Intelligent Inventory

A TIME SERIES ANALYSIS AND LSTM OPTIMIZATION

Edwin Quijano

University of West Florida

Capstone Project In Data Science IDC-6940 | Professor Dr. Shusen Pu

Project Background

- Inventory storage during 2020 to 2022.
- Normalization of inventories in 2023 and 2024.
- U.S political policies affecting supply chain in 2025.
- Shipments are re-routed to South America to reduce costs.
- The challenge of limited warehouse space.

Why Forecasting Is Important

- Helps Plan Warehouse Capacity
- Improves Labor & Staffing Plans
- Reduces Congestion and Cycle Time
- Enables Better Inventory Strategy
- Supports Financial and Operational Planning
- Improves Vendor and Inbound Scheduling
- Creates Early Warning Signals

Data & Quality

- 10,000 rows
- 4 product categories
- Product dimensions
- Stock quantity

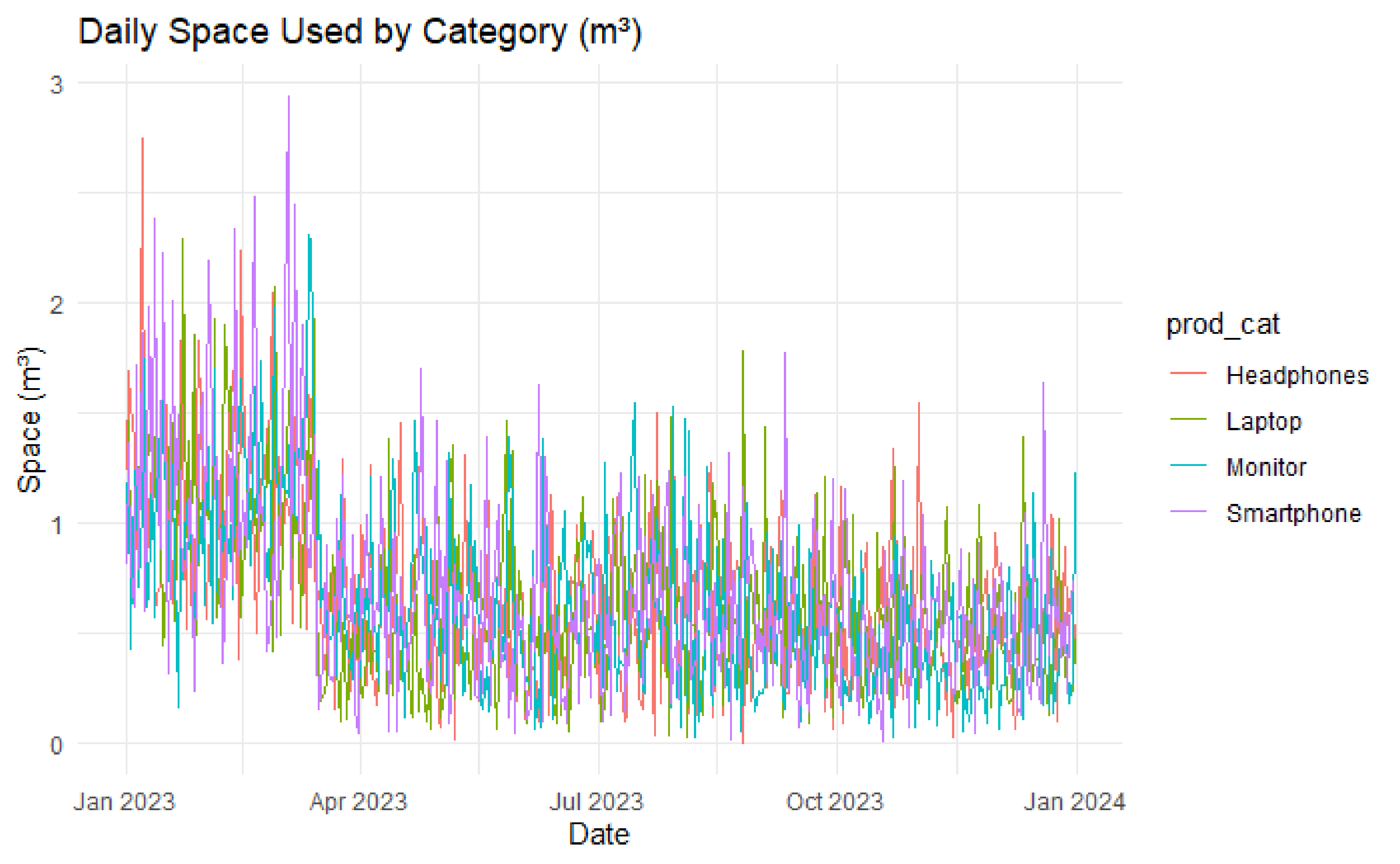
Key variables

prod_id <chr>	prod_cat <chr>	prod_sub <chr>	prod_desc <chr>	prod_price <dbl>	warranty_period <dbl>	prod_dimensions <chr>
93TGNAY7	Laptop	Home Appliances	Product_XU5QX	253.17	2	16x15x15 cm
XBHKYPQB	Monitor	Clothing	Product_8SBDO	403.33	1	7x13x5 cm
27R9M103	Smartphone	Electronics	Product_Z5CGR	18.87	3	11x16x7 cm
JDOVOMY2	Headphones	Home Appliances	Product_7IBNL	81.29	3	14x18x14 cm
0KHFMXFN	Monitor	Home Appliances	Product_GFUOS	21.35	2	11x6x11 cm

