Hotel web site SEO analysis : Segmentation and valorization as a precondition for discovering and understanding insights to improve online visibility

Šimunić, Mislav; Pilepić Stifanich, Ljubica; Car, Tomislav

Download date / Datum preuzimanja: **2025-01-08**

SVEUČILIŠTE U RIJECI
FAKULTET ZA MENADŽMENT
U TURIZMU I UGOSTITELJSTVU
OPATIJA, HRVATSKA

Repository / Repozitorij:

Repository of Faculty of Tourism and Hospitality Management - Repository of students works of the Faculty of Tourism and Hospitality Management

UNIri DIGITALNA KNJIŽNICA

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

Mislav Šimunić
University of Rijeka
Faculty of Tourism and Hospitality
Management, Opatija
51410 Opatija, Croatia
mislavs@fthm.hr

Tomislav Car
University of Rijeka
Faculty of Tourism and Hospitality
Management, Opatija
51410 Opatija, Croatia
tcar@fthm.hr

Ljubica Pilepić Stifanich
University of Rijeka
Faculty of Tourism and Hospitality
Management, Opatija
51410 Opatija, Croatia
ljubicap@fthm.hr

HOTEL WEB SITE SEO ANALYSIS: SEGMENTATION AND VALORIZATION AS A PRECONDITION FOR DISCOVERING AND UNDERSTANDING INSIGHTS TO IMPROVE ONLINE VISIBILITY

Abstract

Purpose: The purpose of this article, in which the authors present the results of the scientific project "The relationship between the quality of the main SEO ranking factors and the position of a hotel brand's website in the SERP" ZIP-FMTU003-11-2021, is to (1) identify and segment SEO variables relevant for better online visibility, (2) define a valorization model of the SEO variables, (3) valorize and define the mutual relationships of the observed SEO variables, and (4) identify and interpret the influence of the quality of the observed SEO variables on the quality of the websites of the observed sample of hotels.

Methodology: The research was conducted in two main phases. In the first phase, the available scientific and professional literature was reviewed and examined. In the second phase, the analysis model was designed, the graphical interface of the analysis model was created, the hotel sample was defined, the necessary data of the observed variables were collected, and finally, data analysis and interpretation was performed.

Results: In this article, (1) an approach to creating an analysis model and its graphical interface is described and explained, (2) a graphical interface and the analysis model are presented, (3) the algorithm for valorization of selected SEO elements is explained, (4) and finally, the results of the conducted analyses are presented.

Conclusion: The results of the conducted research show the importance and complexity of researching the topic of online visibility, which opens numerous possibilities for conducting similar research.

Keywords: Hotel, website, SEO, segmentation, valorization

1. Introduction

The quality of websites is determined by a large number of variables. The quality of each of these variables impacts the overall quality of the website and online visibility. Google uses a variety of SEO ranking factors to determine where a particular hotel website will appear on search results pages. Not only are some of these ranking factors more important than others, but the degree of their importance changes as Google's algorithm evolves. Therefore, ensuring high quality in the presentation of online hotel information and maintaining high quality in the optimization of key SEO ranking factors is of great importance to the success of any business, including hotels, Law (2019). This is especially important in the hotel industry, where today more than 90% of sales activity ("booking" - sales) in Europe is done through online channels (D-edge hospitality, 2021; Phobs CRS, 2021). Today, when we talk about online sales in the hotel industry, sales through indirect sales channels dominate (Vukasović & Mihač, 2021). However, it is best for each hotel to realize as much of its sales as possible directly, i.e., through its own hotel website. Direct sales via the hotel website/website largely depend on the position in search engine results (visibility in search engines). The reason for writing this article is the increasingly complex influence and importance of SEO factors and search engine algorithms, artificial intelligence and machine learning on visibility in Internet search engines, which is especially evident when we talk about increasing direct sales through the hotel website.

Design/Methodology/Approach - Based on the available literature, preliminary research was conducted to obtain new scientific knowledge. After that, empirical research was conducted, which was divided into 2 parts. The first part of empirical research was conducted on a defined sample of websites of 5-star hotels in the Republic of Croatia (19 hotels - 45.24%), using one of the world's best commercial "all in one" software for analyzing and increasing search engine visibility. The second part of empirical research involved the creation of an analytical model and its graphical interface, as well as an algorithm for secondary data processing. This will provide a basis for relevant decisions in SEO activities. The first part of research (collection and processing of primary data) was conducted in the period from 11/2021 to 01/2022. The second part of research was conducted in 2 phases, namely: (phase 1) preparation and creation of an analytical model, graphical interface and software solution algorithms for secondary data processing from 01/2022 to 06/2022, and (phase 2) secondary processing, analysis and interpretation of the obtained data (06/2022 - 08/2022).

Hypotheses/Research question - The proof or rejection of the established hypotheses is based on the conclusions derived from the facts obtained from the interpretation of the results of theoretical research through the analysis of available literature and empirical research through the described research processes and their sub-processes on the websites of a selected sample of 5-star hotels in the Republic of Croatia. In the initial phase of research, 3 basic research hypotheses were established:

H1: The position of the hotel website in the search results of Internet search engines depends on the quality of optimization of the main SEO ranking factors for ranking,

H2: The segmentation and valorization of the main SEO factors provides a better basis for decision making with the aim of increasing online visibility in search engines.

