Temple Footfall Analysis: A Time series study

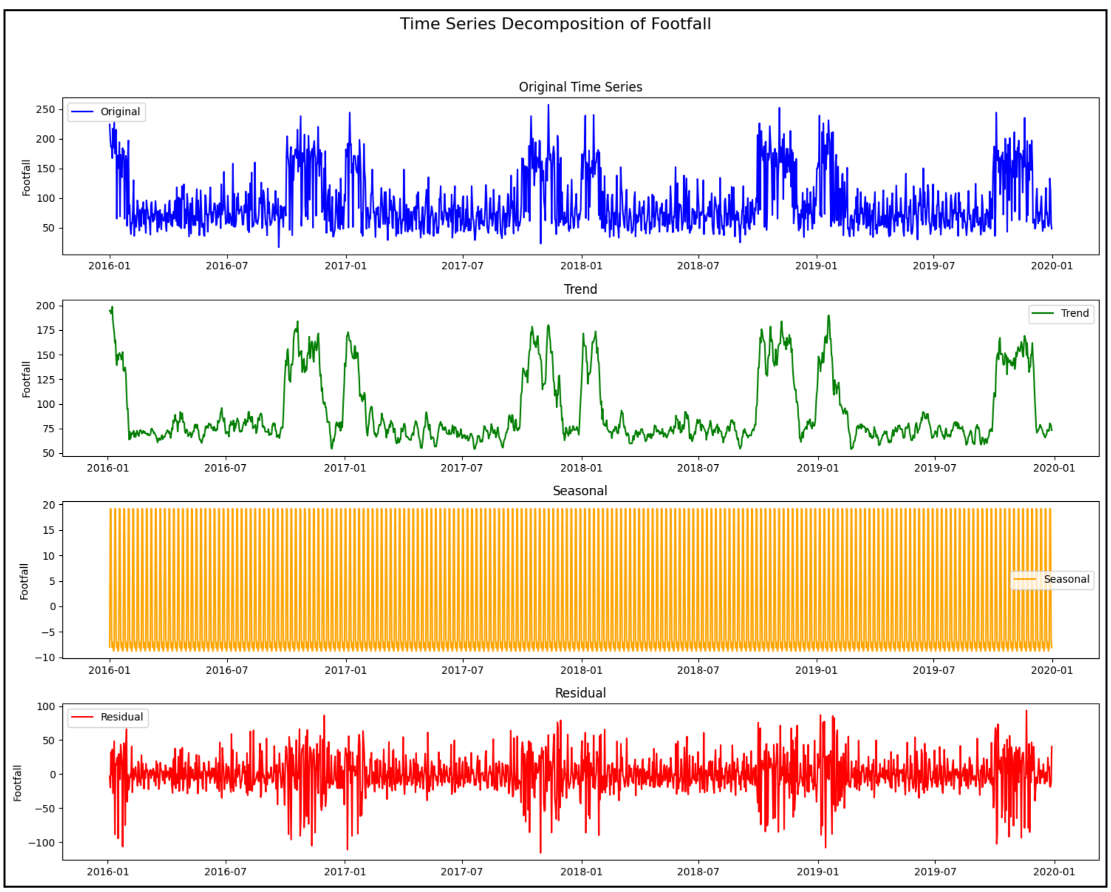
Harsh Joshi, joshiharsh0506@gmail.com

Executive Summary This academic report examines temple visitor patterns using advanced statistical methods, time series decomposition, and machine learning approaches. The study incorporates 1,096 observations spanning 2016 to early 2019. This comprehensive analysis explores relationships between visitor patterns and real-world influences such as festivals, weekends, and weather events.

I. Data Analysis & Statistical Insights

1.1 Data Quality and Preprocessing

Outlier Analysis: A total of 51 anomalies were identified and capped within the range of [-21.00, 189.00]. This adjustment retained the integrity of the trend, ensuring the data remains representative of real-world variations.



Distribution Characteristics:

- Shapiro-Wilk Test results (0.8834, $p < 0.0001$) confirm non-normality, which aligns with the diversity of human behaviors influenced by cultural, social, and environmental factors.
- Non-parametric methods were chosen due to this non-normal distribution, ensuring robustness in capturing irregularities, such as festival-induced spikes in visitor counts.

