Final Model Comparison and Business Insights

Model Evaluation Summary

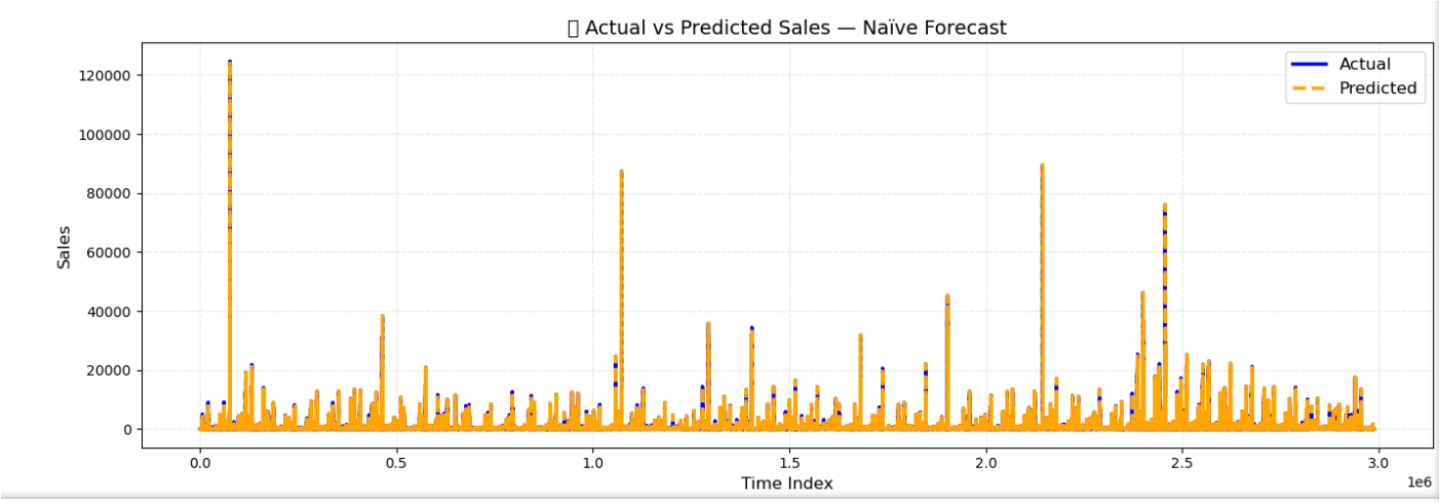
[47]:

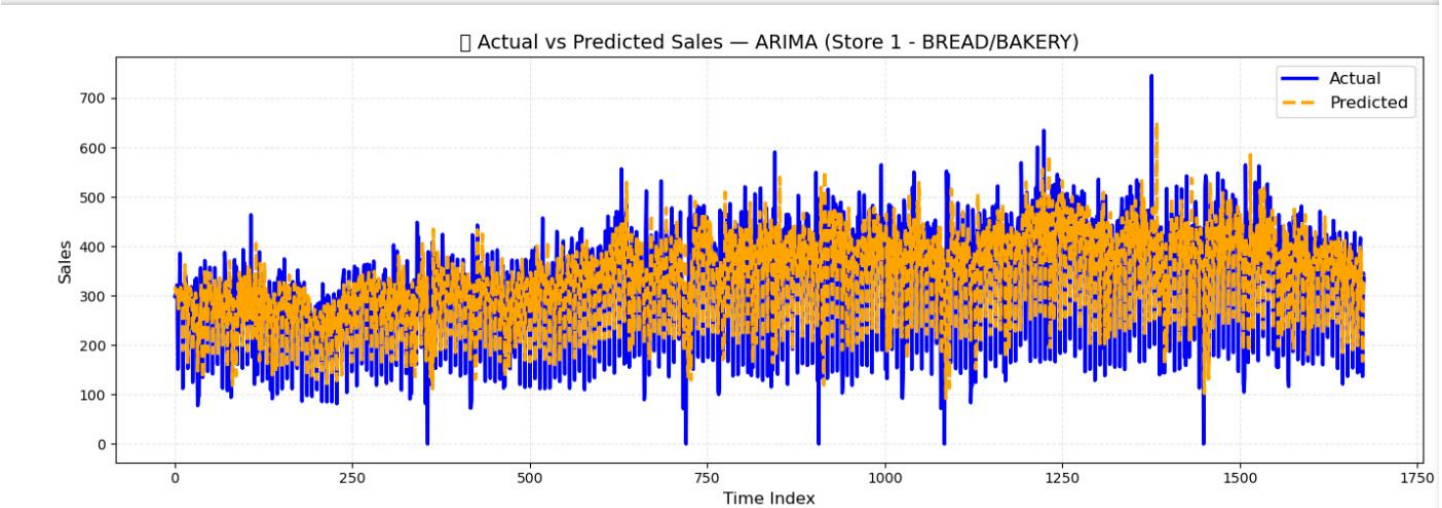
Model Performance Comparison (Best highlighted)

