UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

DYNA
http://dyna.medellin.unal.edu.co/

Digital gap of the colombian tourism sector

Oscar Jose Ospino-Ayala [a], Alexandra Barrios-Sandoval [b] & Ramon Antonio Matos-Mareño [b]

[a] *Faculty of Basic Sciences, Department of Mathematics, Universidad del Atlántico, Barranquilla, Colombia & Faculty of Business Sciences, Economics Program, Universidad del Magdalena, Santa Marta, Colombia. ojospino@mail.uniatlantico.edu.co*

[b] *Faculty of Basic Sciences, Department of Mathematics, Universidad del Atlántico, Barranquilla, Colombia. alexandrabarriossandoval@gmail.com, ramonmatos@mail.uniatlantico.edu.co*

Received: August 30th, 2019. Received in revised form: March 17th, 2020. Accepted: April 13th, 2020.

Abstract
The purpose of this research is to know the current situation of digitization that exists in the Colombian tourism sector. For this, the data of the Annual Survey of Service-2018 were used, to which the multivariate analysis techniques of categorical main components and factorial analysis were applied, in the construction of an index that measures the digitization from the Technologies of the Information and communication. With the results of this application it was demonstrated that the techniques of multivariate analysis guarantee objective knowledge that allows the planning and implementation of strategies to face digital changes and generate innovation.

Keywords: digital divide; tourism; categorical principal components; factor analysis; composite index.

Brecha digital del sector turístico colombiano

Resumen
El propósito de la presente investigación es conocer la situación actual de la digitalización que existe en el sector turístico colombiano. Para esto, se usaron los datos de la Encuesta Anual de Servicio-2018, a los que se les aplicaron las técnicas multivariadas de análisis de componentes principales categóricos y análisis factorial, en la construcción de un índice que mide la digitalización desde las Tecnologías de la Información y Comunicación. Con los resultados de esta aplicación se evidenció que las técnicas de análisis multivariado usadas de forma adecuada garantizan conocimientos objetivos garantizan conocimientos objetivos que permiten la planificación e implementación de estrategias para afrontar los cambios digitales y generar innovación.

Palabras clave: brecha digital; turismo; componentes principales categóricos; análisis factorial; índice compuesto.

1. Introduction

The use of digital tools greatly affects the way of doing business, this is reflected in aspects such as the creation, destruction, and modification of jobs [1,2]. The digital transformation for the Colombian tourism subsector has had positive impacts; for example, in the first quarter of 2018, the profits of companies that used digital resources in the development of their activities exceeded by 6.9% the profits of companies that did not use those resources [3].

Despite the above, the progress of the digital transformation is not uniform in the set of companies in the tourism sector, because not all companies can overcome the difficulties, they encounter in the digitization process. Inequalities to access information, knowledge and education through new technologies [4], known as the gap, hinder the design and implementation of policies that guide improvements in the Colombian tourism sector.

From this objectivity, it is not possible to speak of a single digital gap, given that it is multifaceted depending on the context of comparison; thus, there exists a semantic application of digital gaps in plural [5]. For this reason, the relevance of partial ICT indicators for measuring the digital divide (gap) has been recognized [6].

Consequently, a first unavoidable task is the statistical measurement of the digital gap in order to know the situation of digitization in the Colombian tourism sector. For this, a good tool is the index evaluation, which is a synthetic

How to cite: Ospino-Ayala, O.J, Barrios-Sandoval, A. and Matos-Mareño, R.A, Digital gap of the colombian tourism sector. DYNA, 87(213), pp. 173-177, April - June, 2020.

measure that accounts for an abstract definition [7].

In this scenario, the multivariate statistical analysis that mentions the set of techniques used to examine, represent and interpret the associations between multiple variables of an individual or sample of individuals, at the same time [8,9]. It is suitable for the construction of indexes, from the application of different analysis methods.

Among the multivariate techniques most used for the elaboration of indexes or indicators, the Main Components Analysis (PCA) and the Factor Analysis (AF) are efficient, by summarizing a set of original quantitative variables in a few synthetic variables, with minimal loss of information [9,10].

The use of techniques analogous to PCA is also common when dealing with qualitative variables. For this reason, the Main Categorical Components Analysis (CATPCA) is widely used for the construction of indexes that address qualitative original variables, summarizing them in a few synthetic variables with the least possible loss of information [11,12].