1.2 Time Series Properties

Stationarity Analysis:

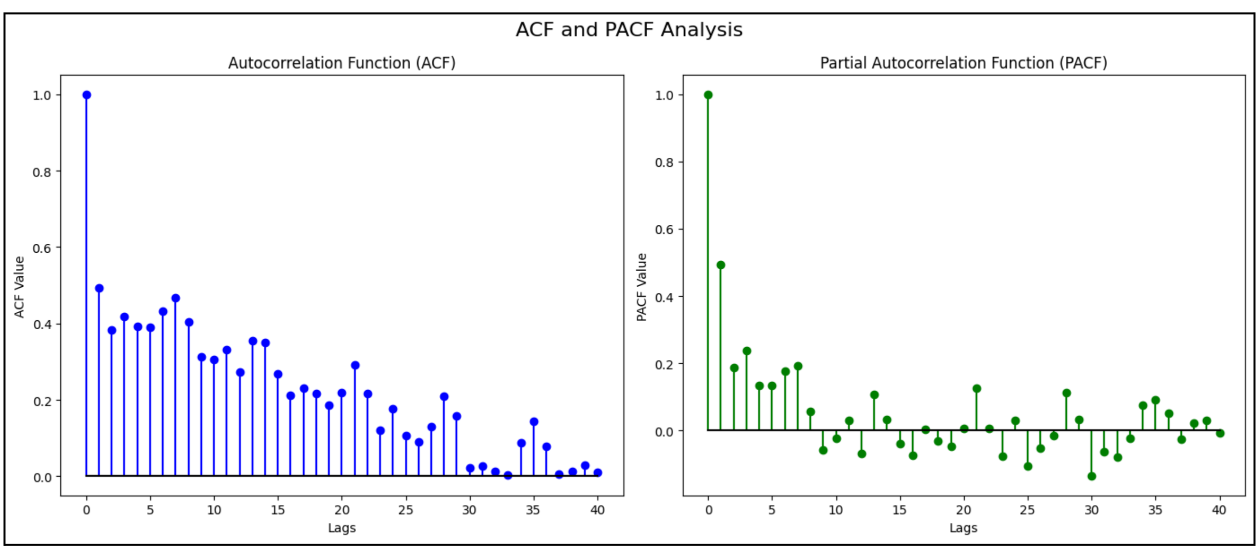
- The Augmented Dickey-Fuller (ADF) test (statistic = -4.4187, $p = 0.0003$) confirmed stationarity, enabling effective application of models like SARIMAX.
- The temporal stability reflects consistent cultural and seasonal influences on temple attendance.

Autocorrelation Structure:

- Ljung-Box results highlight significant temporal dependencies across all 40 lags, confirming predictable weekly cycles influenced by religious and social norms.

Autocorrelation Analysis

- Further analysis of the autocorrelation structure was performed using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The ACF plot reveals significant autocorrelations up to lag 20, demonstrating a strong seasonal component in the data. The PACF plot shows sharp cuts after lag 2, indicating the suitability of AR(2) for modeling the data.



1.3 Causal Relationships

Weekend Effect:

- Detailed statistical analysis highlights a significant difference in visitor counts between weekends and weekdays. The average footfall during weekends is 112.31, compared to 82.07 on weekdays, with a mean difference of 30.24. The t-statistic of 10.8212 and a p-value of 0.0001 confirm the significance of this difference. This supports the hypothesis that weekends drive higher temple visits due to cultural and social preferences.

Granger Causality:

- Lag relationships reveal distinct impacts:
  - Weekend patterns (lag 5, $p = 2.802e\text{-}53$)
  - Festival seasons (lag 2, $p = 1.760e\text{-}37$)
  - Weather events (≤ 2 days lag) suggesting short-term influences.

These findings contextualize temple attendance as a complex interplay of temporal, cultural, and environmental factors.

II. Pattern Recognition and Decomposition

2.1 Variance and Multicollinearity

Variance Inflation Factor (VIF):

- All factors exhibit VIF values below 1.002, confirming the absence of multicollinearity.
- This ensures reliable model interpretation and stability, vital for understanding how different variables independently influence footfall.