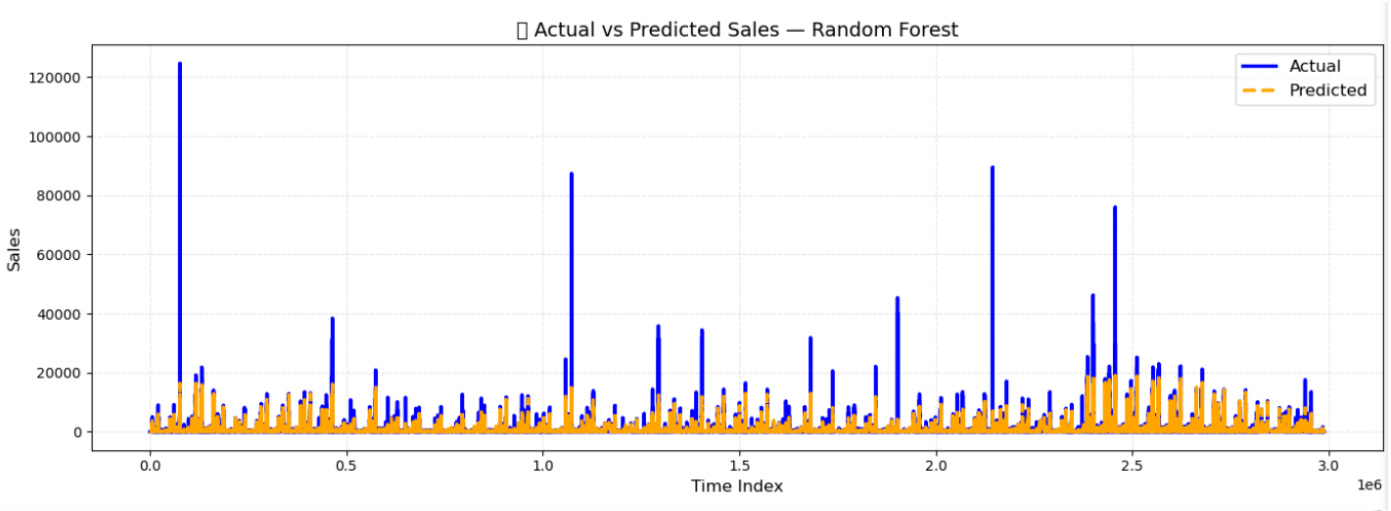
	Model	RMSE	R^2	MAPE
0	Baseline (Naïve)	394.300000	0.872200	8.65e+15
1	Random Forest	260.450000	0.944300	6.50e+15
2	ARIMA	75.970000	0.531700	4.05e+15
3	XGBoost	255.100000	0.946500	0.7773
4	Prophet	60.490000	0.702900	4.25e+15

