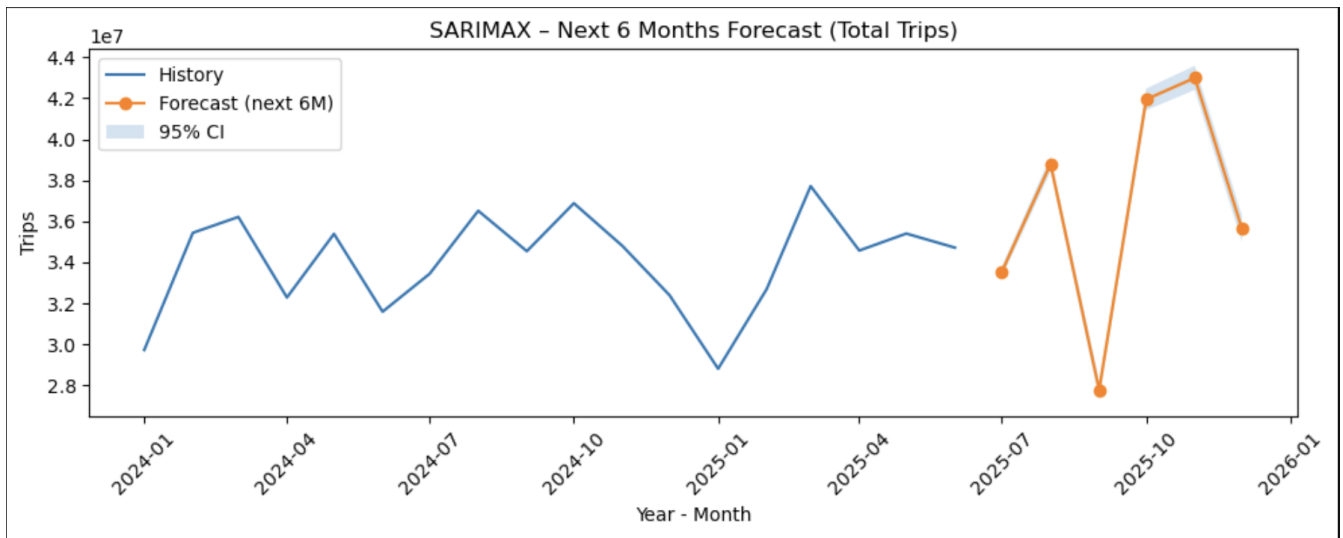# Public Transport Trip Forecasting

Business Summary – Linear Regression vs SARIMA vs SARIMAX

SARIMAX – Next 6 Months Forecast (Total Trips)

Model Performance Comparison (Test Window)

Model	Order (p,d,q)	Seasonal (P,D,Q,s)	MAE	RMSE	MAPE / R²
Linear Regression	—	—	782,018	1,196,300	R²: 0.572
SARIMA	(0, 2, 0)	(0, 1, 0, 12)	8,169,985	8,836,844	23.36%
SARIMAX	(0, 1, 0)	(0, 1, 0, 12)	5,016,261	5,164,030	14.40%

Prepared by Sahil Nagpal • LinkedIn: https://www.linkedin.com/in/sahil-nagpal-318999172/

Key Business Insights

- SARIMAX delivered the most reliable time■series forecasts with the lowest percentage error (MAPE 14.4%).
- Linear Regression explained 57% of historical variation (R² 0.572) but does not model seasonality explicitly.
- SARIMA (seasonality only) underperformed on the test window, indicating the value of exogenous structure in SARIMAX.
- Forecasts indicate seasonal peaks (Oct–Nov), useful for capacity planning and roster allocation.
- Recommendation: retrain monthly, extend history when available, and consider adding holiday/weather features to improve accuracy.

Contact: Sahil Nagpal • LinkedIn: https://www.linkedin.com/in/sahil-nagpal-318999172/