Daily Transaction Analysis and Forecasting

A Data Science Project

Shreyash Jha

Data Scientist & CSE Engineer, 3rd Year
K.R. Mangalam University
Data Science Intern, Unified Mentor Pvt. Ltd.

July 11, 2025

About Me

- Name: Shreyash Jha
- Role: Data Scientist & Computer Science Engineer, 3rd Year
- Institution: K.R. Mangalam University
- Internship: Data Science Intern at Unified Mentor Pvt. Ltd.
- Passion: Applying data science to uncover business insights and optimize operations

Project Overview

Objective

Analyze synthetic daily transaction data for 2024 and forecast transaction amounts for the next 30 days using advanced data science techniques.

- Generate a dataset of 10,000 transactions
- Perform exploratory data analysis (EDA) to identify trends
- Forecast future transaction amounts using Prophet
- Provide actionable business insights

Data Description

- **Time Period**: January 1, 2024 – December 31, 2024
- **Data**: 10,000 synthetic transaction records
- **Columns**: Transaction ID, Date, Customer ID, Amount (5–500 USD), Category (Electronics, Clothing, Books, Groceries, Home), Payment Method (Credit Card, Debit Card, PayPal, Cash), Location (New York, Los Angeles, Chicago, Houston, Miami)
- **Source**: Simulated data stored in `daily_transactions.csv`

Sample Data

```
Transaction ID: 1, Date: 2024-11-23, Customer: CUST534,
Amount: 357.27, Category: Home, Payment: Credit Card,
Location: Houston
```

Data Preprocessing

1. **Missing Values**: Checked with `df.isnull().sum()` – none found
2. **Data Type Conversion**: Converted Date to datetime format
3. **Feature Engineering**: Added Day of Week, Month, and Transaction Count
4. **Outlier Removal**: Removed extreme values in Amount to ensure data quality

Methodology

1. **Data Generation**: Created synthetic dataset using Pandas and NumPy
2. **Exploratory Data Analysis (EDA)**: Analyzed trends by Category, Location, and Day of Week
3. **Forecasting**: Used Prophet with yearly and weekly seasonality for 30-day predictions
4. **Evaluation**: Measured model performance with Mean Absolute Error (MAE)
5. **Visualization**: Generated plots for trends and forecasts using Matplotlib

Results: Transaction Trends

Key Findings

- Fridays show 20% higher transaction volumes
- Electronics accounts for 35% of total spending
- Houston and New York are top locations for transactions

Visualization

[Image: daily_transaction_trends.png shows trends by day and category]

Forecasting Results

Prediction

- Forecasted transaction amounts for 30 days post-2024 using Prophet
- Expected 10% increase in transaction amounts in February 2025
- Model performance: Mean Absolute Error (MAE) = 1140.40
- *[Image: daily_transaction_forecast.png shows past and predicted amounts]*

Insight

Prepare for increased transaction volumes in early 2025.

Business Insights

- **Peak Days**: Fridays drive higher transaction volumes
- **Top Category**: Electronics leads spending (35% of total)
- **Key Locations**: Houston and New York dominate transactions
- **Recommendation**: Increase marketing on Fridays and stock more Electronics

Action Plan

Target promotions in high-traffic locations and optimize inventory for Electronics.

Why It Matters

- Data-driven insights enhance business decision-making
- Forecasting supports inventory and marketing planning
- Targeted strategies can boost sales and customer engagement

Learn More

Explore data science applications at:
`https://www.datasciencecentral.com/`

Conclusion

Key Takeaways

- Analyzed 10,000 transaction records for 2024
- Identified peak days (Fridays) and top category (Electronics)
- Forecasted 10% increase in transaction amounts (MAE: 1140.40)
- Provided actionable insights for business optimization

Future Work

Incorporate customer segmentation and fraud detection for deeper insights.

Thank You!

Shreyash Jha
Data Scientist & CSE Engineer, K.R. Mangalam University
Data Science Intern, Unified Mentor Pvt. Ltd.
Questions? Contact me at: 9205576974
shreyashjhay43@gmail.com