H3: The online direct sales of hotels can be significantly increased through recognition and better optimization of Google ranking factors.

Thus, **the main objective of this research** is to (1) identify, segment and evaluate the most important SEO factors for ranking in order to achieve better online visibility – the position of the hotel website on the SERP, and (2) determine the relationship and interpret the relationship between the quality of SEO factors for ranking and the position in search results based on search engine queries.

2. Preliminary research - theoretical and conceptual background/framework

As mentioned in the introduction, preliminary research included an analysis of the literature and previous research on the described topic. In this sense, Yalçın & Köse (2010) emphasized the importance of SEO and examined the interdependence of search engine optimization, page rank, backlinks, meta tags, and social media. They concluded that SEO is a dynamic process that needs to be continuously monitored to track positive or negative changes to improve the website. Dean (2020) explains and comments on Google's 200 ranking factors that were relevant in 2020 based on his practical experi-

ence in e-business for many years. Yu (2018) writes about the importance of measuring website quality. In his research, he highlights the following factors as the most important qualitative factors to be analyzed: content, features, structure, usability/usability, the quality of navigation, technical performance in terms of compatibility with different platforms, interoperability in terms of testing the quality of interaction with applications and databases, security, and configuration in terms of testing the quality of work with complex hardware and software systems. In addition, knowledge is also expanded by the articles that deal with research on SEO analysis tools such as Google Trends, SEMRUSH, SimilarWeb, Alexa, Moz-Open Site Explorer, etc. (Vyas, 2019), and the importance of individual Google ranking factors for achieving a better position/visibility in search engines (Pan, 2015). Shaolang et al. (2019) highlight the importance of data analysis when drawing relevant conclusions and conducting future SEO activities. In their research, they propose a framework that uses machine learning and search indexes to predict tourist arrivals by comparing the predictive performance with Google and Baidu generated search results. Furthermore, Konidaris et al. (2022) emphasize that selecting appropriate keywords is a fundamental task for all SEO and SEM activities. Iqbal et al. (2022) investigate the key factors to achieve a better SERP position. In their study, they highlight the following 7 main groups of factors that should be considered when developing and maintaining websites: website design, performance, target market, responsiveness, security, fresh content, and SEO usage. In their article, they present an interesting SERP equation for predicting the ranking of business websites. Garcia et al. (2022) present the application of different SEO techniques to improve the indexation of websites on Google, where the implementation of the implemented strategy ensures the improvement of the position within the SERP. For this article, the work of Rodlej (2020) is interesting, in which the author, in the article entitled "Hotel SEO: How to guide", describes modern SEO, its importance and the way of using and implementing it in the modern hotel industry. Sengar (2022) explores the advantages, disadvantages and difficulties of the commercial application of search engine optimization. After completing preliminary research, during which a large number of articles were examined (the most relevant ones are all listed in the literature section), the authors compared previously selected articles with their previous research and, taking into account the new

findings, conducted the rest of the research, which is described in the introductory part of the article under the design/methodology/approach part.

3. Empirical research process and methodology

The first part of empirical research was conducted on a defined sample of hotels in the Republic of Croatia using Semrush SaaS – "all in one" software for website optimization analysis. The value of Semrush lies precisely in the fact that it contains a wide range of available analytical tools - the software contains over 40 tools for SEO as well as for content, advertising, competitive research, reporting and social media management. The analysis by means of the aforementioned tools allows us to gain new insights for the purpose of better management of the online presence of hotel websites.

The tasks in this part of research were as follows:

1. define individual projects for data processing,
2. parameterize the processes in the defined projects,
3. monitor the processes within the projects in the defined period,
4. collect, analyze and isolate the data defined for primary research, and
5. group and prepare primary data for secondary processing.

In the second part of empirical research it was necessary to:

1. define the approach and design and create a graphical interface for data segmentation and exploitation,
2. define the required functions and procedures,
3. define and develop an algorithm for secondary data analysis and processing,
4. analyze and evaluate the obtained data,
5. interpret the research results, draw sound conclusions and present our own views and opinions.

In this way, the basis for relevant, scientifically sound conclusions is established. Such an approach not only brings new scientific knowledge, but ultimately ensures the possibility of application in modern hotel management theory and practice.

3.1 Empirical research - the first part

For the first part of empirical research, a sample for conducting research was determined at the beginning. The sample consists of 19 5-star hotels in the Republic of Croatia, which represents 45.24% of the total number of 5-star hotels in the Republic of Croatia. In addition, a license (2-month plan GURU) was purchased for the use of Semrush SaaS, so that a detailed analysis of the quality of the websites of the hotels selected in the sample could be performed. Then, based on the procedures available in Semrush SaaS, (1) a new analysis project was defined for each observed hotel (Figure 1), (2) analysis parameters for each analysis project were defined (Figure 2), (3) based on the preliminary research results, Semrush components and functions/analytical interface (dashboard) were analyzed in order to create a relevant model for secondary data processing, and (4) the data were isolated in order to develop and create a graphical interface and algorithm for secondary analysis.

