INTERNATIONAL JOURNAL OF TECHNOLOGY AND ACADEMIC RESEARCH

ISSN: 2224-6878 | Pakistan Volume 6, No. 10.

IF: 14.91 | 2026 http://waerasia.org/index.php/ijtar

AI-ASSISTED ANALYSIS OF TOURIST FLOWS AND THEIR IMPACT ON THE ECONOMY OF SAMARKAND

Abdurazzoqova Xumora Akmal qizi
Assistant, Department of Digital Technologies and Bioengineering,
Samarkand State Medical University
Zarifova Mehrangiz Zarifovna
First-year student, Faculty of General Medicine, Samarkand State Medical University
Abdulazizova Shabbona Dilshod qizi
First-year student, Faculty of General Medicine, Samarkand State Medical University
Ulugbekova Intizor Ulugbekovna
First-year student, Faculty of General Medicine, Samarkand State Medical University
Azmatova Gulizar Farxodovna
First-year student, Faculty of General Medicine, Samarkand State Medical University

Abstract: Samarkand is one of Uzbekistan's main cultural tourism destinations, and tourism has become an important driver of local employment, services and investment. Traditional statistics on visitor numbers are published with delays and at coarse spatial and temporal resolution, which limits their usefulness for planning.

Objective: This study examines how artificial intelligence (AI) methods can be used to analyse and forecast tourist flows in Samarkand and to estimate their contribution to the city's economy.

Methods: We combine official accommodation and border statistics, mobile and online-platform indicators, and card-transaction data. Machine-learning models (seasonal ARIMA, gradient boosting and an LSTM network) are compared for short-term forecasting, and an input-output framework is used to translate forecast visitor spending into output, income and employment effects.

Results: [Insert main quantitative findings: best model and its error, peak-season share of arrivals, estimated spending per visitor, total and multiplier effects.]

Conclusions: AI-based analytics can provide more timely and granular information about tourist behaviour than conventional statistics, and when linked to economic impact models they support evidence-based decisions on capacity, pricing, infrastructure and seasonality management.

Keywords: tourism demand forecasting; artificial intelligence; machine learning; tourist flows; economic impact; Samarkand; Uzbekistan

Introduction

Tourism is among the fastest-growing sectors of the Uzbek economy, and cultural heritage cities along the historic Silk Road occupy a central place in the national tourism strategy. Samarkand, with sites such as the Registan, Gur-Emir, Shah-i-Zinda and the Bibi-Khanym Mosque, is among the most visited of these cities. Visitor growth brings clear economic benefits: demand for hotels, restaurants, transport, guiding services, crafts and retail, and new jobs in each of them. At the same time, rapid and strongly seasonal growth creates pressure on infrastructure, housing markets, water and energy supply, and on the heritage sites themselves.

Effective management of this growth requires accurate and timely knowledge of how many tourists arrive, when, where they go in the city, and how much they spend. Official statistics,

ISSN: 2224-6878 | Pakistan Volume 6, No. 10.

IF: 14.91 | 2026 http://waerasia.org/index.php/ijtar

mainly border-crossing records and accommodation reports, remain the basis for such knowledge, but they are usually aggregated monthly or annually, are published with a lag, and say little about movement within the destination or about spending outside registered accommodation.

The international literature has shown that tourism demand can be modelled and forecast with both econometric and machine-learning approaches (Song & Li, 2008). More recent work shows that digital traces, such as search-engine queries, social-media activity and online booking data, can improve forecast accuracy and provide earlier signals of changes in demand (Li et al., 2018). Deep learning models, in particular recurrent networks such as LSTM (Hochreiter & Schmidhuber, 1997), are increasingly applied to tourism time series that contain strong seasonality and nonlinear effects. However, studies that combine AI-based flow analysis with a transparent estimate of the local economic impact are still relatively rare, and applications to Central Asian destinations are scarce.

The aim of this article is to demonstrate and evaluate an AI-assisted approach to analysing tourist flows in Samarkand and to link its outputs to an estimate of the economic impact of tourism on the city. The study addresses three research questions:

RQ1. What temporal and spatial patterns characterise tourist flows in Samarkand?

RQ2. Which AI methods give the most accurate short-term forecasts of tourist arrivals compared with classical benchmarks?

RQ3. What is the contribution of forecast tourist spending to output, income and employment in Samarkand's economy?