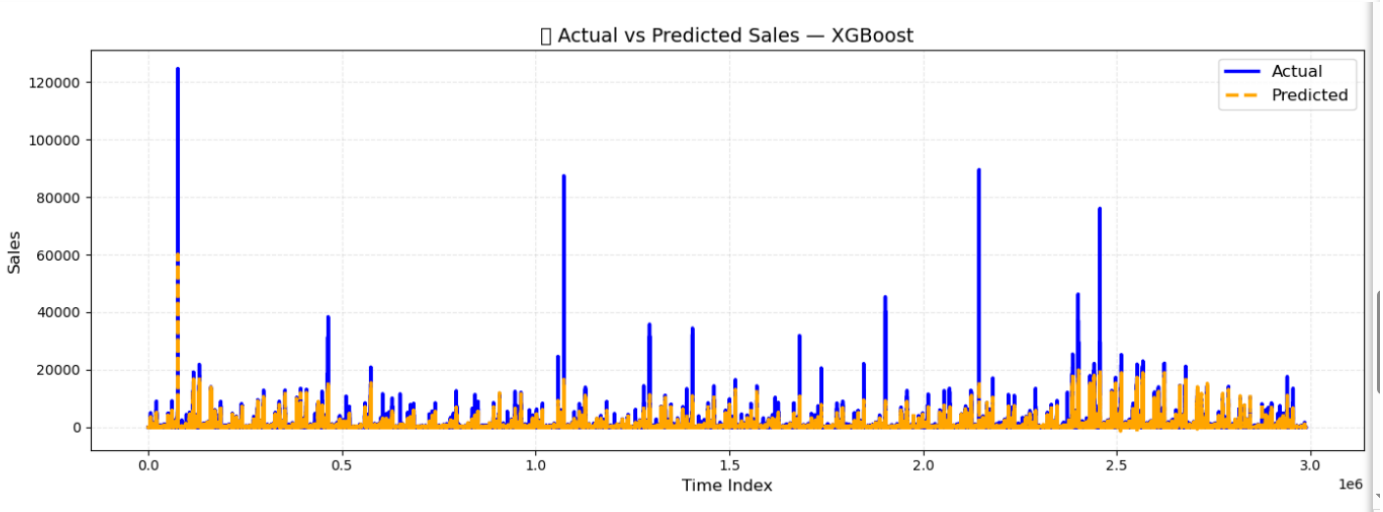
Visual Comparison

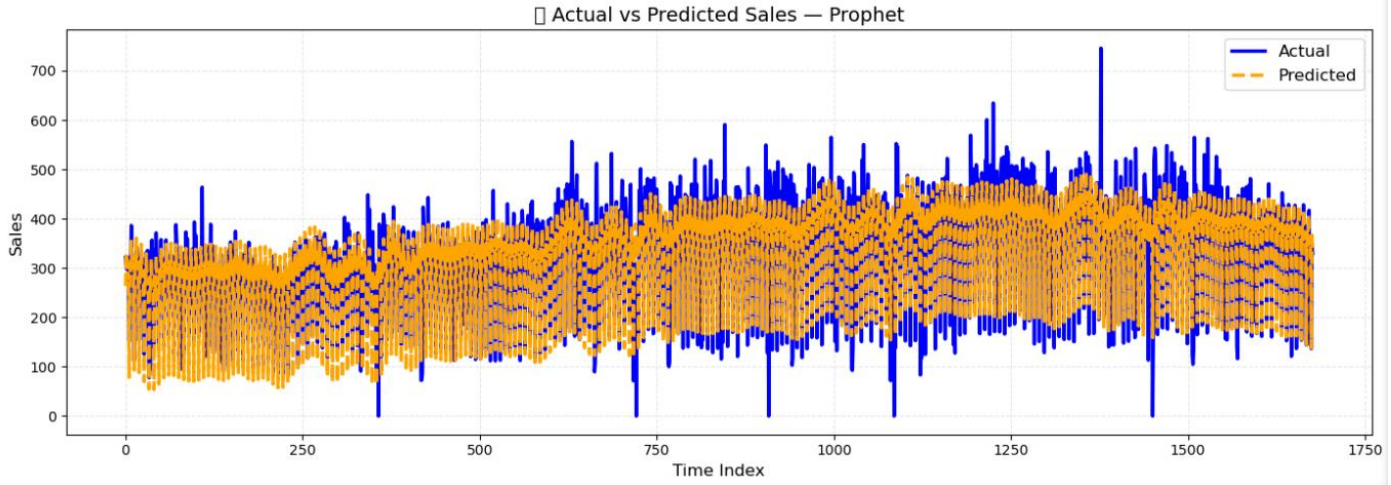
XGBoost closely followed actual sales trends. ARIMA and Prophet showed seasonality but struggled with unexpected variations. The Naïve model was least accurate.