Figure 1 Defining analytical projects within Semrush SaaS (each hotel – one project)
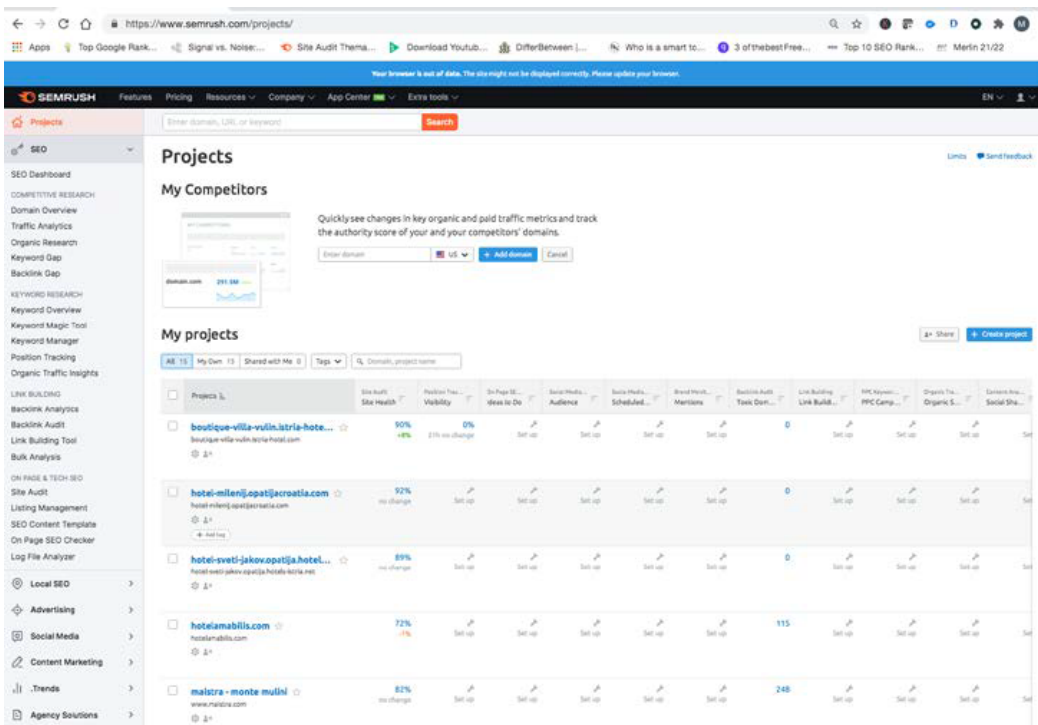
Source: Authors - Screenshot – Semrush graphical interface

Figure 2 Semrush SEO Dashboard - displaying analytical reports for a hotel
(Definition of analytical parameters for each project, where each hotel represents one project)
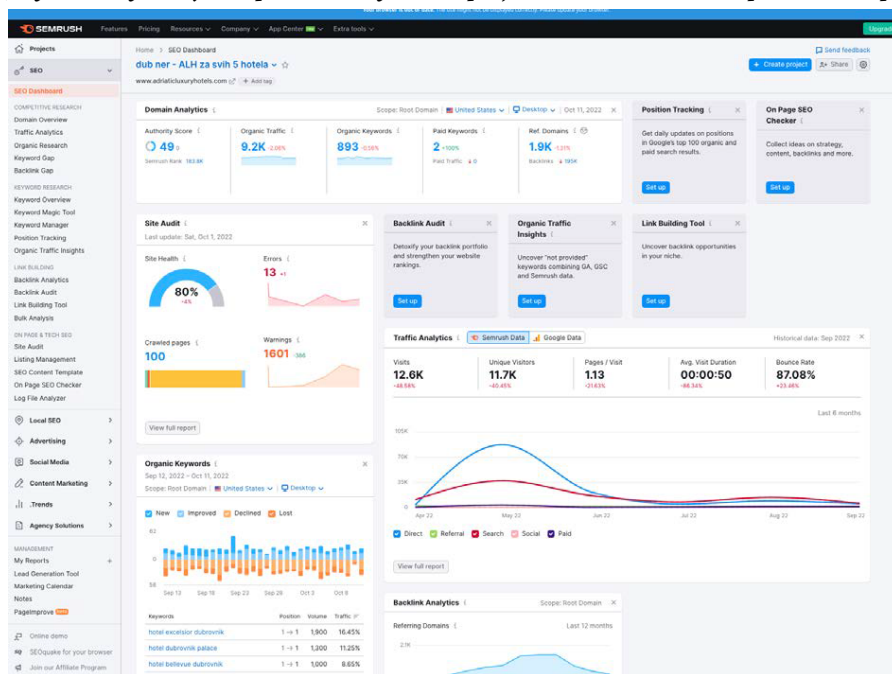
Source: Authors - Screenshot – Semrush graphical interface

As shown in Figure 2, a complex SEO analysis was performed for each hotel. Based on the analysis performed in this way, a very large amount of data related to SEO, PPC: pay-per-click ad campaigns, social media marketing, keyword research, competitive research, PR, content marketing, marketing insights and campaign management were collected for each hotel. Since the focus of this research is on the study of algorithmic (organic) search in the context of online search engines, the rest of the article focuses on the analysis of the data obtained through the segment of SEO analysis. Semrush SEO toolkit contains over 20 tools and reports that are very helpful in optimizing the parameters for better online visibility of the hotel website. The result of the first phase of empirical research is a huge amount of data, from which it was necessary to extract the most relevant ones for the further course of research.