stock_qty <dbl>	stock_xfer_date <date>	manuf_date <chr>	exp_date <chr>	sku <chr>	prod_tag <chr>	color_var <chr>	rating <dbl>
3	2023-03-05	1/1/2023	1/1/2026	8NMFZ4	VNU,NZ6	Green/Large	2
40	2023-02-12	1/1/2023	1/1/2026	65MQC3	RPP,M40	Green/Large	1
98	2023-01-01	1/1/2023	1/1/2024	UKNO5L	KMQ,39Z	Green/Large	5
32	2023-01-16	1/1/2023	1/1/2026	L72TO4	8JR,Z6A	Green/Large	2
48	2023-01-20	1/1/2023	1/1/2026	4C11JZ	M0V,5ES	Blue/Medium	1

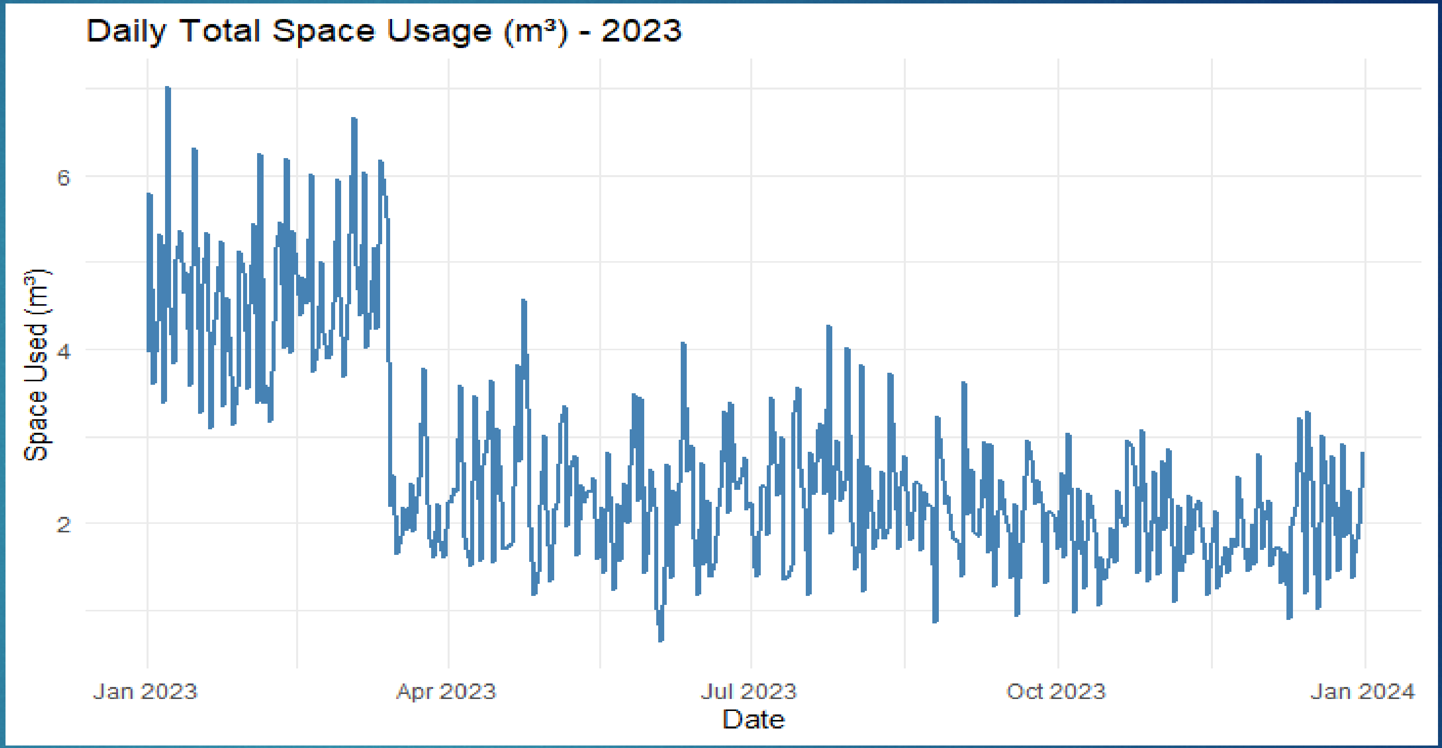
MODEL & METHODOLOGY

- Data Preparation for Modeling
- Stationarity Testing (ADF Test)
- SARIMA Model Structure
- SARIMA Residual Diagnostics
- VAR Model



