Business intelligence and open data: The possibilities for the derivation of valuable information in tourism domain

Marija Blagojević¹, Miloš Papić¹, Dušan Garabinović²*

¹ University of Kragujevac, Faculty of Technical Sciences Čačak, Serbia
² Higher Business School of Vocational Studies “Prof. dr Radomir Bojković”, Kruševac, Serbia

Abstract: This paper aims to introduce the concept of data analysis which could easily be implemented by anybody involved in the subject matter with basic IT knowledge and skills. The paper is divided into two parts, the first of which presents an overview of related research from two points of view: (1) publications which refer to the analysis, or the overall use of open data from the tourism domain and (2) publications which use business intelligence tools to analyse tourism data. Results indicate that there is a significant number of publications but none of them combines the two issues in the field of tourism (open data and business intelligence). The second part refers to the possibilities of using Power BI, the business intelligence tool for analysing available open data about tourism in Serbia.

Keywords: tourism, open data, business intelligence, overview, related research

JEL classification: Z32, Y10, C55

Poslovna inteligencija i otvoreni podaci: Mogućnosti za izvođenje vrednih informacija u oblasti turizma

Sažetak: Ovaj rad ima za cilj da predstavi koncept analize podataka koji bi svi koji se bave predmetnom materijom s lakoćom mogli primenjivati, pri čemu im je potreban osnovni nivo IT znanja i veština. Rad je podeljen u dva dela pri čemu prvi predstavlja pregled srodnih istraživanja i to sa dve tačke gledišta: (1) pregled publikacija koje se odnose na analizu i generalno upotrebu otvorenih podataka iz oblasti turizma i (2) pregled publikacija u kojima se alati poslovne inteligencije koriste za analizu podataka u vezi turizma. Rezultati ukazuju da postoji značajan broj publikacija, ali nijedna od njih ne obraduje istovremeno upotrebu otvorenih podataka i poslovne inteligencije u oblasti turizma. Drugi deo se odnosi na mogućnosti korišćenja Power BI alata za poslovnu inteligenciju za analizu dostupnih otvorenih podataka o turizmu u Srbiji.

Ključne reči: turizam, otvoreni podaci, poslovna inteligencija, pregled, srodna istraživanja

JEL klasifikacija: Z32, Y10, C55

*dusan.garabinovic.032@gmail.com
** This study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and these results are parts of the Grant No. 451-03-68/2020-14/200132 with University of Kragujevac – Faculty of Technical Sciences Čačak.
1. Introduction

As a land of rich history, Serbia is located from a cultural point of view on the border between East and West, geographically speaking, located at a place which enables its future to be built in the direction of its tourism development potential. Serbia was ranked 83rd in 2019 in the tourism competitiveness chart, which is an astonishing 12-position rise compared to 2017 (World Economic Forum, 2019).

Numerous spots in Serbia which have enormous tourism potential are yet to be open to a wider population of tourists. A need for marketing and promotion quality improvement is self-imposed. However, it is vital to appropriately target marketing, and for that, valid information obtained by analysing appropriate data is needed.

In the past couple of years, the world has been aiming towards opening up data in different areas of creativity. That trend has been followed by Serbia as well where data is usually opened up by public institutions. Consequently, they offer the possibility to their citizens to obtain, process and analyze them in the desired manner. Thus, public institutions achieve a higher level of transparency in their work, and the citizens are able to indirectly contribute to the important decision-making.

Collections of open data for Serbia are located at specialized portals such as Open data Portal (data.gov.rs) and are usually stored in one of the open formats such as CSV, XML, and JSON. The main characteristic of open formats is their machine-readibility, which implies that the collection of data can be automatically processed and analysed through one of the open softwares. There are many open-source business intelligence tools. In this paper, we used Microsoft Power BI. Although not entirely open, Power BI has enough free features for beginners.

One of the goals of this paper is to determine the possibility of implementing this tool to the analysis of available open data with regards to the tourist visits in Serbia in the last decade. The second important goal of this paper is an overview of related research from two perspectives: (1) publications which refer to the analysis, or the overall use of open data from the tourism domain and (2) publications in which business intelligence tools are used to analyse tourism data.