3.2 Empirical research - the second part

Based on the findings from preliminary research and the data collected in the first phase of empirical research, the second phase of empirical research

began. The most relevant data were isolated from the Semrush GUI. The isolated data formed the basis for the creation of an analysis model and graphical user interface for secondary data analysis (a new analysis model and user interface design for data analysis). The isolated data model for secondary analysis was realized by 3 main analytical modules (M_1, M_2, M_3), and 1 valorization module (M_V) for the final analysis, which includes 1) the interpretation of the obtained absolute values of the observed variables - *where absolute value Av $\in$ 1,2,3,...n*$\rangle$, 2) the definition of the range and class of the absolute values of the variables - X_{imax} - X_{imin}, 3) in terms of the absolute values of the variables - the association of the score/point value - y_j, and 4) the analysis of the impact of the values of each variable on a) the total value of each analytical module, and b) overall, on the search engine optimization, giving measures of dispersion, arithmetic mean $\overline{xi}$, standard deviation σ, variance $σ^2$ and Spearman's rank correlation coefficient - R_s. The data analysis model was created by isolating and segmenting the most relevant data obtained by performing defined analytical projects using Semrush

SaaS. The created analytical model is based on the analysis of 337 variables for each analyzed hotel in the selected sample - $h_r \in 1,2,3,...19$. Since the subject of this study is a sample of 19 5-star hotels in the Republic of Croatia, 6,403 analytical variables were treated in this study. Based on their independent interpretation and the interpretation of their mutual relationships, the final conclusions of the study are drawn and their own opinions and attitudes are presented. The modules of the analytical data model are described below.

Module 1: M_1 - Domain overview

This module contains a total of 81 analytical fields analyzed for each hotel. Module M_1 analyzes:

(1) Authority Score: a metric used to measure the overall domain quality and SEO influence. The score is based on the number of backlinks, referring domains, organic search traffic, and other data. 100 is the highest possible authority score.

(2) Organic search traffic: The organic search section displays widgets related to the domain's organic traffic, the keywords for which the domain ranks in the top 100 of Google organic search results, the distribution of keyword rankings, and the domain's organic competitors.

(3 Backlinks: These data provide information about the number of backlinks the domain has.

(4) Distribution of organic traffic of the domain and keyword by country.

(5) Organic research (includes keyword analysis, organic position distribution analysis, and keyword analysis by intent). Additional analyses allow the study of the relationship between organic/algorithmic domain traffic and keywords by the analyzed countries.

The aforementioned analyses were used to examine the qualitative values of the observed variables of the observed hotel sample. $M_1TP_{(hr)}$ - M1Total-Points represents the total qualitative value of the domain whose quality is examined by analyzing the variables for the hotel (hr) just observed. This value was determined based on the sum of the individual qualitative point values of the observed variables x_i f_j within this module:

$$M_1TP_{h_r} = \sum_{i=1}^{3} x_i \ f_j \qquad (1)$$

Module 2: M_2 - Backlink analytics overview

This module contains 97 analytical fields. The data are analyzed using 7 main analyses and 9 auxiliary analyses, namely:

(1) Authority Score: A metric used to measure the overall quality and SEO influence of a domain. The score is based on the number of backlinks, referring domains, organic search traffic and many other data. 100 is the highest possible authority score.

(2) Referring domains: the total number of referring domains that have at least one link pointing to the analyzed domain/URL. Only referring domains from the last few months are considered.

(3) Backlinks: the total number of backlinks pointing to the analyzed domain/URL. Only backlinks from the last few months are taken into account.

(4) Monthly visits: the total number of unique visits to the root domain in the past month.

(5) Keywords: a number of keywords that lead users to the root domain via the top 100 Google organic search results.

(6) Outbound domain: the total number of domains the analyzed domain is linked to. Only domains referred to by the analyzed domain in the last few months were taken into account.

(7) Toxicity: The toxicity rating of a website is based on the number of toxic backlinks leading to the analyzed website and the importance of the detected toxicity markers. This applies only to the root domain and subdomains. In the analysis, n/a means that the number of reference domains is too low.

Other 9 additional analyses relate to analyses such as New and lost referring domains in the observed time unit, Relationship between referring domains by Authority Score - Backlink Quality, New and lost backlinks in the observed time unit, Top anchors, Analysis of backlinks by type, Analysis of link attributes and backlink top countries.

The analyses within this module show the total backlink analytics qualitative score explored by the $M_2TP (h_i)$ module and are rated as follows:

$$M_2TP_{h_r} = \sum_{i=4}^{10} x_i \ f_j \qquad (2)$$

based on the sum of individual qualitative scores/point values of the observed variables $X_i f_j$ within this module.

