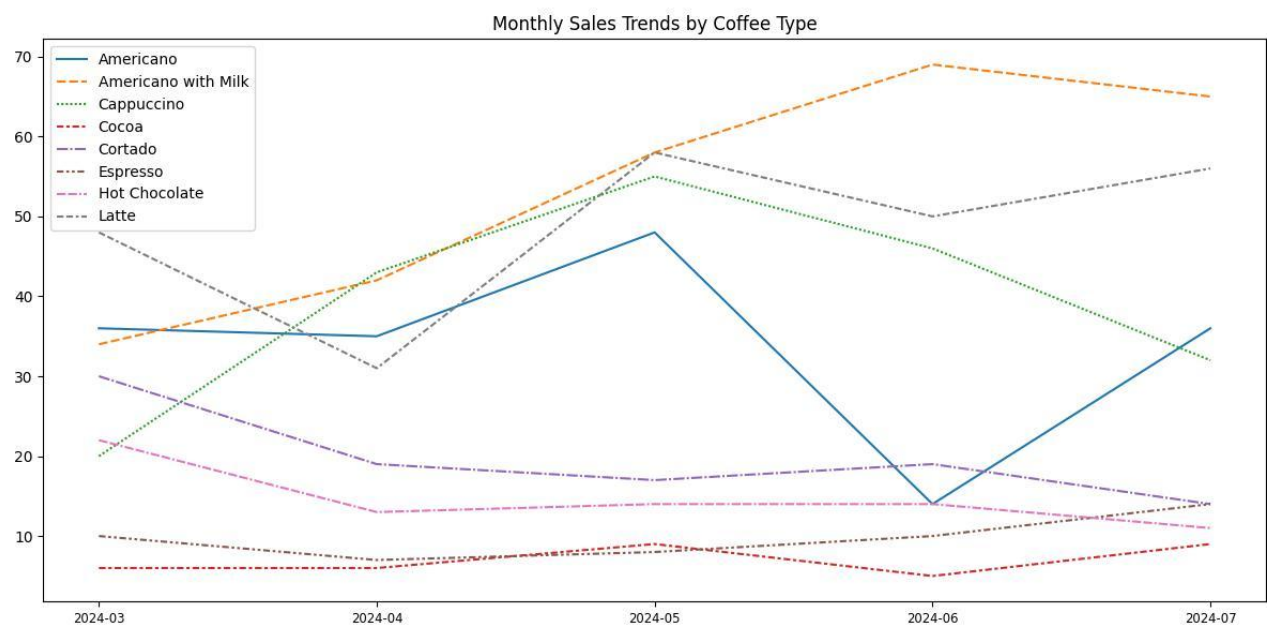
(i)



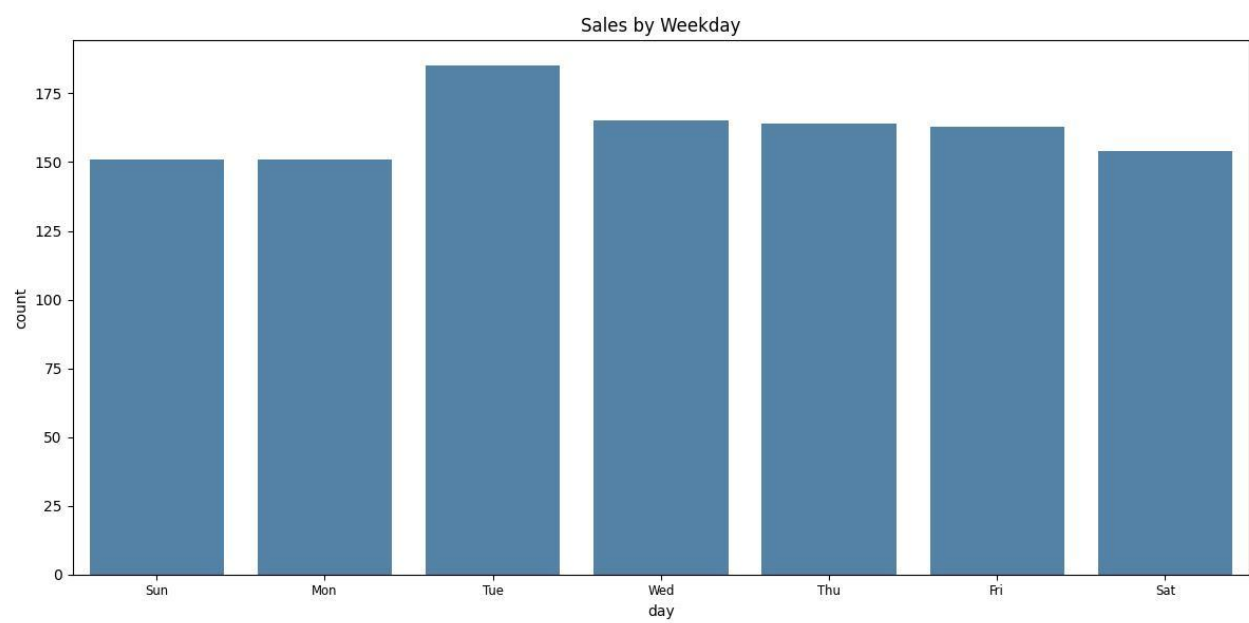
(ii)



(iii)



(iv)



4.2 Key Insights

- ❖ Peak Sales Hours: **10:00 AM, 4–7 PM**
- ❖ Weekday Impact: **Highest on Tuesday, lowest on Monday**
- ❖ Top Selling Coffees:
 - **Latte (₹9009.14)**
 - **Americano with Milk (₹8601.94)**
 - **Cappuccino (₹7333.14)**
 - Least Selling: **Cocoa and Espresso**

5. BUSINESS APPLICATIONS

- ❖ **Stock Optimization:** Refill top-selling items before peak hours
- ❖ **Dynamic Pricing:** Adjust prices during low-demand periods
- ❖ **Forecasting:** Enable demand planning using predictive analytics
- ❖ **Sales Strategy:** Promote underperforming products or bundle offers.

6. CONCLUSION

This project successfully demonstrates that sales data from vending machines can be effectively used for prediction using ML techniques. The current model, though basic, provides highly accurate forecasts and actionable insights.

7. FUTURE SCOPE

- ❖ Use advanced models like Random Forest or Neural Networks
- ❖ Deploy the model as a dashboard for real-time monitoring
- ❖ Incorporate weather, seasonality, or customer segmentation