Although each multivariate method can be used separately in the construction of indexes. These become more efficient when used together with other data analysis techniques. Indeed, the combined use of the aforementioned multivariate techniques for the transformation of partial indicators (original variables) into a few synthetic variables (indexes) is notable [13,14].

The measurement of the digital gap in the Colombian tourism industry that is proposed in this research is done with an approach from multivariate statistics with the results referring to Information and Communication Technology (ICT) obtained in the Annual Service Survey (EAS) -2017 [15], within the dimensions of ICT Preparation, Use and Impacts.

2. Methodology

For the construction of the composite Index of Use, Preparation and Impacts of ICT (IUPITIC) in the PIMES [16], which is a mathematical combination of sub-indexes or partial indicators [9,11], the information of the National Department of Statistics [15] was available through of the EAS. This survey measured the technological component of 5,989 Colombian companies in the service sector in general. Which were distributed according to the economic activity described by the International Standard Industrial Classification-ISIC Revision 4.

Now, in order to obtain the target population, the definition of a tourist industry made by the World Tourism Organization [17] was used, thus framing the ISIC (CIIU)-4 divisions by group, together with their number of companies, to the tourist activities of accommodation, foods and beverages, transportation and equipment rental, and travel agencies and other reservation services. In effect, a target universe of 1,678 companies was obtained. The variables of interest of the EAS- 2017 [15] were selected as stipulated by [18], thus choosing 14 qualitative variables for the Use sub-index, 5 quantitative variables in the Preparation sub-index, and 1 quantitative variable for the Impact sub-index.

The IUPITIC index compromises the use of quantitative and qualitative variables, consequently, in order to elaborate the Use sub-index, the method of Analysis of Main Categorical Components-CATPCA is used, which allows treating qualitative variables, for the construction of the Preparation and Impact sub-indexes is used in Main Components Analysis -ACP, and for the IUPITIC composite index the weighted weights method is applied. Finally, a single factor ANOVA is used in order to measure the gap.

3. Results

The measurement of the digital gap in the Colombian tourism sector was carried out in this research evaluating a composite index (IUPITIC) defined from the Use (Y_i with $\mathrm{i} = 1,2$), of Preparation n (X_i with $\mathrm{i} = 1,2$), and of Impact (C) of Information and Communication Technology.

In these calculations $Y_1 -$ corresponds to the value of the form of business, Y_2 - speed and communication of digital information, X_1- occupation of internet and computer and X_2- salary and telework.

The variables considered in the aforementioned sub-indexes ($Y_i, X_i,$ and C) are presented in the following Table 1, also the notations of the variables in this tabulation correspond to those established in the information capture instruments used by DANE (2018).

Table 1.
Variables used in the construction of subindexes and composite index.

Subindex	Variables	Kind*	Definition	Methods	IUPITIC
Uses (Applications)	B_1_6	Ord	Number of ICT goods	CAT-PCA	Weighing
	B_3	Dic	Internet use		
	B5	Dic	Website availability		
	B_6	Dic	Use of Networks for communication and information sharing		
	B_9_1	Ord	Maximum speed (bandwidth)		
	B_9	Ord	Main connection type		
	B_11	Dic	Use of Networks for communication		
	B_12_A	Dic	Sending or receiving emails		
	B_12_B	Dic	Information search		
	B_12_C	Dic	Electronic banking and other financial services		

	B_12_D	Dic	Government transactions	
	B_13	Dic	Selling services or products	
	B_13_B	Dic	Purchase of products or services	
Skill or abilities for use				
Prepa-ration	B_2	Cont	Total percentage of people who used a computer	PCA
	B_4	Cont	Total percentage of people who used the internet	
	B15PER	Cont	Percentage of people employed with Telework	
	SUSAL	Cont	Average salary of employed personnel	
Preparation				
Impact	ISERPER	Cont	Average net earnings per worker	Coding (*c*)

Source: Authors
*Ord: Ordinal; Dic: Dichotomous; Cont: Continuous

The values obtained for the subindexes and the composite index are detailed below:

3.1 Use subindex

The coefficients that appear in the following equalities (eq. 1-2) were obtained from the Analysis of Main Categorical Component (CATPCA) technique.