Module 3: M_3 - Website audit overview

This module contains a total of 159 analytical fields distributed among 11 main analytical procedures. The M_3 module is used to examine the quality of the following data/parameters:

(1) website health level: the score of the quality (health) of the website is based on the number of errors and warnings found on the website and their uniqueness. The higher the score, the fewer problems the website has, the better it is optimized for search engines, and the easier it is to use;

(2) crawled pages: this analysis shows the total number of pages indexed by SiteAuditBot and analyzes the distribution of pages according to their status;

(3) Robots.txt file status: the robots.txt file is used to tell search engines which content on a web page to index;

(4) crawlability: analysis of website navigation, search and indexing;

(5) HTTPS: HTTP status codes refer to requests made to a web server by search engines or website visitors to a web server. Many pages on a website that return 4xx or 5xx status codes can have a negative impact on the website's usability and crawlability, which can lead to a decrease in visitors;

(6) international SEO level: international SEO helps optimize a website for search engines to better understand which countries and languages which part of the website is intended for;

(7) core web vitals metrics: this metric consists of three key measurements to evaluate a website's loadability, interactivity, and visual stability, i.e., Largest Contentful Paint (LCP) - the time it takes for a browser to load the largest block of content on your homepage, Total Blocking Time (TBT) - indicates the amount of time your homepage is unavailable for input, and Cumulative Layout Shift (CLS) - detects content shifts on the homepage by combining the shifts of all elements as they load. Percentage of good pages - a page is considered good if all three metrics (LCP, TBT, CLS) are above or equal to the recommended values. The result does not include n/a pages;

(8) website performance score level: this metric deals with the loading speed of the website;

(9) internal linking score: a metric that analyzes the internal links of a website;

(10) top issues: the top issues discovered on the website based on their number and priority level;

(11) markup score: the markup score is based on the "Invalid Structured Data Elements" check and uses the Site Health algorithm. Using markup data in website analytics is a powerful way to increase a hotel's online visibility.

The analyses within the M_3 module show the total qualitative score of the website audit analysis M_3TP (h_i) explored by this module and it is calculated according to the following formula:

$$TP_{(h_r)} = \sum_{j=1}^{3} M_j \ TP \tag{3}$$

based on the sum of the individual qualitative scores of the observed variables $x_i f_j$ within this module.

The overall qualitative hotel website score (h_r) of the observed hotel sample (N) is the sum of qualitative scores of individual modules and it is described as follows:

$$M_3TP_{h_r} = \sum_{i=11}^{26} x_i \ f_j \tag{4}$$

Module 4: M_V - Valorization module for final analysis

The valorization module provides a segmented view of the main variables monitored in the M_1, M_2 and M_3 modules. Through these modules, a total of 26 summary variables are extracted and evaluated, of which M_1 - (3 cumulative variables), as part of M_2 - (7 cumulative variables), and within M_3 - (16 cumulative variables). M_V module consists of two main parts: (1) the part/column that represents the extracted cumulative data most suitable for valorization, and (2) the part/column that represents the valorization of the extracted data. The valorization part is based on an independently created algorithm. The algorithm of the valorization module is based on (a) determining the range of determined absolute values of the variables, (b) defining classes within the determined ranges of absolute values of

the variables, and (c) defining the point values of the created classes. Since the ranges (from min-value to max-value) of the absolute values of the variables are very different for the evaluation of the variables, the model of assigning a point value to the variables was used in terms of assigning the absolute value of the variable to a certain class within the observed ranges of the absolute values. Table 1 shows the minimum and maximum absolute values of each variable and the distribution of points according to the observed classes.

Table 1 Range classes & class values (variables from x_1 to x_{26})

X_i	X_i min	X_i max	Range classes & class values *in points* (y_j)												
			y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}
M_1 Domain Analytics															
X_1 Authority Score:	0	62	0	1	2	3	4	5	6	7	8	9	10		
X_2 Organic Search Traffic:	45	12,700	0	1	2	3	4	5							
X_3 Backlinks:	0	4,400	0	1	2	3	4	5							
M_2 Backlink Analytics															
X_4 Authority Score:	19	82	0	1	2	3	4	5	6	7	8	9	10		
X_5 Referring Domains:	6	550	0	1	2	3	4	5	6	7	8	9	10	11	12
X_6 Backlinks:	28	22,900	0	1	2	3	4	5	6						
X_7 Monthly Visits:	0	611,000	0	1	2	3	4	5							
X_8 Keywords:	0	135,000	0	1	2	3	4	5	6						
X_9 Outbound Domains:	0	36	0	1	2	3	4	5							
X_{10} Toxicity Score:	L(Low)	H(High)	0	1	3										
M_3 Website Audit Analytics															
X_{11} Website Health Level	61%	0.92	0	1	2	4	5	6	7	8	9	10			
X_{12} Crawled pages	6	100	0	1	2	3									
X_{13} Healthy	0	7	0	1	2	3	5								
X_{14} Broken	0	17	5	3	2	1	0								
X_{15} Have issues	3	96	5	3	2	1	0								
X_{16} Redirects	0	53	55	3	2	1	0								
X_{17} Blocked	0	23	5	3	2	1	0								
X_{18} Robot.txt File Status	0=(NA)	3=A	0	3											
X_{19} Crawlability	0%	1	0	1	2	3	4	5	6	7	8	9	10		
X_{20} Non-indexable pages	1	100	10	9	8	7	6	5	4	3	2	1	0		
X_{21} Indexable pages	0	97	0	1	2	3	4	5	6	7	8	9	10		
X_{22} HTTPS implementation	0%	100%	0	1	2	3	4	5	6	7	8	9	10		
X_{23} International SEO level	0%	100%	0	1	2	3									
X_{24} Core Web Vitals Metrics	0%	100%	0	1	2	3	4								
X_{25} Website Performance	73%	100%	0	1	2	3	4	5	6	7	8	9	10		
X_{26} Internal Linking Score	84%	100%	0	1	2	3	4	5	6	7	8	9	10		