RESULTS & INTERPRETATION

- Total space usage

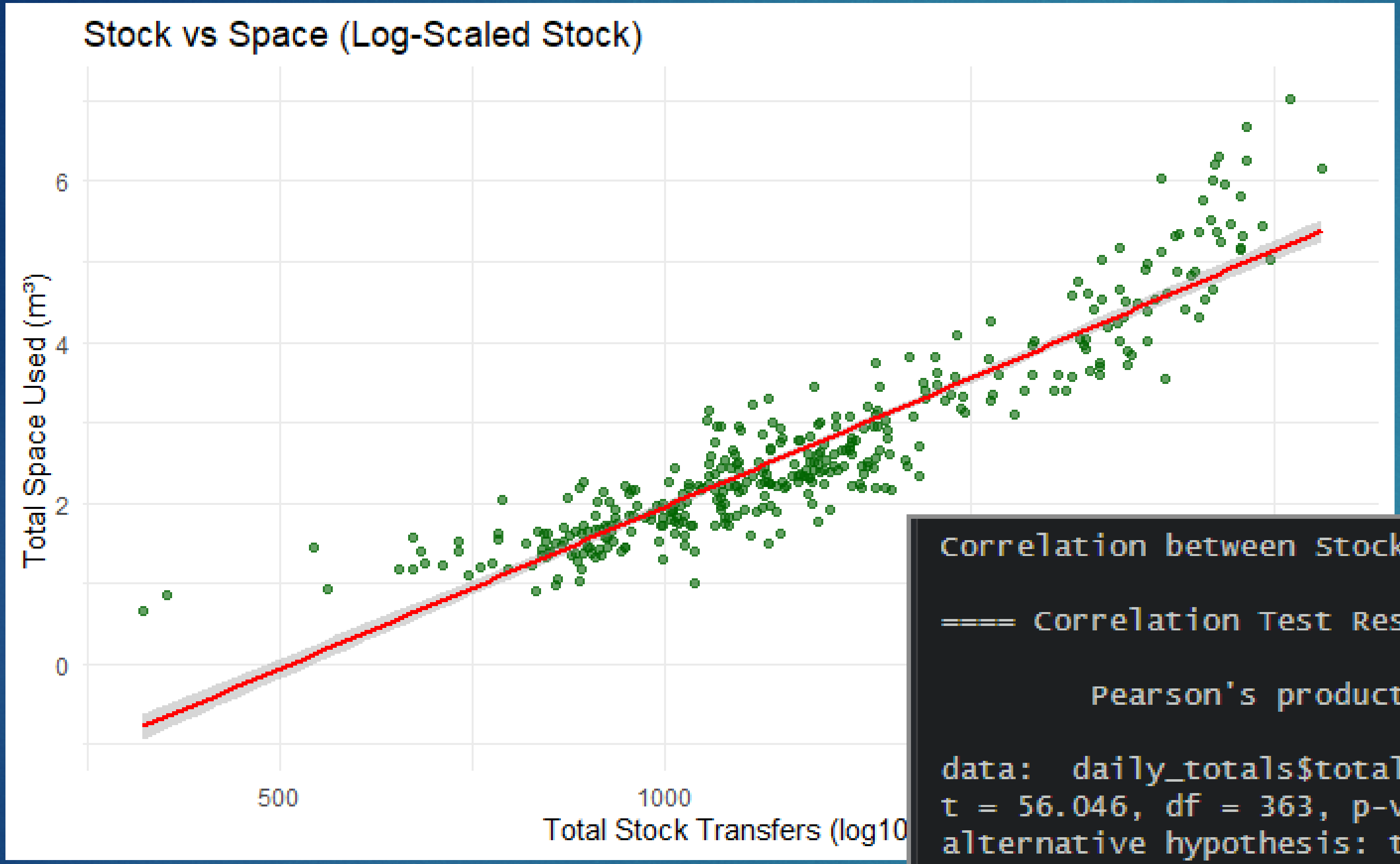


A tibble: 1 x 8

mean_space	sd_space	min_space	max_space	median_space	q1_space	q3_space	iqr_space
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
2.680113	1.224966	0.651418	7.016069	2.339436	1.797563	3.264237	1.466674

