Predicting Tourist Arrivals: A Google Trends-Based Model for Destination Management

Tsvetanka Georgieva-Trifonova [1], Olga Mancheva-Ali [2]

[1] *"St. Cyril and St. Methodius" University of Veliko Tarnovo, Faculty of Mathematics and Informatics, Veliko Tarnovo, Bulgaria*
[2] *"St. Cyril and St. Methodius" University of Veliko Tarnovo, Faculty of Economics, Veliko Tarnovo, Bulgaria*

Abstract **– In this paper, a review of the scientific literature on the importance and knowledge of the tourist destination for the formation of a correct strategy for its development and promotion is made. The possibility of managing the tourist destination by modeling with Google Trends data is represented. A review of studies using the mentioned tool in the field of tourism is conducted. The potential benefits of using Google Trends data on searches for a city as a keyword in combination with archival data are summarized. The possibilities for modeling the visits of tourists using data from Google Trends are investigated, as a result of which a model is built by applying polynomial regression to predict the overnight stays of foreign citizens in the city of Veliko Tarnovo based on archival data from the National Statistical Institute and Google Trends data. Through data analytics with Google Trends, it is concluded that key value is created for the external and internal environment, influencing the destination. The data creates conditions for the development of management strategies for attracting tourists based on primary and secondary data for analysis. The application of the received data sets through data analysis are also important for tourism organizations and institutional bodies in tourism (macro- and meso-level). The resulting analysis affects issues related to the management and marketing of the destination, and on the other hand, it directly affects the segmentation of the market and the distribution of products and services to the end user (micro level).**

DOI: 10.18421/TEM133-22
https://doi.org/10.18421/TEM133-22

Corresponding author: *Tsvetanka Georgieva-Trifonova, "St. Cyril and St. Methodius" University of Veliko Tarnovo, Faculty of Mathematics and Informatics,Bulgaria.*
Email: cv.georgieva@live.uni-vt.bg

Received: 05 February 2024.
Revised: 08 May 2024.
Accepted: 27 May 2024.
Published: 27 August 2024.

Keywords **– Management, tourist destination, consumer demand, regression analysis, Google Trends, tourism.**

1. Introduction

The management of the tourist destination is a key process in the development of tourism as a whole. The creation of models, approaches, and criteria for exploring the tourist potential of destinations leads to making key decisions based on certain facts. Knowledge of the peculiarities of the tourist destination and its proper promotion has several important aspects, some of which have economic, social, health, and cultural significance with a scope depending on the size of the area studied.

The study of a tourist destination is relevant to the planning and management of tourism on a regional, national, and global scale. Taking into account the availability of tourist resources, the state of the adjacent infrastructure and specialized superstructure, the uniqueness of the region, the offering of tourist experiences, and opportunities to achieve complexity in tourist services are a prerequisite for specific consumer behavior.

Consumer behavior, in turn, explores the attitudes, desires and preferences of potential users of tourism services. At the same time, this type of research is complex and multi-layered in terms of the diversity of human interests. Generational characteristics also lead to prerequisites that expand or limit the interests of getting to know and visiting tourist destinations. The study of tourist behavior is a topic of leading importance in determining attitudes and creating strategies for attracting tourists in certain areas. With the advancement of digitalization worldwide, opportunities for studying tourism demand and using tools for analyzing information in real time and through archival data are increasing. Specialized information products add value with their capabilities and deliver much-needed information to tourism professionals for timely and real data for consumer demand analysis. This, in turn, leads to planning the most accurate and appropriate tourist supply to satisfy potential tourists.

The purpose of the paper is to present the peculiarities of the tourist destination and to identify the possibilities of using Google Trends as a tool for analyzing tourist search.

The tasks in the study are:

- a brief theoretical overview of the knowledge for the management of the tourist destination;
- representing the opportunities for research of tourist search with the specialized tool Google Trends;
- analysis of the data from the survey conducted from a selected region (on the example of destination Veliko Tarnovo).