Note: (NA – Not Available), A - Available
Source: Authors

With regard to the position (X_8) of the absolute data field of the Authority Score variable within the analytical model for secondary analysis, the algorithm for assigning the score/point value of the Authority Score variable (X_1) within the M_1 module is presented below:

$$X_{1\,Pts\,=} \quad \begin{aligned}&\text{IF}(X_8=0;"0";\text{IF}(X_8<=10;"1";\text{IF}(X_8<=\\&20;"2";\text{IF}(X_8<=30;"3";\text{ IF}(X_8<=40;"\\&4";\text{IF}(X_8<=50;"5";\text{IF}(X_8<=60;"6";\text{IF}\\&(X_8<=70;"7";\text{ IF}(X_8<=80;"8";\text{IF}(X_8<=9\\&0;"9";\text{IF}(X_8<=100;"10"))))))))))\end{aligned} \quad (5)$$

From the presented *if* algorithm, it appears that 11 classes are defined for the assignment of the Authority Score variable. For each class, a score value ranging between 0 and 10 points is assigned.

In the next step of the research, the absolute values of the observed variables x_i are assigned point values. In this way, the monitoring of the variable values and a uniform addition of the values are ensured. At the same time, the analysis of the dominant keywords and search queries revealed that in almost 100% of the cases the word hotel is among the top 3 keywords. In the top 5 organic search terms, the word hotel is combined with the name of the city or the name of the hotel in more than 90% of the cases through various phrases. It can be concluded that potential guests mainly use these words or the phrase "hotel - hotel name - city". Thus, (1) considering the search queries from the perspective of potential buyers (from the country and abroad), (2) taking into account the results and findings obtained in the course of previous research, and (3) understanding what a 5-star hotel represents in the context of supply, the 3 most relevant search engine queries are derived based on demand. The fourth search query was to research the online visibility of hotels using the phrase "hotel - hotel name - Croatia". The rest of the research is based on the following search queries: search query 1: five-star hotel Croatia, search query 2: hotel in Croatia, search query 3: luxury hotel in Croatia, and search query 4: syntax: hotel - hotel name - Croatia. Based on the above search queries (based on the most relevant common important terms), the online visibility of each hotel was studied in terms of 4 search queries. The achieved positions (online visibility) based on the defined search queries were evaluated as shown in Table 2.

Table 2 Evaluation of the online visibility of the hotel websites of the observed sample

SERP POSITION	Score / points
into **1st** page of organic (algorithm) results	5
into **2nd** & **3rd** page of organic (algorithm) results	3
into **4th** & **5th** page of organic (algorithm) results	1
other	0

Source: Authors

4. Research results and discussion

Table 3 below provides the summarized results of the values of variables of the analyzed modules, as well as the points awarded based on the achieved position within the SERP for the observed search queries. Measures of dispersion of the variables are also provided, showing dispersion of the numerical characteristics of the variables from the arithmetic mean. Dispersion of the analyzed numerical characteristics is presented using absolute measures of dispersion, namely, range of variation, variance and standard deviation.

Table 3 Point values of the observed analytical modules

Hotel (hr)		Analytical modules				
		M1TP (hr)	M2TP (hr)	M3TP (hr)	TP (hr)	SERPp
H1		11	24	78	113	11
H2		13	35	55	103	9
H3		5	15	80	100	5
H4		7	34	85	126	5
H5		9	25	77	111	5
H6		10	24	88	122	6
H7		8	21	79	108	5
H8		6	24	78	108	5
H9		11	28	91	130	5
H10		12	28	78	118	5
H11		12	28	90	130	6
H12		7	25	73	105	5
H13		6	14	73	93	5
H14		8	24	74	106	5
H15		4	38	68	110	5
H16		9	16	88	113	5
H17		11	26	81	118	5
H18		11	27	86	124	5
H19		12	30	91	113	18
Dispersion measurement	Min	4.0	14.0	55.0	93.0	-
	Max	13.0	38.0	91.0	130.0	-
	$\overline{xi}$	9.1	25.6	79.6	114.3	-
	Range of variation	9.0	24.0	36.0	40.0	-
	St.dev. σ	2.7	6.3	9.0	11.1	-
	Variance σ2	7.2	40.1	80.8	123.0	-