2. Methodology

The data on publications which deal with the subject of this paper was obtained by using the Google Scholar search engine (https://scholar.google.com/) on May 2020. The criterion for choosing the publications was the keyword in the title which directly refers to open data and business intelligence while combining terms which refer to tourist (e.g. tourism, tourist, tourists, touristic, hotel, hotels, hospitality, etc.). Dimitrovski et al. (2019) used similar methodology for conducting “A bibliometric analysis of Crossref agritourism literature”. For the realisation of the aforementioned search, the advanced search provided by Google Scholar was implemented by finding articles which include all the keywords and at least one keyword in the title. A title suitability check for the theme analysed was subsequently carried out by the authors. Papers found in journals, conference proceedings and parts of thematic collections, like books and monographs written entirely in English were used.

For the purposes of the practical part of research, the data from The Statistical Office of the Republic of Serbia were used, available at the Open Data Portal (Open Data Portal, 2020). Choosing the catering and tourism category one arrives at several collections of open data for which analysing the collection named Tourist arrivals-monthly data was chosen. Located in it is the data on the number of local and foreign tourists by months, years, and regions of the
Republic of Serbia (the Region of Šumadija and Western Serbia, the Belgrade Region, the Region of Southern and Eastern Serbia and the Vojvodina Region).

The data is available in the.xls format and prior to the analysis it was necessary to perform data pre-processing. After that, a *Microsoft* tool for business analytics Power BI was used to carry out the analysis. The tool allows connection of different data sources and provides powerful reports. Power BI provides a possibility of integration with Excel which is significant for the users who are used to working in Microsoft environments.

3. Results and discussion

3.1. Results of the search for topic-related papers

There is a significant number of publications which deal with issues regarding open data use in the tourism field (*Miele & Mola, 2005; Cao et al., 2011; Groen et al., 2013; Cannataro et al., 2013; Longhi et al., 2014; Wu et al., 2014; McNaughton et al., 2014; Okuno, 2014; Pereira et al., 2015; Fermoso et al., 2015; Bue & Machi, 2015; Villa, 2015; McLeod & McNaughton, 2015; Keler & Mazimpaka, 2016; Pesonen & Lampi, 2016; Kršak et al., 2016; McNaughton et al., 2016; Mekhabunchakij, 2016; De Vocht et al., 2016; Li & Hsia, 2016; Pantano et al., 2017; Urata et al., 2017; Mekhabunchakij, 2017; Sidor et al., 2017; Scorza et al., 2018; Ammur & Meidelfi, 2018; Sedlak & Ivanišević, 2018; Maita, 2018; Al-Ghossein et al., 2018; Ocampo & Palaoag, 2019; Pantano et al., 2019; Duca & Marchetti, 2019; Le & Cao, 2020; Mountasser et al., 2020; Yochum et al., 2020*), as well as the implementation of business intelligence tools for tourism data analyses (*Minghetti et al., 2000; Carson et al., 2003; Zimmerman et al., 2004; Vrdoljak-Salamon et al., 2007; Salguero et al., 2008; McKnight, 2008; Galicic, 2009; Lozada et al., 2010; Vizjak et al., 2010; Custis, 2012; Crockett, 2012; Fuchs et al., 2013; Angelaccio et al., 2013; Bazdan, 2013; Sharma et al., 2013; Korte et al., 2013; Daryaei et al., 2013; Alzua-Sorzabal et al., 2014; Verma, 2014; Höpken et al., 2015; Fuchs et al., 2015; Martins et al., 2015; Baggio, 2016; Höpken & Fuchs, 2016; Teimouri et al., 2016; Vajirakachorn & Chongwatpol, 2017; Hyseni, 2017a; Hyseni, 2017b; Ramos et al., 2017; Mariani et al., 2018; Bilandzic & Lucic, 2018; Chen et al., 2018; Chen, 2018; Stylos & Zweigelaar, 2019; Nyanga et al., 2019; Godnov & Redek, 2019*). The interesting fact is that there are no papers combining the two issues in the area of tourism (open data and business intelligence).

The following chart (Figure 1) shows a movement in the number of publications about open data and business intelligence in tourism domain.