To achieve the purpose of the paper, different research methods are used – primary and secondary. Among the former, the method of mathematical modelling based on polynomial regression and statistical analysis can be counted. with the authors used archival data combined with data from Google Trends. The methodology is complemented by a theoretical overview, analysis and synthesis of information, and use of empirical data.

2. Methodology

The proposed methodology includes two aspects of the research problem. On the one hand, this is knowledge about the specifics of the tourist destination, and on the other, existing research on the use of Google Trends data in the field of tourism.

2.1. Theoretical Overview of the Knowledge about the Specifics of the Tourist Destination

Tourism is a sector of the economy that helps to solve a wide range of socially significant issues that are relevant to economic indicators such as employment, welfare, competitiveness, etc. Its influence often goes beyond the borders of a country and creates a prerequisite for a global perception of the tourist supply, creating expectations for an attractive tourist product no matter its area. In order to create such a product, knowledge of the specifics of the regions is needed, taking into account, first of all, the availability of tourist potential – these are the available tourist resources, adjoining infrastructure, tourism policy, and many others. In this regard, the study of these areas is not only key but also structurally decisive for the specialization of supply and marketing orientation to consumers. If we have to combine all these factors and conditionalities into one, we should talk about managing tourist destinations and finding the means to increase their competitiveness.

The World Tourism Organization (UNWTO) defines a local tourist destination as a physical space where the tourist spends at least one night [1].

Its market competitiveness has an impact at territorial and administrative level, taking into account its image, perception and level of development and management [1].

In the Tourism Act of Bulgaria (amend. SG 17/2020), the tourist destination is defined as a regulated set of economic, social and cultural activities carried out in tourist sites in order to create, realize and consume goods and services forming the attractiveness of the tourist product (package) with sustainable competitiveness and efficiency [2].

According to Davidson and Maitland [3] destinations are centers that stimulate tourist visits, where a predominant part of the tourist product is formed and consumed leading to significant impacts on tourism.

If we have to define the significance of the destination from a tourist point of view, it is a place with the presence of natural and anthropogenic resources, the development and promotion of which determines its characteristics and interest among tourists.

Based on the existing theoretical concepts and combining different knowledge of destination management, Marinov and Kazandjieva [4] propose a model of the destination, within which its essence, content elements and management features are specified. The model is defined as systemic and includes two types of factors – the external and internal environment.

First of all, the factors of the external environment of the destination are put first, which is divided into:

- macro- (the political, economic, social, technological, environmental, and legal factors);
- meso- (market factors such as tourism demand, tourist supply, and competitive environment in tourism);
- micro (suppliers, intermediaries, direct competitors, public attitudes) destination environment.

Secondly, the authors of the systematic approach determine the factors of the internal environment of the tourist destination, which are relevant to the managerial subjects in the destination. These are:

- the resources of the destination (including availability, diversity, staff, investments, needs, motives);
- infrastructure and superstructure that represent the primary attractiveness of the territory (including facilities, stay and service of tourists, amenities for tourist stay);
- public and private sector (including state institutions at national and regional level, municipal authorities, non-profit and profit organizations in tourism);

- tourism industry (a set of enterprises that create and offer services and goods for tourists - hotel, restaurant, transport, intermediary and additional and specific tourist services);
- destination marketing mix, which has four components: total tourism product (experiences and perceived benefits of tourists – attractions, amenities, services, accessibility), cost of the expenditure incurred by the tourist for travel to the destination, distribution channels of the destination (the ways in which tourist services reach consumers – through on-site intermediaries, online distribution systems, reservation, information computer systems) and image of the destination (the way which is perceived destination on the basis of ideas, experience, imagination, information).

The presented model allows for extraction and analysis of valuable information based on the concentration of tourist resources in a particular region, available facilities, and developed tourism industry [4]. These indicators create a prerequisite for tourist demand, but on the other hand for competition between different economic actors in tourism. The separate consideration of a tourist destination makes it possible to optimize the process of building strategic policies that aim to realize the economic potential of tourism in the region [5]. Here the question arises of determining the competitive advantages of the destinations, which select the level of development of each of them and make it attractive for tourists.