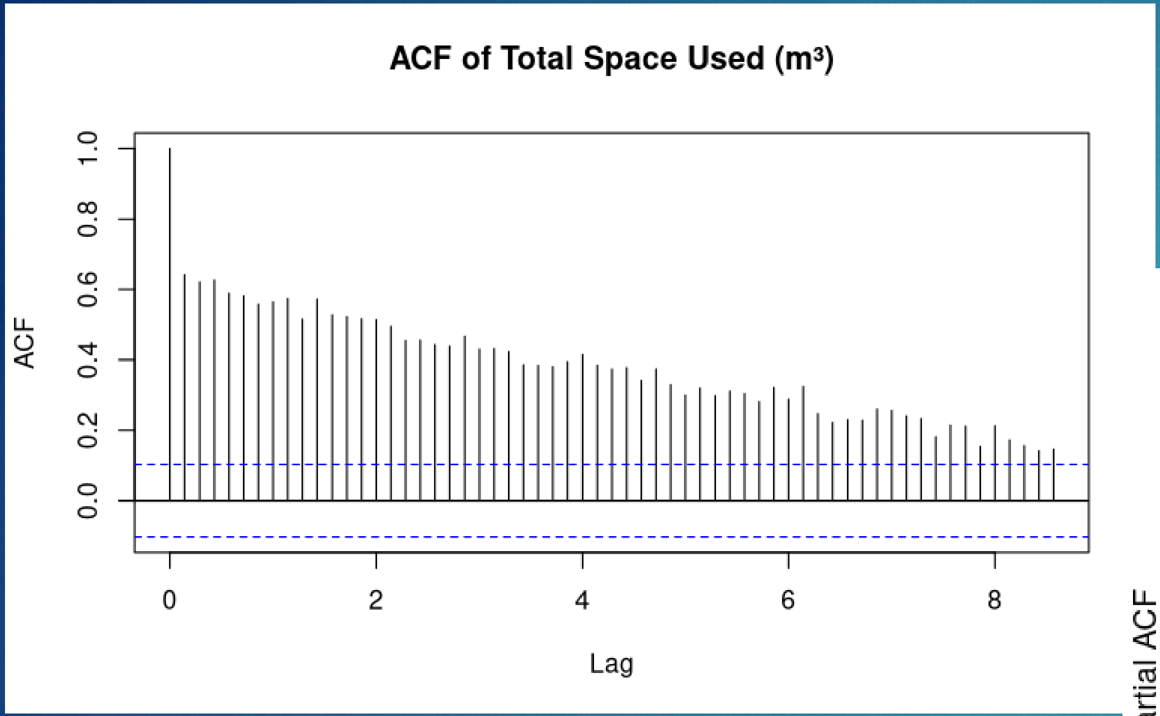
1 row

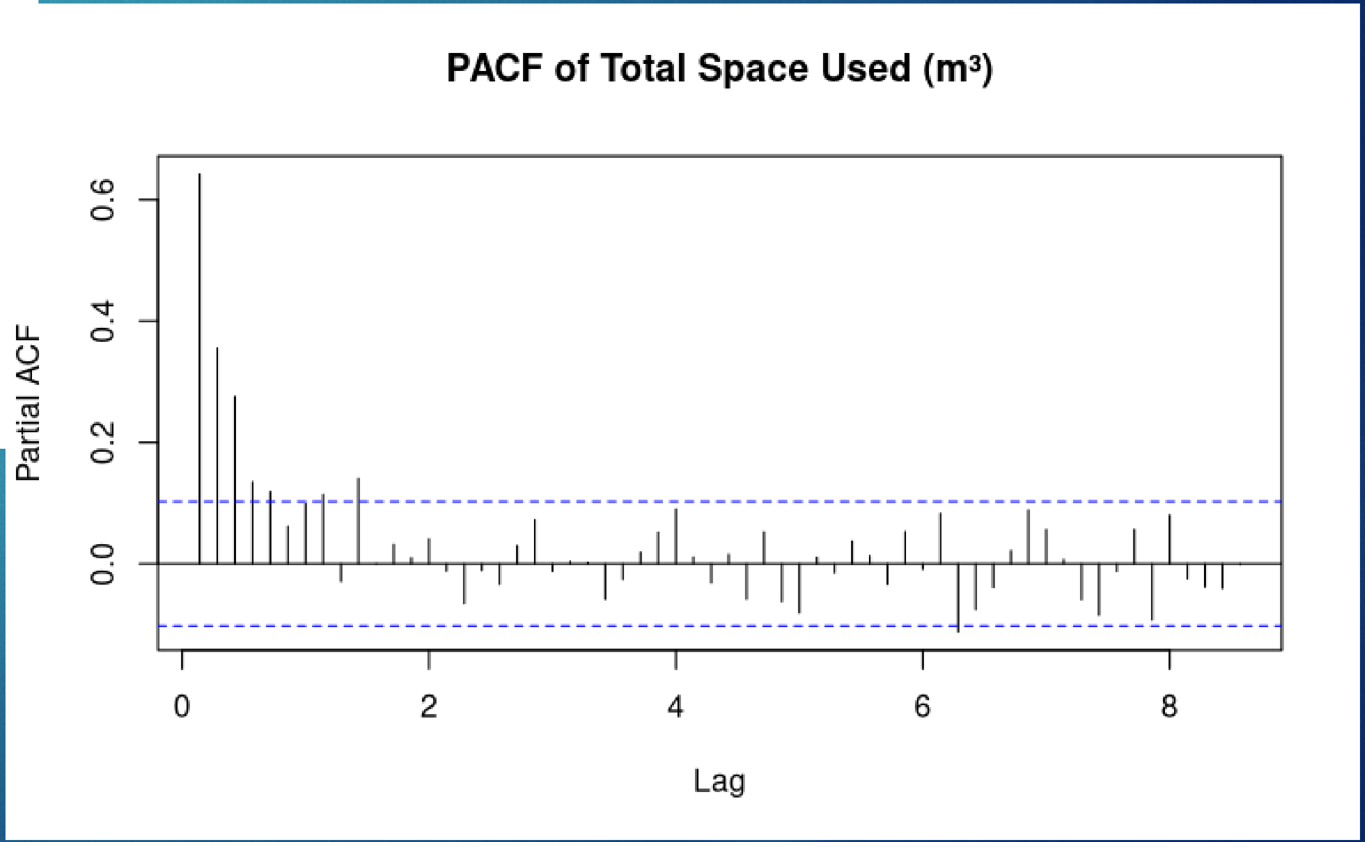
Stock vs. Space Correlation




```
Correlation between Stock Transfers and Space Used (m³): 0.9468

==== Correlation Test Results ====

	Pearson's product-moment correlation

data:  daily_totals$total_stock_qty and daily_totals$total_space_m3
t = 56.046, df = 363, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9350109 0.9564809
sample estimates:
      cor
0.9467892
```