$$y_1 = -(0.646 \times B\_1\_6) + (0.9 \times B\_3) + (0.936 \times B\_5) + (0.178 \times B\_6) - (0.048 \times B\_9) - (0.101 \times B\_9\_1) + (0.096 \times B\_11) + (0.987 \times B\_12\_A) + (0.987 \times B\_12\_B) + (0.987 \times B\_12\_C) + (0.987 \times \mathrm{B\_12\_}D) + (0.918 \times B\_13) + (0.918 \times B\_13\_B) \quad (1)$$

$$y_2 = -(0.216 \times B\_1\_6) + (0.021 \times B\_3) - (0.042 \times B\_5) + (0.559 \times B\_6) + (0.518 \times B\_9) - (0.669 \times B\_9\_1) + (0.566 \times B\_11) - (0.048 \times \mathrm{B\_12\_A}) - (0.048 \times B\_12\_B) - (0.049 \times B\_12\_C) - (0.050 \times B\_12\_D) - (0.067 \times B\_13) - (0.067 \times B\_13\_B) \quad (2)$$

3.2. Preparation subindex

The partial indicator was incorporated into a linear model that relates its coefficients from the multivariate of Principal Components Analysis (PCA) technique in the way expressed in eq. (3)-(4):

$$x_1 = (0.961 \times B_2) + (0.961 \times B\_4) + (0.18 \times B15PER) + (0.164 \times SULSAL) \quad (3)$$

$$x_2 = (-0.016 \times B\_2) - (0.015 \times B\_4) - (0.655 \times B15PER) + (0.764 \times SULSAL) \quad (4)$$

3.3. Impact subindex

The subindex was transformed from the following methodology (eq. 5). Where μ is the mean of the *Iserper* variable and σ is its deviation.

$$c = \frac{(iserper - \mu)}{\sigma} \quad (5)$$

For each of the 1,678 companies, the sub-index values were calculated and with uniform weightings for each sub-index, the *IUPITIC* composite index was obtained, which yielded the following value -1.5x10-5 determined with eq. (6):

$$IUPITIC = (0.20 \times Y_1) + (0.20 \times Y_2) + (0.20 \times X_1) + (0.20 \times X_2) + (0.20 \times ISERPER) \quad (6)$$

3.4. Impact subindex

The results of the composite-*IUPITIC* index were classified by company size based on the definition made by Law No. 905 (2004) of the Colombian Legislative Branch on the size of companies (1-micro, 2-small, 3-medium, and 4-large) by number of employed workers. Obtaining the bar chart on the means of the results of the *IUPITIC* index by type of company (Fig. 1). From this primary graphic analysis, there are strong suspicions that i) the composite index is affected according to the size of the organization, ii) in micro (1) and large (4) companies, the highest *IUPITIC* values are found, and iii) small and medium-sized companies have the lowest index.

Carrying out a more objective study, an Analysis of Variance-ANOVA is applied in order to test the hypothesis of equality of means. It was concluded at a confidence level of 95% that there are significant differences in the Use, Preparation and Impact of ICTs between micro, small, medium-sized and large companies in the evaluated sector, with a $F_{exp} = (46.413) > F_{Teo} = (2.61)$, , as presented in the ANOVA of Table 2.

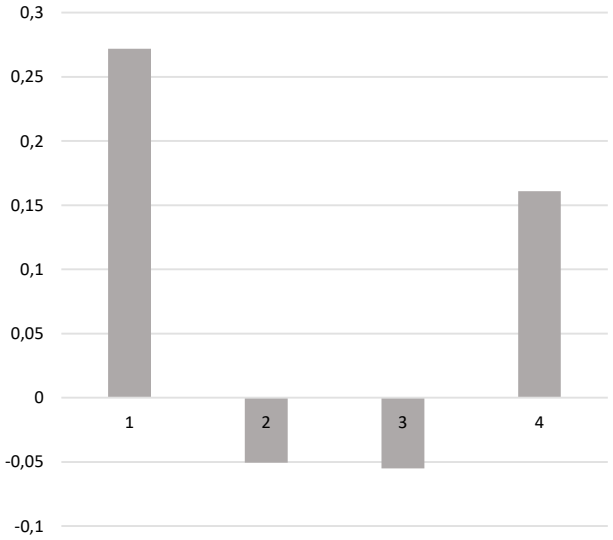

Figure 1. Bar chart of the IUPITIC index by company size.
Source: The Authors.