In any destination, the presence of tourist enterprises to create basic and additional services is of key importance for the formation of a complex tourist product and its marketing on the market. This whole process is related to the success rate of economic entities (enterprises) offering services, and hence the whole region and its economic indicators.

The choice of a tourist destination by tourists is related to all of the above. Their consumer behavior is the result of everything presented so far, but also of the personal interests of each of them. This is where the question of consumer awareness arises. Tourism behavior generally goes through four stages. The first is an awareness of the need to travel, namely the motives and personality characteristics of the individual. The second is the search for information, the third stage is the evaluation of travel alternatives and the fourth stage is decision-making.

Digitalization plays an important role here, as a result of which decisions are made through the analysis of certain consumer behavior, the development of business models and the management of tourist destinations. New users of tourist goods and services have already introduced technology into their daily lives.

The question arises what and which tools are suitable for analyzing user behavior and what benefits the data obtained brings.

2.2. Overview of Existing Research on the Use of Google Trends Data in the Field of Tourism

Google Trends is a tool that allows obtaining information about the popularity of keywords and its change over time. Its use in the field of tourism attracted research interest for the purposes of forecasting tourism demand. This section represents an overview of such studies, which shows that they vary in the approaches used and the destinations considered.

In [6], an auto-regressive distributed lag (ARDL) approach is applied in order to improve the accuracy of tourism search forecasting in relation to cities (Vienna and Barcelona) and countries (Austria and Belgium), based on web and Google Trends image indexes (Google Trends web and image indices). The archival data are monthly arrivals in Vienna, Barcelona, Austria, and Belgium and are extracted from an online tourism statistics database TourMIS (http://tourmis.info/).

Volchek *et al*. experiment to establish the existence of a correlation between Google Trends data regarding London's most popular museums and actual visits to these attractions [7]. In [8], a heuristic algorithm is applied by using Google Trends data in combination with tourist arrival data from the Taiwan Tourism Bureau.

Clark *et al.* [9] create a model for forecasting visits to U.S. National Parks using Google Trends data. Xiao *et al*. [10] propose a hybrid model combining ensemble empirical mode decomposition (EEMD) and deep belief network (DBN). For this purpose, Google Trends data regarding pre-identified keywords that indicate interest in traveling to Shanghai are included.

In [11], forecasting of tourist search for Amsterdam has been implemented through the Hidden Markov Model (HMM) with Google Trends data and archival data from the electronic database StatLine of the Central Bureau of Statistics (CBS) 2018.

In [12], an analysis is made of the behavior of Czech and Slovak tourists traveling to seaside resorts for a certain period of time based on Google Trends and statistics on the trips of Czech and Slovak tourists to beach resorts.

Çekim and Koyuncu [13] are building multivariate time series models for forecasting tourist arrivals based on data on international trade, weather conditions, and flight counts in combination with Google Trends data for Antalya [13].

In [14], the application of regression, decision tree, and random forest is investigated to predict the attendance of tourist attractions along the Moche Route in Peru. These algorithms are implemented after combining public data from official sources (such as data from the Peruvian government's open data platform) and Google Trends data.

The review of existing studies using Google Trends data in the field of tourism confirms the usefulness of this tool for forecasting tourist search.

2.3. Possibilities for Using Google Trends Data About Searches for a Given City

Google Trends search data for a city can significantly enrich a model based on archival visitation data for that city. Some of the ways this data can contribute are:

- Early warning of trends;

Google Trends data provides real-time information about interest in certain keywords or topics. This can serve as an early warning of future trends in tourist arrivals before they are reflected in archival data.

- Identify interest in attractions or events;

Search analysis can help detect growing interest in attractions in the city that may represent new tourist sites, events, or activities.