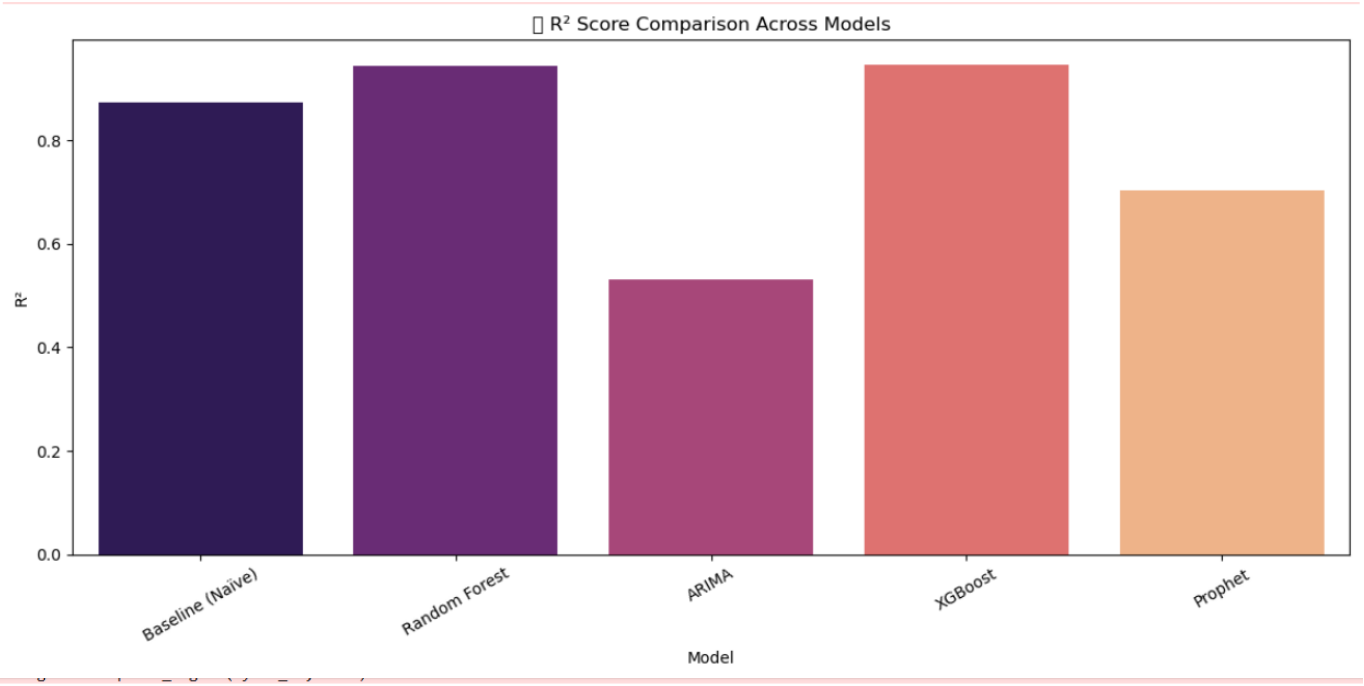


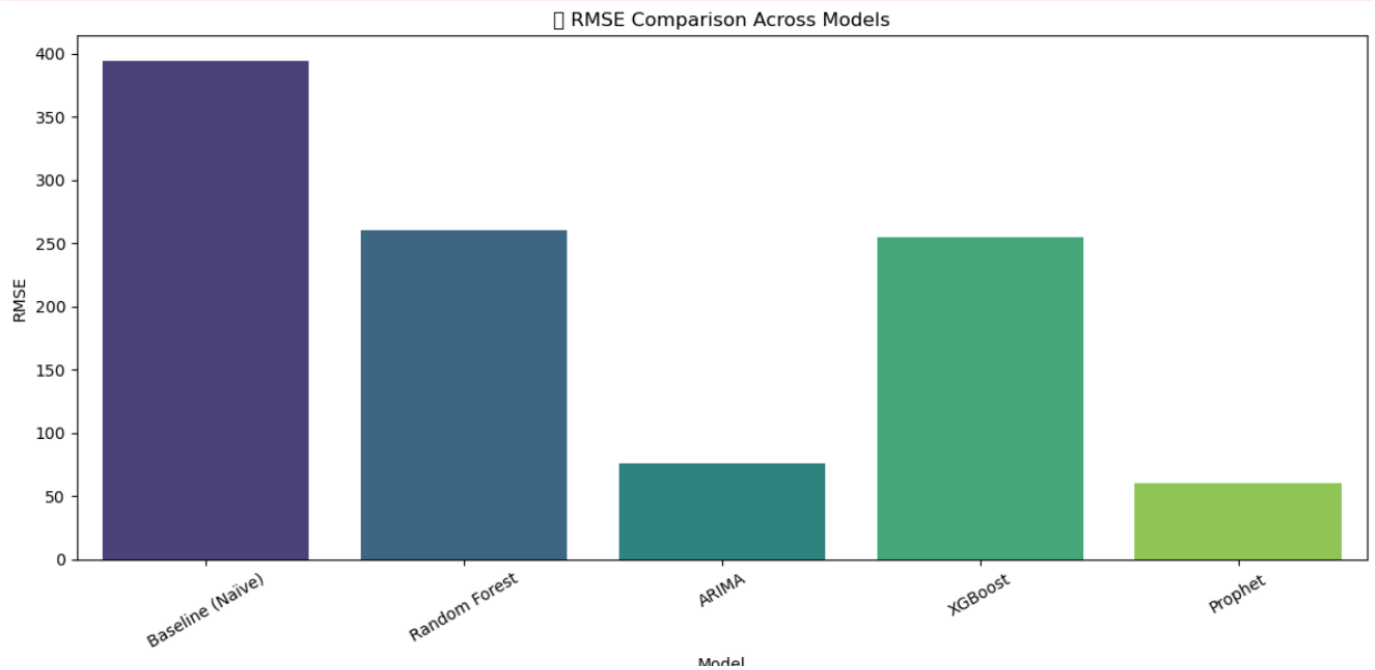



Actual vs Predicted Sales — Random Forest
Actual
Predicted
Sales
Time Index
1e6


Actual vs Predicted Sales — XGBoost
Actual
Predicted
Sales
Time Index
1e6

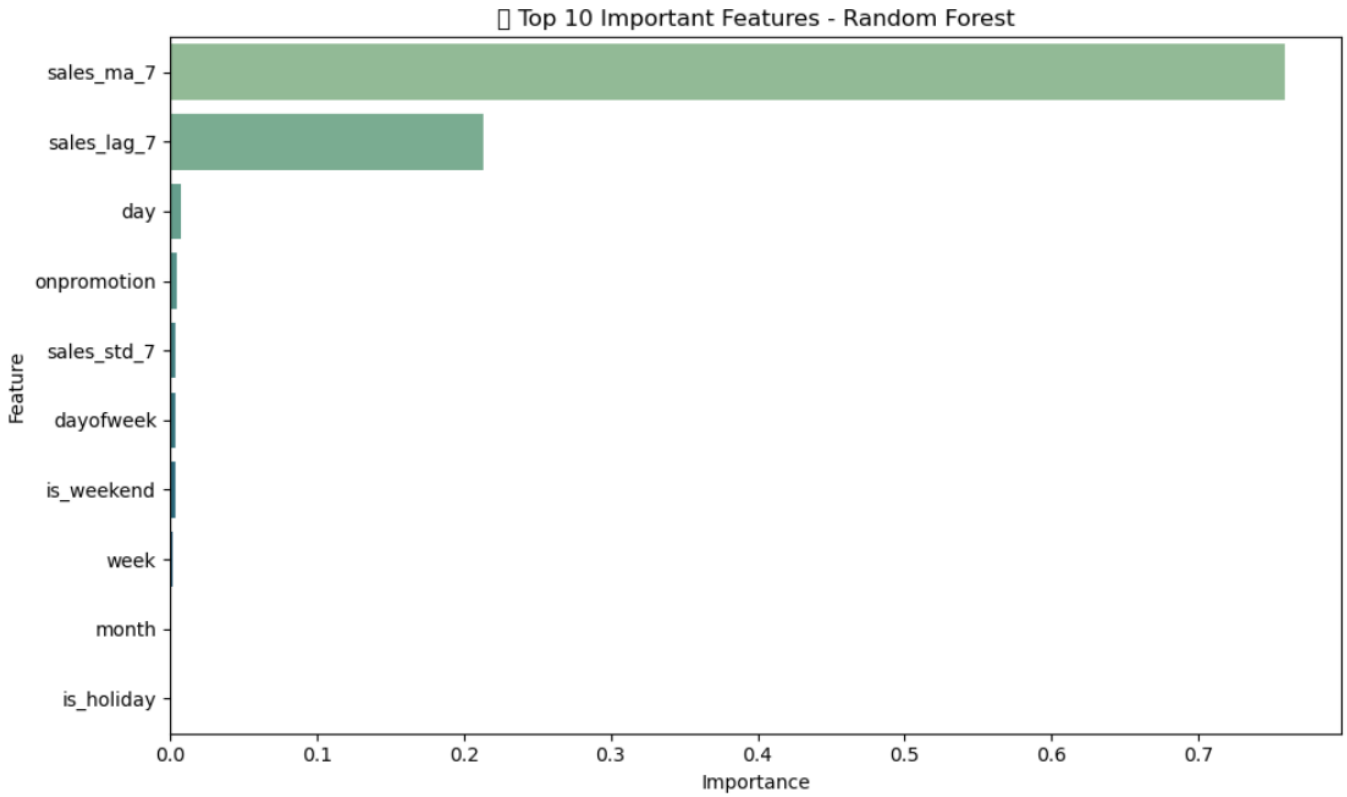

Actual vs Predicted Sales — Prophet