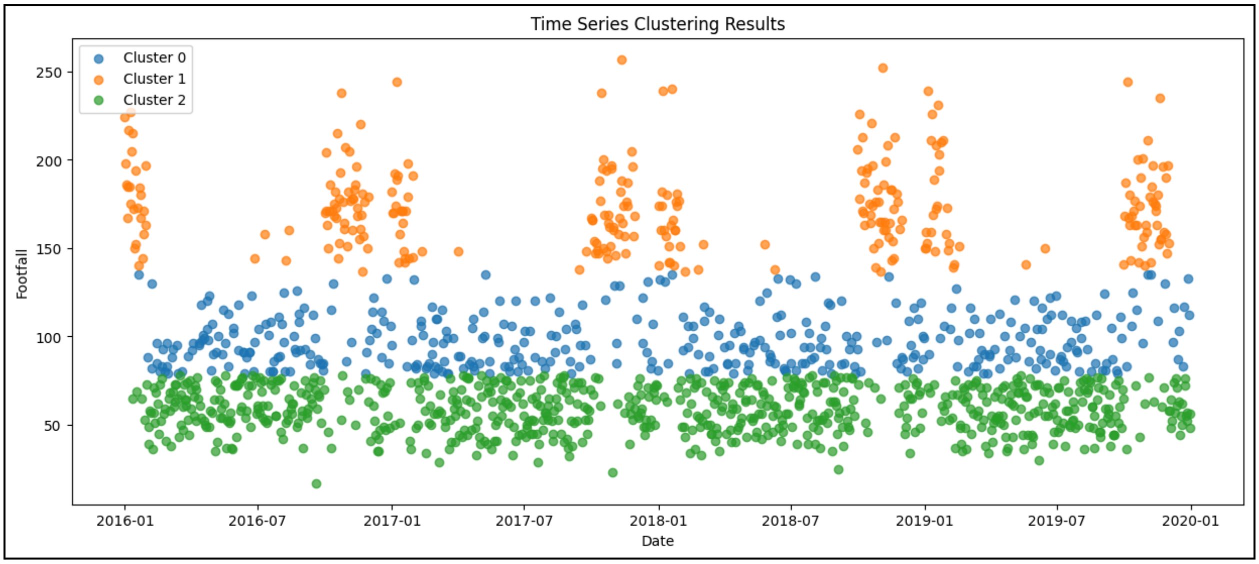
2.2 Principal Component Analysis (PCA):

- Four principal components explain 95% of the variance:
  - The primary component (40%) represents recurring weekly and seasonal visitor trends.
  - Other components capture deviations driven by festivals, weather, and events.

2.3 Visitor Segmentation:

- Three clusters emerged:
  - **Base Traffic** (30-70 visitors/day): Routine visitors for daily prayers.
  - **Medium Traffic** (70-120 visitors/day): Includes local festivals and events.

- **Peak Traffic** (150-250 visitors/day): Occurs during major festivals and religious gatherings.