- Evaluation of the effectiveness of marketing campaigns;

Combining Google Trends data with archival visit data allows evaluating the effectiveness of marketing campaigns. Tracking changes in searches and matching visitor activity can help measure the results of various marketing efforts.

- Adaptation of tourist services;

Searches can provide information about the specific tourist services or experiences that potential visitors are interested in. This type of information can be used to personalize the services offered and improve the tourist services.

- Measuring interest on a global level;

Google Trends data can reveal interest in a city on a global level.

- Comparison with competitive destinations;

Google Trends allows comparing data for several different destinations to retrieve a clearer picture of how they are performing online. The information obtained could be useful in formulating strategies for competition with other tourist destinations.

- Identify specific trends and interests;

Search analytics can provide insight into specific tourist trends and interests.

- Improving the accuracy of predictions.

Combining the two types of data (archival and Google Trends) can improve the accuracy of predictions by taking into account how online interests are reflected in actual visits over time.

In conclusion, combining archival visitor data with Google Trends data provides a broader and more detailed perspective on tourism dynamics and can support making more informed decisions in the tourism sector. Combining data obtained from Google Trends with archival data on real visits (such as data from statistical institutes) allows for further analysis, the creation of prediction models and the assessment of their accuracy.

3. Results from Modeling Tourist Visits Using Google Trends Data

For the purposes of this study, data for Veliko Tarnovo District are represented. By administrative division, Veliko Tarnovo is a regional city in north-central Bulgaria, located on the 15th largest in the country. From the point of view of Tourist Zoning Concept (2015), Bulgaria has nine tourist regions, one of which is Stara Planina Tourist Region, with Veliko Tarnovo as its center.

To begin with, data are downloaded from the website of the National Statistical Institute of the Republic of Bulgaria [15]. The dataset used refers to the activity of accommodation establishments by statistical zones, statistical regions, and districts for the period January 2021 – October 2023. Data on the number of overnight stayers who are foreign citizens in Veliko Tarnovo District are extracted from it (Figure 1).

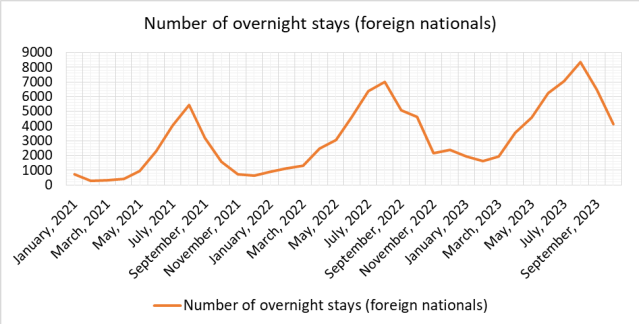

Figure 1. Number of overnight stays (foreign nationals) in Veliko Tarnovo District for the period January 2021 – October 2023 according to NSI data

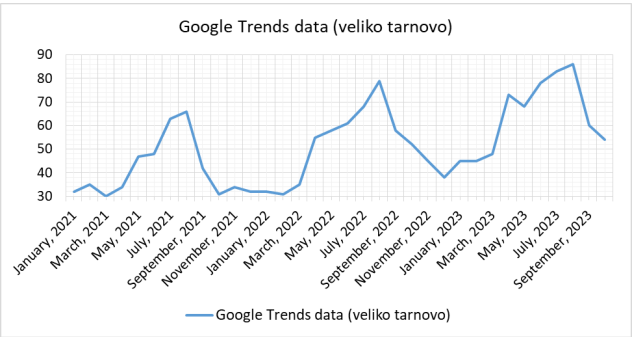

Figure 2. Keyword searches "veliko tarnovo" according to Google Trends for the period January 2021 – October 2023