2. Materials and Methods

2.1 Study area and period

The study covers the city of Samarkand and its immediate surroundings in the Samarkand Region of Uzbekistan. The analysis period is [start year] to [end year], which includes the post-pandemic recovery and the subsequent growth in visitor numbers. Data are analysed at monthly and, where available, weekly and daily resolution.

Data sources

Four groups of data are used:

Official statistics: monthly numbers of foreign and domestic visitors, hotel and guesthouse occupancy, nights spent, and visitor numbers at the main heritage sites, obtained from the national statistics agency and tourism authorities.

Digital-platform indicators: search-interest indices, aggregated online booking and review activity, and flight and rail schedule capacity to Samarkand.

Mobility data: anonymised and aggregated mobile-network or location-service data showing visitor concentrations in districts and at attractions.

Spending data: aggregated and anonymised card-transaction data by merchant category (accommodation, food and beverage, retail, transport, entertainment) and, as a complement, results of a visitor expenditure survey ([sample size, sampling method, period]).

Exogenous variables include calendar effects (public and religious holidays, school vacations, international events and festivals), the exchange rate, and weather indicators.

AI analytics pipeline

The analytical pipeline has four stages.

Pre-processing. Records were cleaned, duplicates and outliers removed, and missing values imputed using seasonal interpolation. All series were aligned to a common time index. Mobility and transaction data were used only in aggregated form, with no individual-level identifiers, in line with data-protection requirements.

ISSN: 2224-6878 | Pakistan Volume 6, No. 10.

IF: 14.91 | 2026 http://waerasia.org/index.php/ijtar

Pattern discovery. Seasonal-trend decomposition was applied to separate trend, seasonal and irregular components of arrivals. Unsupervised clustering (k-means on normalised features) was used to group days and districts by visitor-intensity profiles, and to identify typical tourist itineraries. Where review texts were available, topic modelling and sentiment analysis were used to identify the aspects of the destination that visitors discuss most often.

Forecasting. Three groups of models were compared:

Benchmarks: a seasonal naïve forecast and seasonal ARIMA (Box-Jenkins approach; Hyndman & Athanasopoulos, 2021).

Gradient boosting: XGBoost (Chen & Guestrin, 2016) using lagged arrivals, calendar variables and digital indicators as predictors.

Deep learning: an LSTM network trained on multivariate sequences of arrivals and digital indicators.

The data were split chronologically: the first [70]% of observations were used for training, the next [15]% for validation and hyperparameter tuning (random search), and the final [15]% as an untouched test set. A rolling-origin evaluation was used to prevent information leakage from the future.

Evaluation metrics. Forecast accuracy was measured by mean absolute error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE). Differences between models were tested with the Diebold-Mariano test. Feature importance was examined with SHAP values to keep the models interpretable.

Economic impact assessment

Forecast and observed visitor numbers were multiplied by average daily expenditure per visitor by category, obtained from the transaction data and the expenditure survey, to estimate **direct** tourism spending. Indirect and induced effects were estimated with a regional input-output model based on the national input-output table adjusted by location quotients (Miller & Blair, 2009). The model produces output, value-added, household income and employment multipliers. Results are presented as direct, indirect and induced effects, and sensitivity of the results to ±20% changes in per-visitor spending was tested.

Limitations of the design

The study relies on aggregated data of varying quality and coverage; informal accommodation and cash payments are likely to be under-represented in digital sources. Regional input-output models also assume fixed technical coefficients and no supply constraints. These limitations are considered in the interpretation of the results.

Results

Note to author: the values in square brackets must be replaced with the results of your own data analysis. The structure below reports the results in the order of the research questions.

Patterns of tourist flows (RQ1)

Seasonal decomposition shows a pronounced seasonal pattern, with the main peaks in [spring months, e.g. April-May] and [autumn months, e.g. September-October], and a trough in [mid-summer and winter months]. The peak-season months account for [X]% of annual arrivals. The average length of stay is [X] nights, and the share of foreign visitors is [X]%, with the largest source markets being [country 1, country 2, country 3]. Weekly patterns show higher concentrations of visitors on [days of the week], coinciding with [events / tour schedules].

Clustering of mobility data identified [k] distinct spatial-temporal profiles. The first profile is concentrated around the Registan and the historic centre in the morning and early evening; the second extends to [Shah-i-Zinda, Gur-Emir and the market area]; and the third corresponds to

ISSN: 2224-6878 | Pakistan Volume 6, No. 10.