Source: Authors

Using Spearman's rank correlation coefficient, the relationship and influence of each module and its variables (M_1 Domain Analytics - M_1TP (h_r), M_2 Backlink Analytics - M_2TP (h_r), M_3 Website Audit Analytics - M_3TP (h_r)) on the overall qualitative value of the website was examined. The relationship between the overall qualitative value of the website and online visibility was also examined by analyzing the position within the SERP for specific search queries. Spearman's rank correlation coefficients between the above variables were as follows: (1) rs_1 [M_1TP (h_r) : TP (h_r)] = 0.50131579, (2) rs_2 [M_2TP (h_r) : TP (h_r)] = 0.44780702, (3) rs_3 [M_3TP (h_r) : TP (h_r)] = 0.74429825, and (4) rs_4 [TP (h_r) : SERPp] = 0.31491228. Spearman's coefficients were obtained by observing a sample of hotels (N = 19), and therefore these values are compared with the critical value of the rank correlation coefficient (= 0.391) based on the significance level of 5% (0.05). For the interpretation of Spearman's coefficients, 2 additional hypotheses were made, i.e., (1) independence hypothesis $h_0 : r_s = 0$, and (2) alternative hypothesis $h_1 : r_s \neq 0$. If Spearman's rank correlation coefficient (r_s) is greater than the nominal table value (α =

0.391), independence hypothesis h_0 is rejected and alternative hypothesis h_1 is accepted, implying that there is dependence between the observed modules. Otherwise, hypothesis h_1 is rejected, which assumes that a relationship exists, and independence hypothesis h_0 is accepted.

(1) $rs_1 [M_1 TP (h_r) : TP (h_r)] = 0.50131579 > 0.391$

(2) $rs_2 [M_2 TP (h_r) : TP (h_r)] = 0.44780702 > 0.391$

(3) $rs_3 [M_3 TP (h_r) : TP (h_r)] = 0.74429825 > 0.391$

(4) $rs_4 [TP (h_r) : SERPp] = 0.31491228 < 0.391$

Since $rs1 > \alpha$ and $rs2 > \alpha$, and the value of the variables ranges between ± 0.40 and ± 0.70, indicating a truly significant relationship, independence hypothesis h_0 is rejected for M_1 and M_2 and the alternative hypothesis h_1 is accepted. A truly significant correlation of the influence of quality is shown for the M_1 Domain Analytics module and the M_2 Backlink Analytics module on the overall quality of the website TP (hr).

Moreover, $rs3 > \alpha$, and the value of the variable ranges between ± 0.70 and ± 1.00. These data show a very high correlation between the quality of the M_3 Website Audit Analytics module and the overall quality of the website. In this case, independence hypothesis h_0 is also rejected and alternative hypothesis h_1 is accepted.

Since $rs_4 < \alpha$, hypothesis h_1, which assumes that a relationship exists, is rejected, and independence hypothesis h_0 is accepted. Thus, the analysis of the influence of the overall quality of the website on the position of the website on the SERP shows that there is no strong direct correlation between the points obtained based on the qualitative analysis of the websites and the points awarded based on the position obtained within the SERP for the specified search queries. It can be noted that the analyzed hotels did not pay enough attention to general search queries, but the focus of the keywords and search phrases was exclusively on the name of the hotel. Indeed, search queries in which the name of the hotel was not entered in over 90% of the cases did not result in the hotel's website being listed on the first 5 pages of the search engine results. This fact indicates a relatively weak online visibility of the observed hotels in algorithmic/organic search queries. In this way, hotels are missing an important opportunity for a direct positive impact on website visits and positive action to improve business results. We should also mention the limitations of this research

due to the sample size and a geographic definition. The research results show that there is a problem and a relevant space for the extension of the existing research, and at the same time, open the possibility of conducting similar research. In addition to the scientific aspect, knowledge gained based on the research results can be applied in modern business practice.

5. Conclusion

Nowadays, the quality of the website is a very important factor for a successful hotel business. The intention of every hotel is to sell as many products and services as possible through direct sales channels. Visibility in search engines is very important for selling through the hotel brand website. Visibility in search engines in the context of organic/algorithmic search is the result of the treatment by the Google algorithm. The Google algorithm treats a large number of variables. As highlighted in the introduction, **the aim of the research** in this article was to identify, segment and evaluate the most important SEO ranking factors to achieve better online visibility/positioning of the hotel website on the SERP, and investigate the relationships between the quality of SEO ranking factors and the position in search results based on search engine queries. In this context, the hypotheses of the article were also formulated. Starting from the first part of research (based on the study of the problem and the analysis of the latest theoretical findings), the assumptions that the position of the hotel website in the search results on the Internet search engines depends on the quality of optimization of the SEO ranking factor as a variable crucial for placement on the SERP (Search Engine Result Pages). However, at the same time, the analysis of the influence of the overall quality of the website on the position of the website on the SERP (rs4) shows that there is no strong direct relationship between the points obtained on the basis of the qualitative analysis of the websites and the points awarded on the basis of the analysis of search queries. This only partially confirms the first hypothesis of the article. This result is due to the main **limitations** of this research, namely the sample size (in this article, only 5-star hotels in the Republic of Croatia were studied) and the geographical determination (the search queries were conducted from Croatia, and Google's SERP algorithm also takes into account the location of the user making the search query). Bearing this in mind, and for more