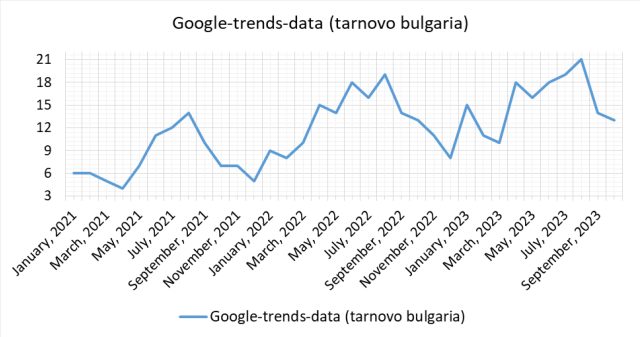

Figure 3. Keyword searches "tarnovo bulgaria" according to Google Trends for the period January 2021 – October 2023

Google Trends has extracted data on the interest in Veliko Tarnovo for the same period of time based on search reports by keywords "veliko tarnovo" (Fig. 2), "tarnovo bulgaria" (Fig. 3).

Polynomial regression is applied to build the model. The purpose of the model is to propose overnight stays prediction based on Google Trends data. The application of polynomial regression assumes that the conditions for normal distribution (C1) and homoscedasticity (C2) are satisfied.

C1. According to the Jarque-Bera test the data are normally distributed (the resulting *p*-values are given in Table 1).

Table 1. The resulting p-values resulting from the Jarque-Bera test

Data	***p-values***
Number of overnight stays (foreign nationals)	0.252586334
Google Trends data (veliko tarnovo)	0.354404561
Google-trends-data (tarnovo bulgaria)	0.545754585

C2. By the Breusch-Pagan test a homoscedasticity check is performed *(the resulting p-value is 0.0265150772, at a significance level α = 0.01),* we can assume that homoscedasticity is present.

The implementation of the model and the computation of the measures for its evaluation are carried out through RapidMiner (https://rapidminer.com). The created process involves normalization of values, resulting in values between 0 and 1. Using the built model, the predicted values are calculated in the manner shown in Figure 4.

PolynomialRegression

```
  0.739 * Google Trends data (veliko tarnovo) ^ 7.000
+ 0.177 * Google-trends-data (tarnovo bulgaria) ^ 8.000
+ 0.281
```

Figure 4. The resulting model by applying polynomial regression

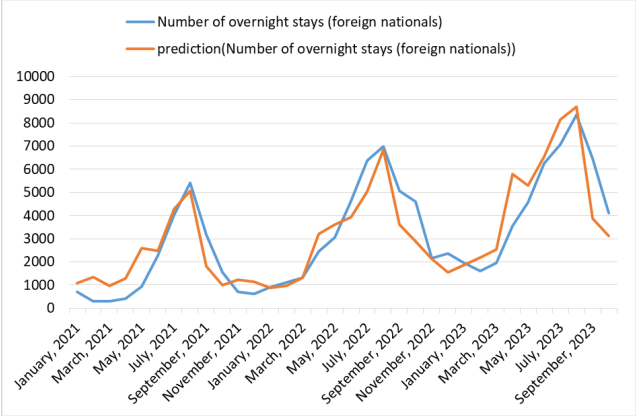

Figure 5. Number of overnight stays (foreign nationals) in Veliko Tarnovo District and predicted values

Table 2 summarizes results of measures obtained from validation of the proposed model.

Table 2. Results of measures obtained from validation of the proposed model

Measure	**Result**
Root mean squared error (RMSE)	0.231
Mean absolute error (MAE)	0.177
Mean Squared Error (MSE)	0.061
Pearson correlation	0.757
Spearman correlation	0.911
Kendall gets correlation	0.788

- Pearson correlation;

The Pearson correlation coefficient is a measure of the relationship between actual and predicted values.

- Spearman correlation;

It represents rank correlation between actual and predicted values. In contrast to Pearson's correlation, the Spearman correlation measures the monotonous association (only strictly increasing or decreasing, but not mixed) between two variables and relies on another order of values.

- Kendall tau correlation.

Kendall's rank correlation coefficient measures the strength of the relationship between two actual and predicted values.

The computed coefficients confirm that in this case the actual and predicted values are highly correlated, i.e. there is a high positive correlation [16].