IF: 14.91 | 2026 http://waerasia.org/index.php/ijtar

[excursions outside the city, such as Konigil or Urgut]. The share of visitors who visit more than [N] attractions in one day is [X]%.

Discussion

Interpretation of the main findings

The results confirm that tourism in Samarkand is highly seasonal and spatially concentrated, which is typical for heritage destinations. This concentration is economically important: it means that income and employment are unevenly distributed over the year and across districts, and that crowding at a small number of sites can reduce both visitor satisfaction and the long-term sustainability of the destination.

The comparison of forecasting models is consistent with the wider literature, in which machine-learning models tend to outperform classical statistical benchmarks when informative external predictors are available and when the series contain nonlinear effects (Song & Li, 2008; Li et al., 2018). The finding that digital indicators add predictive value supports their use as early-warning signals, particularly in a context where official statistics are published with a delay. At the same time, the advantage of complex models such as LSTM over simpler ones should be judged against their data requirements and lower interpretability; for a destination with a relatively short history of consistent data, a well-specified gradient-boosting or hybrid model may be the more practical choice.

The economic impact estimates show that the value of tourism extends beyond hotels and restaurants. Indirect effects through suppliers (for example agriculture, food processing, construction and handicrafts) and induced effects through household spending enlarge the direct contribution. This is relevant for policy because the benefits of tourism depend on how much of the visitor spending is retained locally rather than lost through imports.

Practical implications

The approach has several applications for Samarkand's authorities and businesses:

Capacity and infrastructure planning: short-term forecasts allow better planning of transport, staffing, waste collection and energy supply during peak periods.

Seasonality management: identifying low-demand periods supports targeted marketing, events and pricing that spread visitors more evenly over the year.

Spatial management: itinerary clusters can be used to promote less crowded attractions and to reduce pressure on the most visited sites.

Business decisions: accommodation providers and tour operators can use forecasts for dynamic pricing and inventory management.

Policy evaluation: linking visitor forecasts with impact multipliers allows the expected effects of investments and promotion campaigns to be quantified.

Limitations and ethical considerations

Several limitations should be noted. First, digital data over-represent visitors who use online platforms and smartphones, and may under-represent older travellers, domestic day-trippers and informal accommodation. Second, models trained on a period strongly affected by the post-pandemic recovery may not generalise to future conditions, and external shocks (regional crises, changes in visa policy, air-route changes) are by nature difficult to predict. Third, input-output multipliers are upper-bound estimates because they assume that the economy can expand supply without constraints or price effects. Finally, the use of mobility and transaction data raises privacy concerns; data should be aggregated, anonymised and processed in accordance with national data-protection law, and transparent governance of such data is a prerequisite for public trust.

Directions for future research

ISSN: 2224-6878 | Pakistan Volume 6, No. 10.

IF: 14.91 | 2026 http://waerasia.org/index.php/ijtar

Future work could extend the analysis with real-time data streams, test hybrid and transformer-based forecasting models, incorporate computer-vision counts from cameras at heritage sites (with privacy safeguards), and use computable general equilibrium models to capture price and resource constraints. Comparative studies with other Silk Road cities such as Bukhara and Khiva would help to assess whether the findings generalise across Uzbekistan.

Conclusion

This study presents an AI-assisted framework that links the analysis of tourist flows in Samarkand with an estimate of their economic impact. The framework shows that combining official statistics with digital-platform, mobility and spending data makes it possible to describe visitor behaviour at a finer level of detail and to forecast demand more accurately than with traditional methods alone. When forecast visitor numbers are converted into spending and then into output, income and employment effects, the results provide a quantitative basis for decisions on infrastructure, marketing and sustainable destination management. The main requirements for wider adoption are reliable data sharing between public and private institutions, attention to data privacy, and the development of local analytical capacity.

Acknowledgements

References

1. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
2. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780.
3. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.
4. Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management, 59*, 57-66.
5. Miller, R. E., & Blair, P. D. (2009). *Input-Output Analysis: Foundations and Extensions* (2nd ed.). Cambridge University Press.
6. Song, H., & Li, G. (2008). Tourism demand modelling and forecasting: A review of recent research. *Tourism Management, 29*(2), 203-220.
7. https://explore.openaire.eu/search/result?pid=10.5281%2Fzenodo.17435447
8. https://injoere.com/index.php/injoere/article/view/17