Total Space – Non-Stationarity

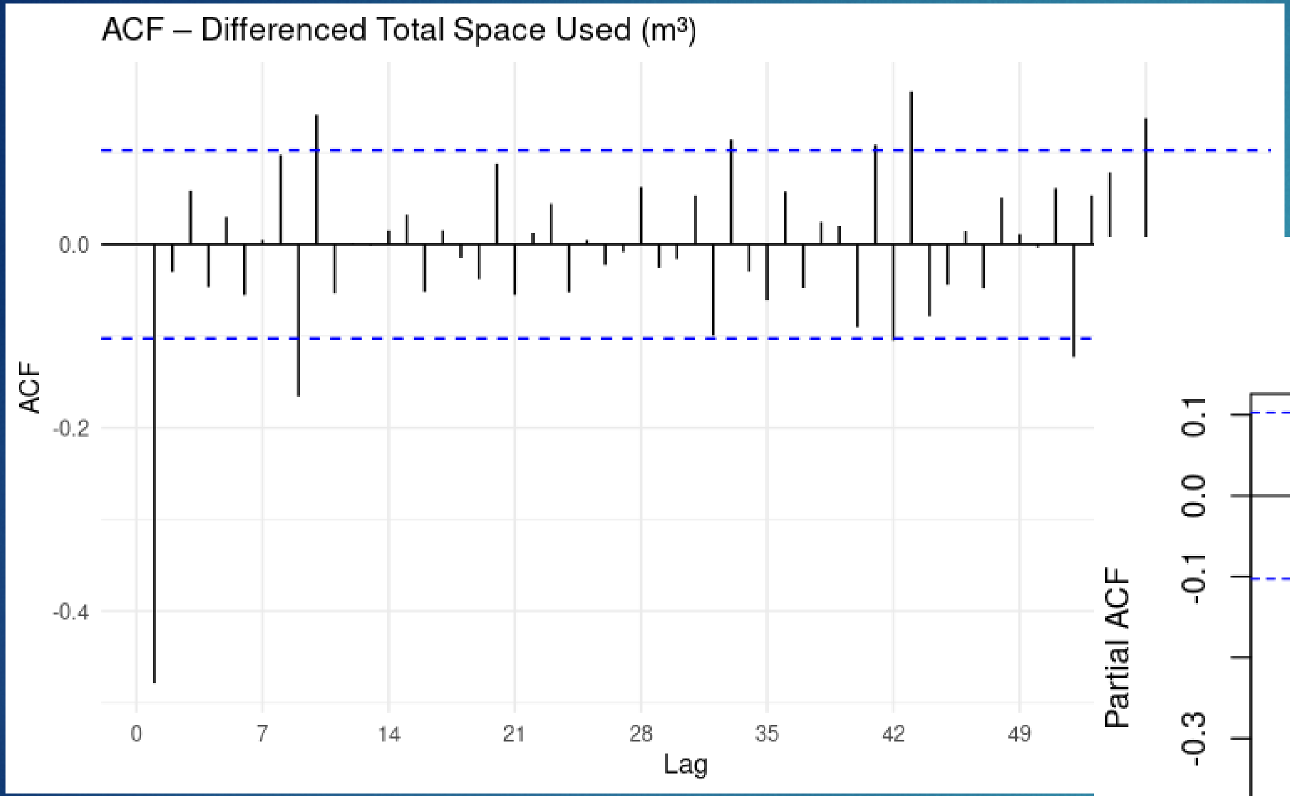


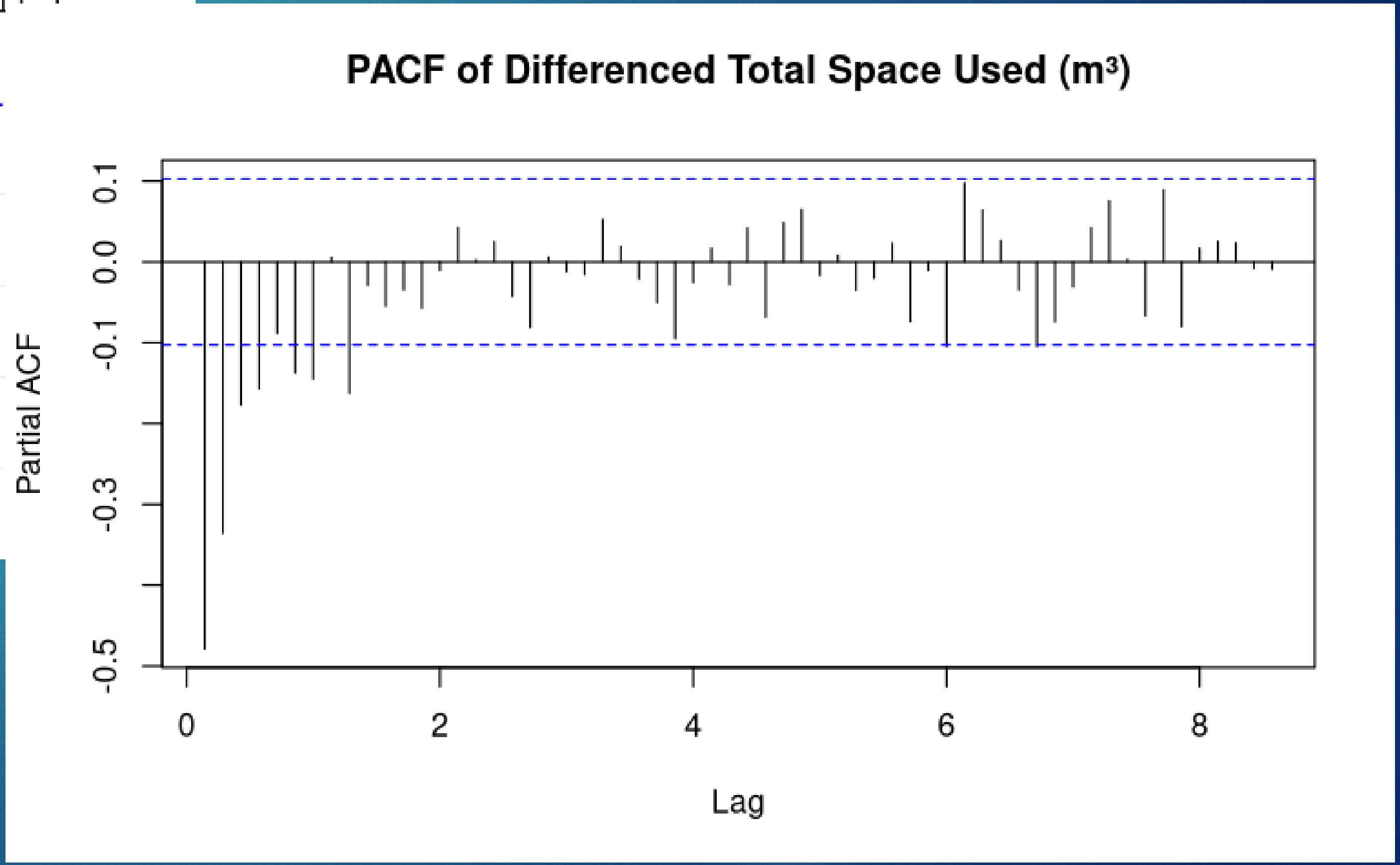


Augmented Dickey–Fuller (ADF) test

```
        Augmented Dickey-Fuller Test

data:  ts_total_space
Dickey-Fuller = -3.0578, Lag order = 7, p-value = 0.1305
alternative hypothesis: stationary


        Augmented Dickey-Fuller Test

data:  diff(ts_total_space)
Dickey-Fuller = -9.9257, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Total Space with difference





SARIMA Model

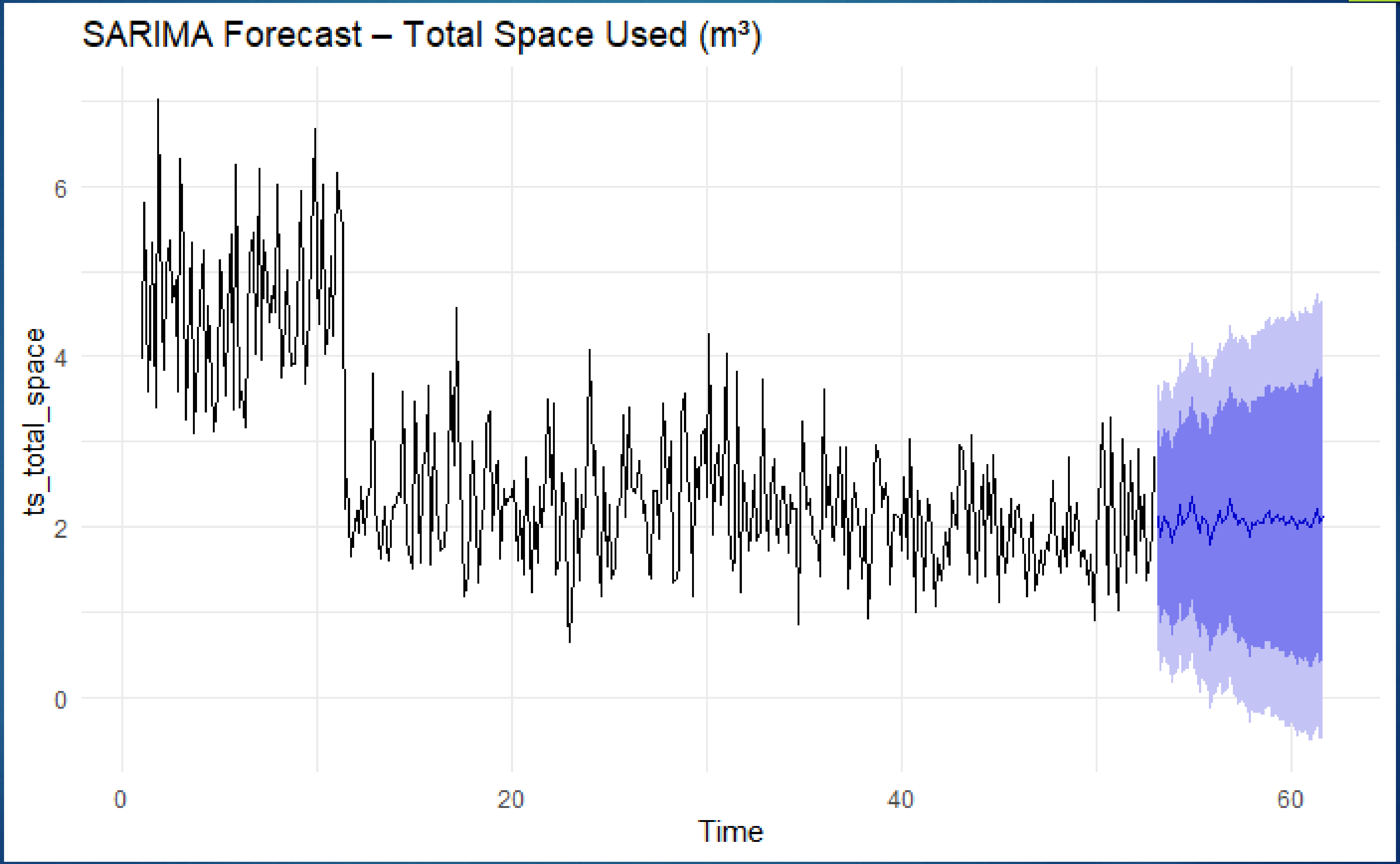
SARIMA (*p,d,q*) (*P,D,Q*)