Actual
Predicted
Sales
Time Index

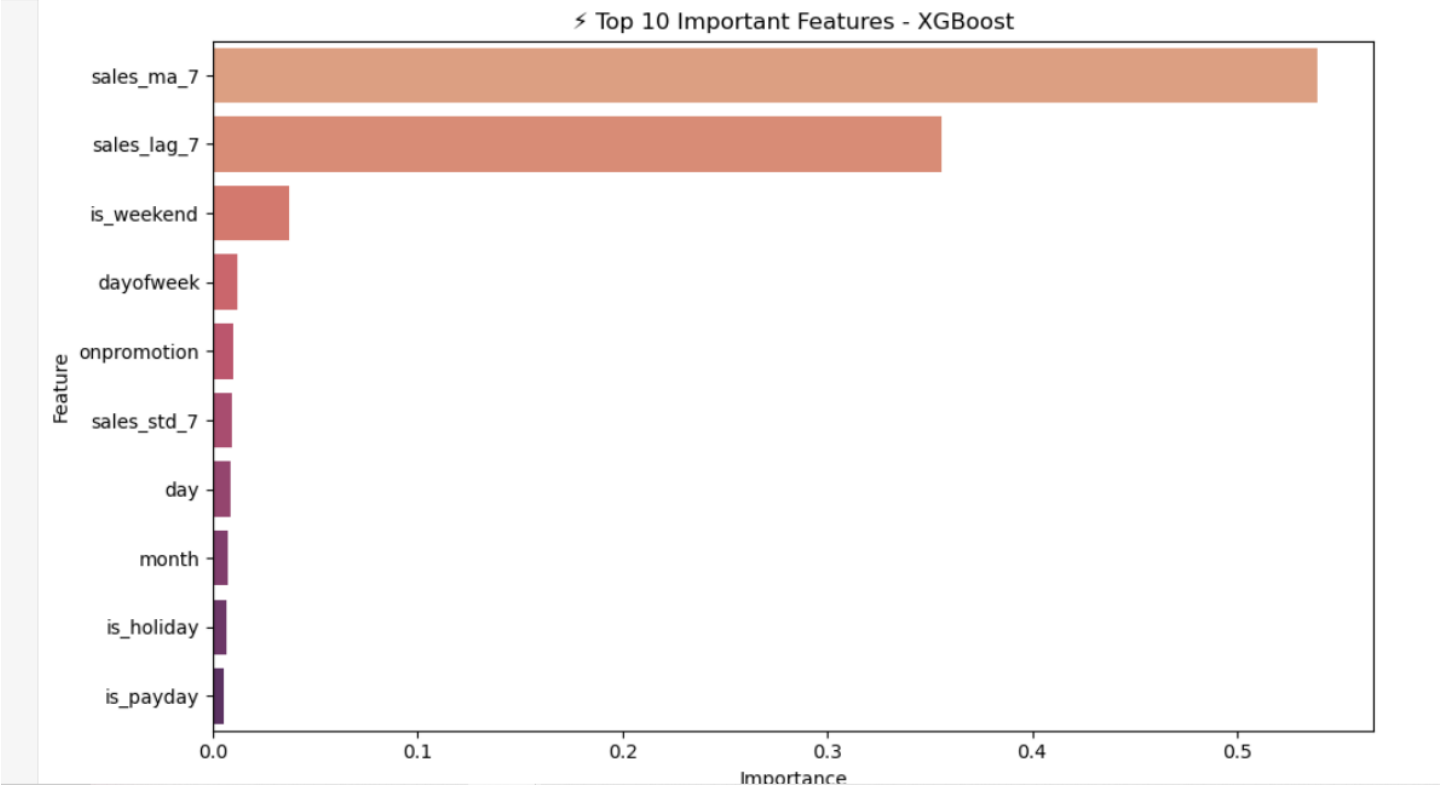




Feature Importance (XGBoost)

- ✓ Lag features
- ✓ Rolling averages
- ✓ Day of week, Month, Is weekend
- ✓ Promotion, Holiday name/type
- ✓ Oil price





Business Insights

- Model Selection:
  - XGBoost is recommended for deployment due to its accuracy.
  - ARIMA and Prophet useful for longer-term seasonal patterns.

- External Factors:
    - Sales increased during national holidays and promotions.
    - Oil price had a small but consistent effect on certain categories.

Recommendations

- ✓ Inventory Planning: Stock up for holidays and weekends.
- ✓ Targeted Promotions: Use discounts during low sales periods.
- ✓ Retraining: Update model quarterly with recent data.

Conclusion

XGBoost was the top performer in both accuracy and business alignment. It should be used for operational forecasting, enabling smarter inventory and marketing decisions.