**Figure 1: Number of publications on open data and business intelligence in tourism**

Source: Author’s research
The first paper on business intelligence in tourism was published in 2000, while the first publication on open data in tourism dates back to 2005 after which follows a five-year break. Since 2011, i.e. 2012 there is a notable increase in the number of publication about the two issues.

The following chart (Figure 2) shows the number of publications by analyzed categories for open data and for business intelligence.

![Figure 2: Number of papers by categories](image)

As it can be seen from Figure 2, the largest number of papers related to open data and business intelligence (total number is 71) was published in conference proceedings (27; 38.03%), followed by papers published in journals which are not listed in Thomson Reuters Web of Knowledge and do not have an impact factor (23; 32.39%), papers in journals from Web of Science database, which are a part of Thomson Reuters Web of Knowledge (11; 15.49%), and papers in other types of publications – chapters in books, monographs, etc. (10; 15.08%).

3.2. An overview of the practical application of Power BI tools on the experimental collection of data

The capabilities of the Power BI tool are presented in the experimental collection of data through specific analyses shown in figures 3-8. As input parameters for the analyses shown in figures 3-6, the following basic analysis results were used:

- an overview of the type of tourists (local, foreign);
- tourist visits by Serbian regions;
- yearly tourist visits spanning from 2010-2020;
- monthly tourist visits.

The abovementioned basic analysis were carried out but were not presented in this paper because they are extremely simple – they can be carried out in any spreadsheet software with a basic skill level. Their results indicate that in the last decade, Serbia had slightly more visits by local tourists; the most visited region was the Šumadija and Western Serbia Region; the number of tourists has been growing yearly since 2010, while in 2014 there was a slight fall. The current year (2020) has been excluded from further analysis in this paper.

Figure 3 shows an analysis of the tourist visits’ decrease in 2014 according to their type (local/foreign). There was a drastic decrease in local tourists’ visits, while the number of foreign tourists increased.
Analysing the decrease in the number of tourists’ visits by region in 2014, the biggest plummet in visits is noticed in the Šumadija and Western Serbia region (Figure 4).

Analysing the number of tourist visits by months in the last 10 years, it was revealed that the largest number of visits was realised during the month of August (11.6%), and the smallest during November, where the largest decrease is noticed compared to the region (Figure 5).
The biggest plummet at the time was noted in the Šumadija and Western Serbia region. Additionally, the decrease in the number of visits was more pronounced in local tourists as opposed to foreign ones whose number was at an average increase (Figure 6).

![Figure 6: Analysis of the decrease in tourist visits in November according to type](source)

In addition to the simple analysis application which is not simple to carry out in the spreadsheet software, Power BI provides the ability for one to intuitively pose a research question to which it provides an answer. An example of the question posed and the answers given is shown in Figure 7. The research question refers to the number of visits by regions and months but with a difference compared to the month of November.

![Figure 7: Question asked and a given answer](source)

Given that the ultimate goal of the tourism sector is the increase of tourist visits, identifying the key factors which could lead to it is crucial for the decision makers in this field. Power Bi has an option which helps one indentify the mentioned key factors – the tool’s answer to the question *What influences the increase of the number of tourists* is that the increase of tourist...
visits in the entire country highly depends on the number of visits in the Šumadija and Western Serbia region.

4. Conclusion

Considering the research goals set in this paper, and acknowledging the results obtained, conclusions are drawn from several directions:

- According to the number of similar research for both aspects of the research (35 for the use of open data in tourism and 36 for the use of business intelligence in tourism), according to the period when research data is realised, it can be said that the subject of this paper is very up-to-date and it has a grounded position in contemporary science;
- The capabilities of Power BI as a business intelligence tool are significant for analysing available open data in the tourism field. Its use is in the forefront when the information obtained after initial analysis is in need of deeper analysis. It is important to note that such analysis is impossible to be carried out in the basic skill versions of the well-known spreadsheet software.
- The literature available, data openness and free access to the business intelligence software should serve as a stimulus for the tourism sector decision makers themselves to arrive at valuable information in similar ways. As it is shown in this paper, an advanced knowledge in statistics or computer science is not necessary for data search and the use of BI software, since the tool itself has indications of artificial intelligence.

The authors’ future work on this issue refers to predicting the number of Serbian tourists by year following regions and months using the data mining technique.

References


Received: 23 July 2020; Accepted: 19 October 2020