meaningful results, **recommendations** for future research should be: (1) to increase the size of the sample, (2) to expand the sample to include the study of hotels from multiple countries, and (3) to conduct analysis of search queries based on queries from multiple locations/countries. In addition, for future research, it is recommended to increase the time period for analyzing the search queries and to compare the search queries with the demand market trends (Google Trends). By segmenting and valorizing the variables, the interrelationships of the variables were investigated. In the empirical part of the study, the variables were observed through defined analytical modules. By examining the relationships between each of the modules studied (and their associated variables), it was found that the $M_3 TP (h_r)$ module (Website Audit) had the greatest influence on the overall quality of the hotel website $TP (h_r)$. In addition, the influence of the variables of M_1 and M_2 modules on the overall quality of the website was evaluated. In this way, new and more detailed findings were obtained, providing a better basis for decision making to increase online visibility in search engines. Hypothesis 2 of this thesis is thus proven. The article introduces the concept of a more detailed analysis of variables/data (segmentation and valorization of variables) obtained through Semrush SaaS. Such an approach provides a better insight into the shortcomings in the quality of variables and a better understanding of the relationship between the variables that influence the online visibility of the hotel website. The research results indicate the qualitative deficiencies in the optimization of the variables, which at the same time shows that there is a significant potential to increase online direct sales of hotels through better detection and optimization of the most important Google factors for ranking. This confirms the third hypothesis of this article. The **scientific contribution** of the article is evident in the original approach of identifying, segmenting and evaluating the most important SEO factors with the aim of improving online visibility in search engines. By exploring the relationship between the quality of SEO ranking factors and the position in the research results, new scientific knowledge was gained and at the same time room was opened for new scientific research, especially related to the analysis of the impact of the quality of a website on the position of a website on the SERP. Apart from the above limitations, the results of the conducted research show that there are a wide range of approaches to this topic, which open up numerous possibilities for conducting this or similar research.

References

1. Dean, B. (2021). *Google's 200 ranking factors: The complete list (2022).* Backlinko. https://backlinko.com/google-ranking-factors

2. D-edge hospitality (2021). *The rise of direct bookings over OTAs - Hotel distribution trends in EMEA and APAC from 2017 to 2020.* D-edge hospitality solutions. https://www.d-edge.com/the-rise-of-direct-bookings-over-otas

3. Garcia, J. E., Lima, R. & da Fonseca, M. J. S. (2022). Search Engine Optimization (SEO) for a Company Website: A Case Study. In Rocha, A. et al. (Eds.). *Information Systems and Technologies. WorldCIST 2022. Lecture Notes in Networks and Systems*, *470* (pp. 524-531) Cham: Springer. https://doi.org/10.1007/978-3-031-04829-6_47

4. Iqbal, M., Khalid, M. N., Manzoor, A., Abid, M. M. & Shaikh, N. A. (2022). Search Engine Optimization (SEO): A Study of important key factors in achieving a better Search Engine Result Page (SERP) Position. *Sukkur IBA Journal of Computing and Mathematical Sciences - SJCMS*, *6*(1), 1-15. https://doi.org/10.30537/sjcms.v6i1.924

5. Konidaris, A., Stellatou, O., Polykalas, S. E. & Vardikou, C. (2022). Predictive SEO for Tourism Websites through Transformer Keyword Identification. In Katsoni, V. & Şerban, A. C. (Eds.). *Transcending Borders in Tourism through Innovation and Cultural Heritage. Springer Proceedings in Business and Economics* (pp. 897-912). Cham: Springer. https://doi.org/10.1007/978-3-030-92491-1_53

6. Law, R. (2019). Evaluation of hotel websites: Progress and future developments. *International Journal of Hospitality Management*, *76*(B), 2-9. https://doi.org/10.1016/j.ijhm.2018.06.005

7. Pan, B. (2015). The power of search engine ranking for tourist destinations. *Tourism Management*, *47*, 79-87. https://doi.org/10.1016/j.tourman.2014.08.015

8. Phobs CRS (2021). *Revenue Generating Solutions for Hospitality*, *Statistical data*. https://phobs.net

9. Rodlej, J. (2022). *Hotel SEO: The How to Guide.* Dialed Labs. https://dialedlabs.com/hotel-seo

10. Shaolang, S., Yunjie, W., Kwok-Leung, T. & Shouyang, W. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1-10. https://doi.org/10.1016/j.tourman.2018.07.010

11. Sengar, A. S. (2021). Search engine optimization: A review. *Asian Journal of Multidimensional Research*, *10*(12), 260-264. https://doi.org/10.5958/2278-4853.2021.01206.4

12. Vukasović, T. & Mihač, V. (2021). Trends in the Online Booking of Hotel Accommodation. *International Journal of E-Services and Mobile Applications* (IJESMA), *13*(1), 60-76. http://doi.org/10.4018/IJESMA.2021010105

13. Vyas, C. (2019). Evaluating state tourism websites using Search Engine Optimization tools. *Tourism Management*, *73*, 64-70. https://doi.org/10.1016/j.tourman.2019.01.019

14. Yalçın, N. & Köse, U. (2010). What is search engine optimization: SEO?. *Procedia - Social and Behavioral Sciences*, *9*, 487-493. https://doi.org/10.1016/j.sbspro.2010.12.185

15. Yu, J. (2018). Exploration on Web Testing of Website. *Journal of Physics: Conference Series*, *1176*(2), 022042. https://doi.org/10.1088/1742-6596/1176/2/022042