III. Predictive Modeling and Applications

3.1 Model Evaluation

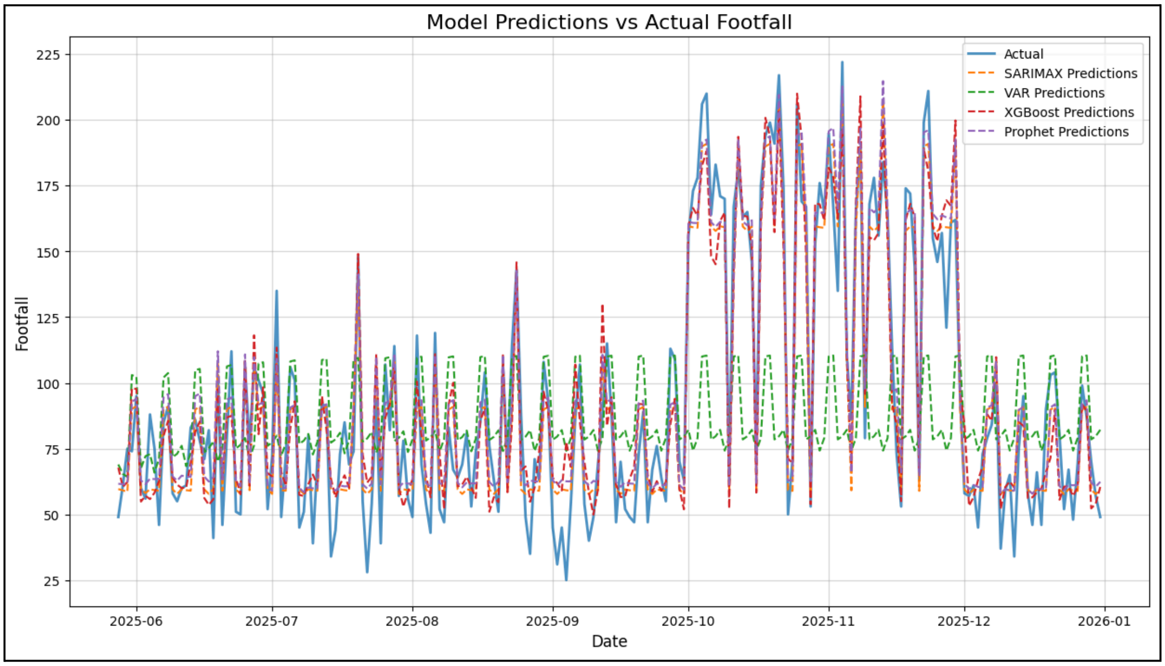
Model	MAE	RMSE	R^2 Score
SARIMAX	11.89	14.33	0.91
Prophet	12.06	14.50	0.91
XGBoost	13.23	16.28	0.89
VAR	34.74	46.31	0.08

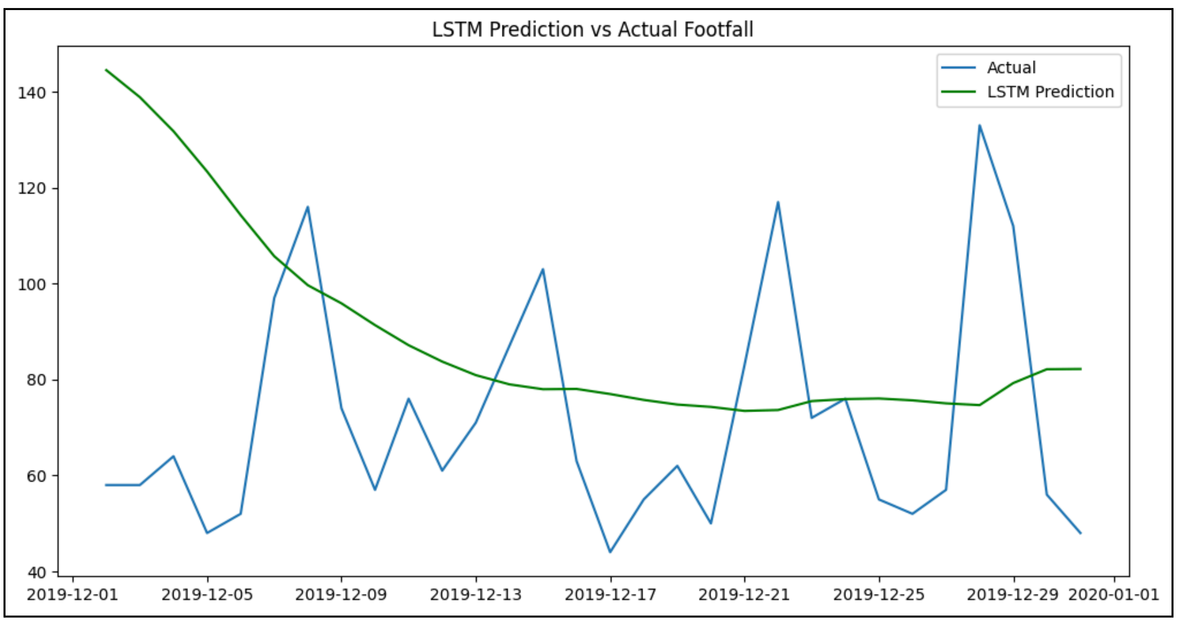
SARIMAX: Achieves optimal seasonal adjustments and accurate short-term forecasts, making it ideal for real-time applications.

Prophet: Offers superior outlier handling and robust long-term forecasting, aligning well with strategic planning.