4. Discussion

The management of a tourist destination is a strong dependence on information, which is key to creating value in the economy. According to Kaplan and Norton [17], the modern kind of information is contained in information capital, which can be a system, a database, a library, a network, specialized software, applications, etc. They define as successful information systems that carry information about user preferences and behavior, improve service and assist in preserving a database [17].

Going back to the presented system model of tourist destination management and putting information capital into this system, data analytics with Google Trends brings key value to the external and internal environment impacting the destination. The data creates conditions for the development of management strategies for attracting tourists based on primary and secondary data for analysis.

The statistical data (NSI) from Figure 1 are the starting point for analyses related to planning the supply in the destination aimed at foreign tourists. They show the current states of the destination and how much of the bed potential of the accommodation establishments in the district were occupied in different periods. From the data there is a curve describing seasonality in the interest and realization of nights spent by foreigners, strongly expressed in the months of July-September. The exported data serve to compare nights spent and tourist search by keywords. The applications of the received data sets are important for tourism organizations and institutional bodies in tourism (macro and meso level). The resulting analysis influences issues related to the management and marketing of the destination. The second directly affects the segmentation of the market and the distribution of products and services to the end user (micro level).

Keyword search using Google Trends data is important for getting information as close as possible to actual subsequent consumption. The indication of dips and peaks of the curves of Figure 2 and Figure 3 are evidence of erratic interest in demand for the period indicated. Another issue is the pre-planning of a tourist trip. If we compare the search for keywords and the time of the nights spent according to NSI data, it can be concluded that less and less time is devoted to planning or it is carried out almost immediately before the decision to travel. This action may be provoked by various factors, but one of the things that have changed long-term and medium-term tourism consumption is the emergence of COVID-19. The global pandemic has changed consumer early planning habits and created a new reality in tourism, namely making travel decisions just before its implementation.

These information models and analysis applications provide a great opportunity for businesses and tourism organizations at all levels and are a prerequisite to monitor current conditions and to make informed and well-grounded management decisions. They are related to the quality and quantity of primary data that goes through the stages of research, information collection, and decision-making. This process introduces the concept of smart destination, precisely because of well-grounded marketing, management, and consumer orientation. These three elements in the process form the stages of destination maturity as part of smart tourism. On this issue, the European Commission to support EU cities and facilitate access to tourism and hotel products and services through technological innovation is setting up the smart tourism destinations initiative [18]. The aim is for EU cities to learn to implement innovative digital solutions to make tourism sustainable and accessible, taking full advantage of their cultural heritage and creativity and to enhance the tourism experience. Such tools raise the question of integrated computer solutions and benefits for both tourists and creators of tourist services.

On the other hand, the cultural and national characteristics of tourists are the next important prerequisite for creating a suitable tourist product with high basic and added value. These may be topics for study in the authors' next scientific search.

5. Conclusion

The management of tourist destinations requires a complex of methods, tools, and data for the analysis of environmental factors. There are different preconditions for the actuality of the offered tourist products and services, one of the most significant being the tourist potential as a basis for the emergence of tourist interest. These processes carry the complexity of planning and forecasting consumer interests and consumption, as well as applying the most appropriate computer solutions for analysis.

This study shows the advantages and applicability of Google Trends data on the example of Veliko Tarnovo District, as well as the methods and means of validating the data obtained when applying the model. The proposed search and analysis tool is suitable for the dynamic environment in which the tourism business operates.

The research gives a real idea of the processes on the topic of destination management through computer systems, but also provokes new questions related to consumer behavior and the emergence of interest to travel in a specific destination.

Planning and making quick decisions to choose a destination are a reality in modern tourist travel and the specific reasons and appropriate tools for analyzing this fact should be sought. Information is a key factor for making correct management decisions and achieving high economic indicators for each tourist destination.

References:

[1]. UNWTO. (2007). *A Practical Guide to Tourism Destination Management*. Madrid: UNWTO. Retrieved from: https://www.unwto.org/global/publication/practical-guide-tourism-destination-management [accessed: 28 January 2024].