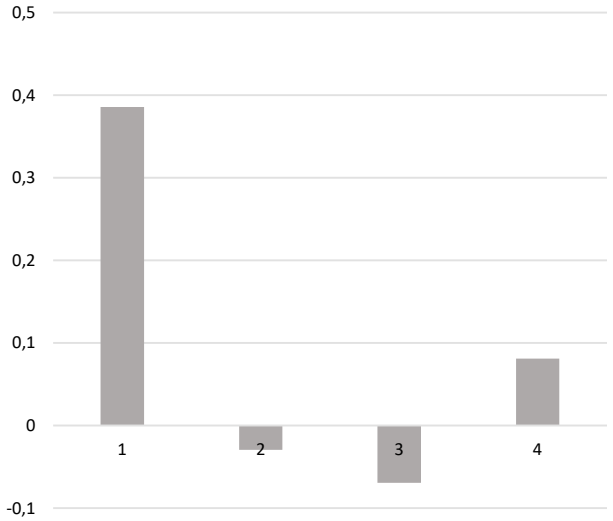

Figure 2. Bar chart of the IUPITIC index by company size for robustness.
Source: The Authors.

This evidences the existence of a large digital gap in the organizational sizes of the Colombian tourism sector.

In order to evaluate the robustness of the general index, the weights of Eq. (6) are modified to the form of Eq. (7) as suggested by Alderete (2011).

$$IUPITIC = (0{,}5 \times Y_1) + (0{,}4 \times Y_2) + (0{,}3 \times X_1) + (0{,}2 \times X_2) + (0{,}1 \times ISERPER) \quad (7)$$

This delivered an *IUPITIC* Composite Index of de -6,412x10^{-6}, slightly smaller than the previous one. The Analysis of Variance of a single Factor - ANOVA (Table 3), continues contrasting to a confidence level of 95% that the means by (according to) company size are different. Which is evidence of the good robustness of the general index, which maintain the same descriptive suspicions that were mentioned (Fig. 2).

Table 2.
Single Factor ANOVA Summary

Source	Sum of Squares	Degrees of Freedom	Squares Average	F_{exp}	Proba bility	F_{Teo}
Between Groups	18.787	3	6.26	46.413	0.000	2.61
Error	225.869	1674	0.135			
Total	244.657	1677				

Source: The Authors.

Table 3
ANOVA summary in order to verify the robustness of the composite index

Source	Sum of Squares	Degrees of Freedom	Squares Average	F_{exp}	Probabi lity	F_{Teo}
Between Groups	23.96	3	7.986	20.47	0.000	2.61
Error	653.081	1674	0.390			
Total	677.039	1677				

Source: The Authors.

4. Conclusions

The execution of composite indexes in order to measure the digital gap from the Use, Preparation and Impact of ICTs in companies, is efficient for the construction of public and business policies for the purposes of competitive sector improvement.

The low level of the *IUPITIC* index shows some important differences in the means by organizational size. Micro and large companies tend to have better Use, Impact and Preparation of ICTs in their activities, while medium sized and small companies are those that least use ICTs in their functions.

The Single Factor Variance Analysis is a good objective tool in order to infer inequalities in the scores of the composite index - *IUPITIC*. Determining the existence of a digital gap from the combined values of the Use, Preparation and Impact of ICT by company size in the Colombian tourism sector.

The evaluation of the robustness of the elaborated composite index is appropriate in order to know the efficiency of the model. In this context, the synthetic variable - *IUPITIC* shows a correct adaptation for the measurement of the digital gap in the tourist sector that was studied.

The use of multivariate techniques in order to measure the digital gap is convenient, especially since there is no universal consensus on how digital gaps can be mediated. Therefore, it is necessary to advance in the construction of composite indexes that enable the evaluating of the Use, Preparation and Impacts of the ICTs of the different companies for the productive strengthening.

References

[1] OECD - Organization for Economic Co-operation and Development. Perspectives de l'économie numérique (numerical economics) of the OECD 2015, 1st ed., Paris, France- OECD, 2015, pp. 30-60. DOI:

10.1787/9789264259256-es
[2] Massachusetts Institute of Technology. MIT. Sloan Management Review [online]. Massachusetts, United States, 2017. [Consulted February 15th, 2019]. Available at: https://sloanreview.mit.edu/article/putting-customer-experience-at-thecenter-of-digital-transformation/
[3] MCIT-Ministry of Commerce, Industry and Tourism. Tourism Monthly Bulletin, March 2018, Bogotá, Colombia, Mincit, 2018, 10 P.
[4] Cabero-Almenara, J. and Ruiz-Palmero, J., Technologies of Information and Communication for inclusion: reformulating the Digital Gap, [Online]. 0(9), 2017. [consultation date April 12th, 2020]. Available at: https://www.upo.es/revistas/index.php/IJERI/article/view/2665.
[5] Camacho, K., Words at stake: multicultural approaches to information societies. [online], Vasco, Spain, C & F Editions, 2005. [consultation date January 20th, 2019], Chapter. 8, The digital divide (GAP). Available at: https://vecam.org/archives/article550.html
[6] Betancourt, V., The World Summit on the Information Society (WSIS) process and debated issues. Tunisia, Association for the Progress of Communications (APC), [online]. 2004. [consultation date January 20th, 2019]. Available at: https://www.apc.org/en/pubs/world-summit-information-society-wsis-process-and
[7] Pasantez, T.E., Construction of an index related to the principle of relevance for national universities using data of the SNIESE 2010. Project and the methodology of the Analysis of Main Nonlinear Components, Thesis Undergraduate, Escuela Politécnica Nacional, Quito, Ecuador, 2013, 28 P.
[8] Álvarez, R., Statistics applied to health sciences, 1st ed., Spain, Ediciones Díaz de Santos Publishing House, 2007, pp. 56-89.
[9] Peña, D., Analysis of multivariate data, Alianza Editorial, Spain, 2002, pp. 137-389.
[10] Cuadras, C., New methods of multivariate analysis, CMC Editions, Barcelona-Spain, 2014, pp. 77-119.
[11] Pérez, C., Multivariate Data Analysis Techniques, Pearson Educación, S.A., Madrid-Spain, 2004, pp. 275-301.
[12] Montanero, J., Multivariate analysis, University of Extremadura, Spain, 2008, pp. 175-193.
[13] CEPAL. The equity gap: a second evaluation, CEPAL, Santiago de Chile, 2000.
[14] CEPAL, Social Panorama of Latin America, 2000-2001. CEPAL, Santiago de Chile, 2001.
[15] DANE-National Administrative Department of Statistics. Annual Survey of Services - EAS - 2017, DANE1-10, Colombia, 2018.
[16] Larrea, C., Health, nutrition and social inequality in eight Latin American countries: a comparative study from the DHS Surveys, Colombia, Pan American Health Organization, 2001, pp. 1-34. DOI: 10.1016/j.gaceta.2012.12.007
[17] SITUR- Magdalena Tourist Information System. Tourist Information System: a reality for Magdalena, SITUR, 1st Ed., Santa Marta Tourist, Cultural and Historical District - Colombia, [online]. 2016. [consultation date January 10th, 2019]. Available at: https://www.siturmagdalena.com/promocionPublicacion/listado
[18] Alderete, M.V., Measurement of Information and Communication Technologies in Colombian Service Companies. Administration Notebooks, 25(45), pp. 39-62, 2012. DOI: 10.11144/Javeriana.cao25-45.mtic

O.J. Ospino-Ayala, received the degree of Sp. in Applied Statistics in 2019 from the Universidad del Atlántico, Colombia. He is BSc Eng. in Industrial Engineer in 2016 at the Universidad del Magdalena, Colombia. He worked as a research engineer in the period 08/2017-12/2017 at the Universidad del Magdalena, Colombia. He worked as a young researcher and innovator for peace - COLCIENCIAS (01/2018 - 01/2019), Universidad del Magdalena. And he is currently a young researcher and innovator - COLCIENCIAS since 2019 at the Universidad del Magdalena.
ORCID: 0000-0002-3876-6111

A. Barrios-Sandoval, received the degree of Sp. in Applied Statistics in 2019 from the Universidad del Atlántico, Colombia. She is BSc. Eng. Industrial Engineer in 2015 at the Universidad del Magdalena, Colombia. In the period 01/2018 - 06/2018 she worked as a mathematics teacher at the Northwestern Educational Institution of Soledad. Currently, she works as a mathematics teacher since 06/2018 at the San Antonio de Padua Industrial Technical Educational Institution of Soledad.
ORCID: 0000-0002-3629-5029

R.A. Matos-Mareño, received in 1996 the Postdoctorate / Postdoctoral stay at the Universidad de la República, Uruguay. He completed a postdoctoral stay at the Inter-American Center for the Teaching of Statistics in 1992. He obtained the degrees of PhD in Mathematics in 1987 and of Mathematician in 1982, both at the Moscow State University M.V. Lomonosov. With work experience at the Simón Bolívar University (January 1997 to July 1998), Faculty of Basic Sciences - UAN (March 1996 to November 2003), Universidad Libre de Colombia (August 1994 to March 2016), Fundación Universidad del Norte (January 1989 to July 1992), Universidad del Atlántico full time professor (February 1988 to March 2016).
ORCID: 0000-0003-1096-034X