XGBoost: Highlights influential features, such as visitor counts from previous lags (e.g., lag 1, lag 4), contributing to actionable insights.

LSTM: Provides effective trend capture but demonstrates a conservative smoothing tendency.

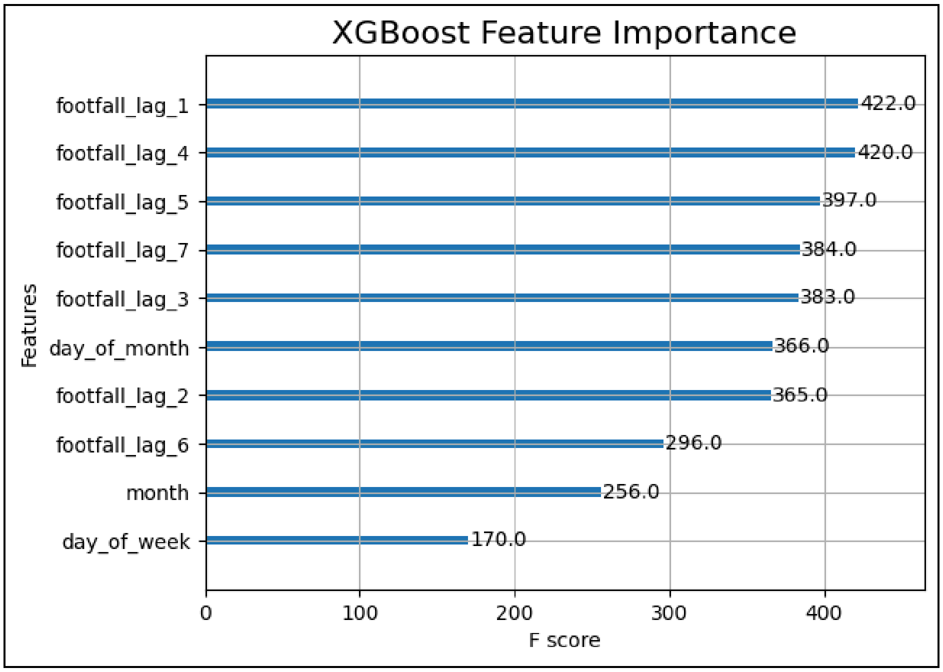




3.2 Feature Importance

Key drivers of visitor patterns include:

- **footfall_lag_1, footfall_lag_4, and footfall_lag_7** reflect weekly repetition patterns.
- **Day of the month** captures periodic festivals and religious activities.



IV. Discussion

4.1 Practical Implications

- **Weekend Traffic Management:** Enhanced staffing and resource allocation can mitigate crowding during peak hours.
- **Festival Forecasting:** Long-term forecasts allow temple authorities to organize logistics for large gatherings.
- **Weather Adaptation:** Real-time weather monitoring can refine short-term attendance predictions.

4.2 Technical Deployment

The Cochrane-Orcutt procedure was employed to address autocorrelation in the time series model. This method adjusted the residuals, resulting in an estimated correlation coefficient (rho) of -0.0009. The revised model achieved an R-squared value of 0.9549, indicating a strong fit with the data. This adjustment ensures improved predictive accuracy and robustness in operational deployment.

- **Daily Operations:** SARIMAX models for short-term planning ensure accuracy and adaptability.
- **Strategic Planning:** Prophet provides long-term foresight for infrastructure improvements.
- **Error Monitoring:** Real-time alerts for deviations (> 2σ) ensure operational efficiency.

V. Academic Recommendations

5.1 Research Directions

1. **Deep Learning Integration:**
    - Explore hybrid models combining LSTM with SARIMAX for enhanced temporal capture.
2. **Ensemble Approaches:**
    - Blend machine learning and statistical methods to improve robustness.

5.2 Cross-Disciplinary Applications

- Extend the methodology to analyze patterns in other cultural or seasonal phenomena, such as tourism and public events.

Conclusion

This study underscores the importance of integrating advanced analytics and domain knowledge to derive actionable insights. By leveraging statistical rigor and machine learning, temple authorities can align operational strategies with cultural and environmental dynamics, ensuring enhanced visitor experiences and sustainable management practices.