[2]. Republic of Bulgaria. (2021). *Tourism Act.* Republic of Bulgaria. Retrieved from: https://www.tourism.government.bg/bg/kategorii/zakoni/zakon-za-turizma [accessed: 29 January 2024].

[3]. Davidson, R. & Maitland, R. (1997). *Tourism Destinations*. London: Hodder & Stoughton.

[4]. Marinov, S. & Kazandzhieva, V. (2023). *Monitoring the Sustainable Development of a Seaside Tourist Destination*. Cambridge Scholars Publishing.

[5]. Kostadinova, N. (2023). Business Hotels and Congress Tourism – Interactions and Practices. Veliko Tarnovo: IB. ISBN: 978-619-7281-99-6, 88.

[6]. Önder, I. (2017). Forecasting Tourism Demand with Google Trends: Accuracy Comparison of Countries vs. Cities. *International Journal of Tourism Research, 19*(6), 648-660. Doi: 10.1002/jtr.2137.

[7]. Volchek, E., Song, H., Law, R., & Buhalis, D. (2018). Forecasting London museum visitors using Google trends data. *E-Review of Tourism Research*.

[8]. Huarng, K.-H. & Yu, T. H.-K. (2019). Application of Google Trends to Forecast Tourism Demand, *Journal of Internet Technology, 20*(4), 1273-1280. Doi: 10.3966/160792642019072004025.

[9]. Clark, M., Wilkins, E.J., Dagan, D.T., Powell, R., Sharp, R.L., & Hillis, V. (2019). Bringing forecasting into the future: Using Google to predict visitation in U.S. national parks. *Journal of Environmental Management, 243*, 88-94.

[10]. Xiao, Y., Tian, X. & Xiao, M. (2020). Tourism Traffic Demand Prediction Using Google Trends Based on EEMD-DBN. *Engineering, 12*(3). Doi: 10.4236/eng.2020.123016.

[11]. Claude, U. (2020). Predicting Tourism Demands by Google Trends: A Hidden Markov Models Based Study. *Journal of System and Management Sciences, 10*(1), 106-120. Doi:10.33168/JSMS.2020.0108.

[12]. Kajzar, P., Dolák, R. & Krkošková, R. (2021). Using Google Trends in International Tourism: A Case Study of the Czech and Slovak Republics. *Academica Turistica, 14*(2), 205-216. Doi: 10.26493/2335-4194.14.205-216.

[13]. Çekim, H. Ö. & Koyuncu, A. (2022). The Impact of Google Trends on the Tourist Arrivals: A Case of Antalya Tourism. *The Journal of Operations Research, Statistics, Econometrics and Management Information Systems, 10*(1), 1-14. Doi: 10.17093/alphanumeric.931652.

[14]. Bravo, J., Alarcón, R., Valdivia, C. & Serquén, O. (2023). Application of Machine Learning Techniques to Predict Visitors to the Tourist Attractions of the Moche Route in Peru. *Sustainability, 15*(11). Doi: 10.3390/su15118967.

[15]. NSI – National Statistical Institute. (2023). *Activity of accommodation establishments by statistical zones, statistical regions and districts*. NSI. Retrieved from: https://www.nsi.bg/en/content/1973/activity-accommodation-establishments-statistical-zones-statistical-regions-and-districts [accessed: 30 January 2024].

[16]. Mukaka M.M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71.

[17]. Kaplan, R. & Norton, D. (2006). *Strategy Maps*. Vlassic and sryle. ISBN: 9543270244, 259-262.

[18]. Galasso, G., Bersani, C., Messina, T. & Venditti, A. (2023). *Mastering Data: A Toolkit for Tourism Destinations.* Smart Tourism Destinations. Retrieved from: https://smarttourismdestinations.eu/wp-content/uploads/2023/03/Smart-Tourism-Destinations_Toolkit_v.3.1-BG.pdf [accessed: 30 January 2024].