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D y_t = \Theta(B^s)\theta(B)\varepsilon_t$$

Where:
B= backshift operator
$\phi(B)$= non-seasonal AR
$\theta(B)$= non-seasonal MA
$\Phi(B^s)$= seasonal AR
$\Theta(B^s)$= seasonal MA
s = 7days (weekly seasonality)

p=0, *d*=1, *q*=2   *P*=2, *D*=0, *Q*=2   *s*=7

SARIMA Model



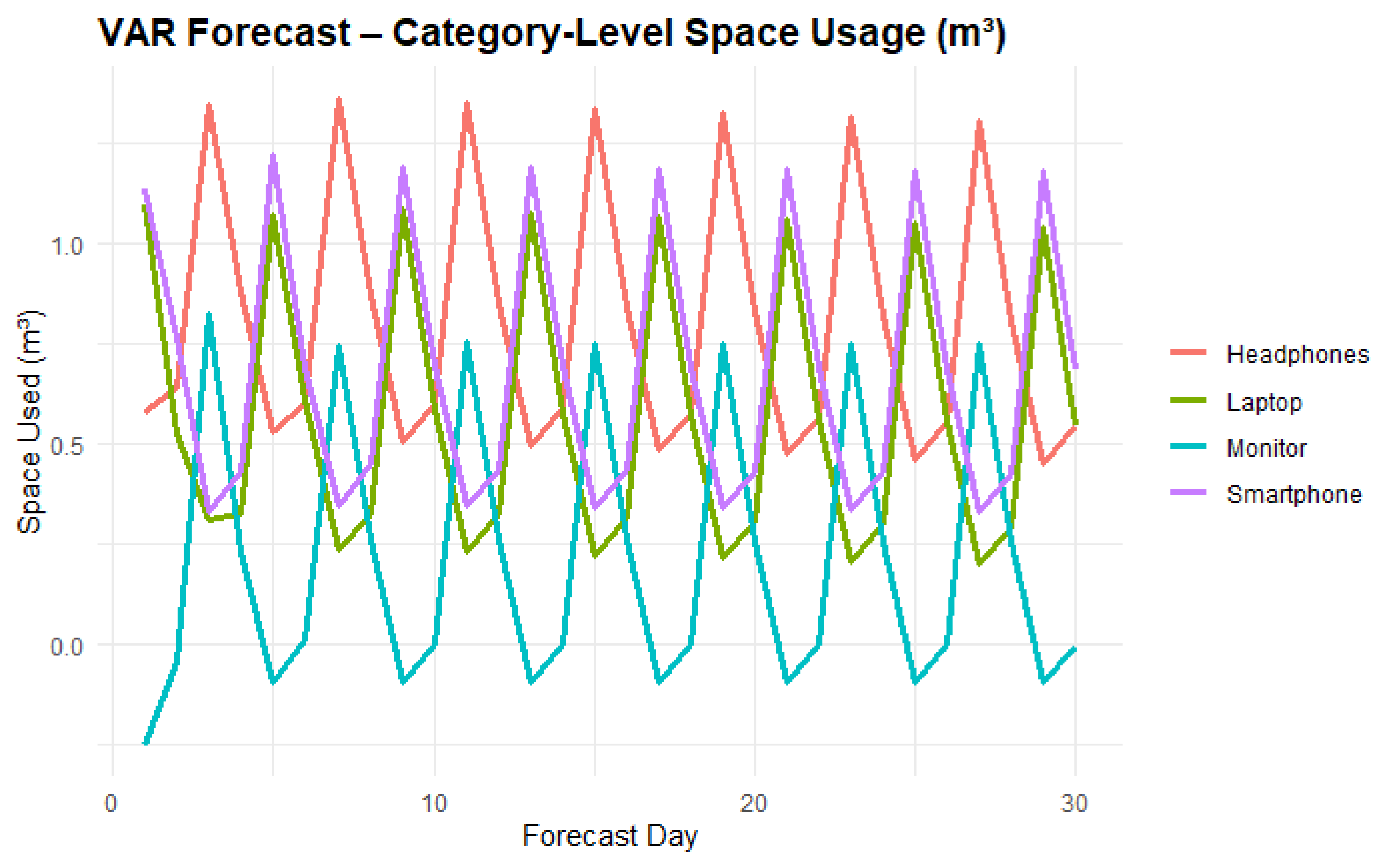
VAR Model

A VAR(2) means today's changes depend on the last two days.
The full model:

$$\Delta y_t = c + A_1 \Delta y_{t-1} + A_2 \Delta y_{t-2} + \varepsilon_t$$

Where:
Δy_t =vector of differenced category space at time t
c= vector of intercept constants
A_1 =coefficient matrix for lag 1
A_2 =coefficient matrix for lag 2
ε_t =vector of multivariate white-noise errors

VAR Model



VAR Forecast – Category-Level Space Usage (m³)
Space Used (m³)
1.0
0.5
0.0
0
10
20
30
Forecast Day
Headphones
Laptop
Monitor
Smartphone

LSTM vs ARIMA

Model	Performance	Results
ARIMA (SARIMA)	Excellent at capturing weekly seasonality and short-term dynamics. RMSE ≈ 0.62 m^3/day	Best-performing model
LSTM	Learns nonlinear patterns, robust to noise RMSE typically 0.75–0.95 m^3/day	Underperforms compared to ARIMA

OPERATIONAL & BUSINESS

- The warehouse should expect **regular weekly fluctuations** in space usage. **Smartphones** are the single most important category for peak capacity planning.
- Forecasts help identify **high-demand periods** early, allowing better resource allocation.
- Category-level forecasts allow teams to plan **which zones** will need the most space in upcoming weeks.

Future Improvements

- Expand the Dataset (More Historical Data)
- Add Key External Predictors (ARIMAX or ML Models)
- Integrate Machine Learning Models

Thank You!