Statistical Brief

DATA FOR NOW

DANE
INFORMACIÓN PARA TODOS

Analysis of accessibility to educational centers

El futuro es de todos

Gobierno de Colombia

SUSTAINABLE DEVELOPMENT GOALS

4 QUALITY EDUCATION

Analysis of accessibility to educational centers

Statistical Brief

NATIONAL ADMINISTRATIVE *DEPARTMENT OF STATISTICS - DANE*

Juan Daniel Oviedo Arango
Director

María Fernanda de la Ossa Archila
Secretary General

Coordination of the series of statistical briefs

Karen Chavez
Advisor - General Directorate

Elaboration

Mateo Cardona Quintero
Andrés Felipe García-Suaza
Sergio Alberto Montoya Cruz
Juan Sebastián Ordoñez Herrera
Advisors - General Directorate

Julieth Solano (Technical Director)
Directorate of Regulation, Planning, Standardization and Normalization

Luisa Natalia Monroy
Natalia Alonso Ospina
Karen Chavez (Coordinator)
SDG Indicators Working Group
General Directorate

Information processing

Mónica Lorena Ortiz Medina (Coordinator)
GIT Statistical Directory
Geostatistics Directorate

Héctor Ricardo Patiño Rivera
Carlos Alberto Durán Gil (Coordinator)
Innovation Working Group
Geostatistics Directorate

Sandra Moreno (Technical Director)
Geostatistics Directorate

Lelio Alejandro Arias (coordinator)
Population Base Register Working Group
Directorate of Censuses and Demography

Cristhyan Leonardo Naranjo
Administrative Records Team
General Directorate

Translation

Angélica Ordoñez
Local Project Coordinator D4N Strategy - Colombia (United Nations)

Design and layout

Santiago Arévalo Gómez
María Andrea Castro Gutiérrez
Brandon Steve Rojas Guerra (Coordinator)
Communication and Design Working Group
Directorate of Dissemination and Statistical Culture

Cover

Francisco Javier Viveros Góngora (*Vigo*)

Foreword

DANE's Statistical Briefs Series encourages analysis and decision-making based on available data about public interest topics. **The contribution of statistical briefs is to gather information from different statistical sources to characterize an issue in a single document.** In this way, the specialized and interested public may have a perspective from a diversity of information sources. The panorama provided in the briefs does not always pretend to be exhaustive, and the analysis of the data presented may be extended according to the reader's interests.

DANE's Statistical Briefs include measurements that belong to the regular statistical production, as well as measurements that belong the 'Experimental Statistics' line. In this context, the Statistical Briefs Series emphasis the differentail and intersectional approach **in the production of data in order to "leave no one behind", in line with the 2030 Agenda for Sustainable Development,** and thus promote analyses that contribute to make visible life situations, particularities, gaps and inequalities between different population groups.

Accordingly, DANE's Statistical Briefs represent an innovative vision of statistical production and dissemination, focusing on the use of data beyond purely statistical purposes. **In addition, the briefs increase the supply of measurements and promote their use in the generation of evidence-based knowledge to enrich dialogues, decision-making, the design of public policies and the monitoring of the country's progress in terms of Sustainable Development.**

On this occasion, DANE presents an analysis of the education sector that addresses coverage measures, regional (departmental) differences, and possible gaps by geographic disaggregation (urban and rural), and introduces a first measurement of the distance from students' homes to educational centers. This note, called Analysis of accessibility to educational centers, aims to present the first approach to education accessibility metrics based on an analysis of distances and their possible relationship with school dropout.

Contents

List of graphics

List of illustrations

List of Tables

List of abbreviations, Spanish acronyms

CNPV 2018: National Population and Housing Census. 2018

DANE: National Administrative Department of Statistics

DIREDU: Statistical Directory of Educational Centers

DUE: Single Directory of Educational Centers

ECTeI: Education, Science, Technology, and Innovation

GIS: Geospatial Information Systems

Km: Kilometers

m: Meters

MEN: Ministry of National Education

NDP: National Development Plan

OBS: Observations

OECD: Organization for Economic Cooperation and Development

OSM: Open Street Map

p.p.: Percentage points

SDG: Sustainable Development Goals

SIECTI: Science, Technology, and Innovation Information System

SIEF: National Educational Information System

SIMAT: ntegrated Enrollment System

SISE: Education Centers Information System

TRENDS: Thematic Research Network on Data and Statistics

Introduction

Education is fundamental for the development of any society since it is a relevant factor for economic growth and one of the main determinants for upward socioeconomic mobility. Thus, education has a direct effect on closing gaps and fighting poverty (Ejaz et al 2019; Hanushek & Woesman, 2007; Ndou & Tsaurai, 2019). Having timely and relevant statistical information that allows both the generation of knowledge and the design, implementation, monitoring, and evaluation of evidence-based public policies is fundamental for strengthening education systems globally.

In Colombia, the production of statistical information related to education issues has been carried out in a decentralized manner by multiple institutional actors that are part of the statistical roundtable on Education, Science, Technology and Innovation (ECTeI, for its acronyms in Spanish). A commitment of the National Statistical System (SEN, for its acronyms in Spanish) seeks to provide timely information to strengthen public policy in this sector. Since 2020, the members of this roundtable have participated in the process of consolidating the Science, Technology, and Innovation Information System (SIECTI, for its acronyms in Spanish), which is composed of a set of prioritized strategic indicators that will be interoperated and available to the specialized and non-specialized public, responding to the requirements of the 2030 Agenda for Sustainable Development, the recommendations of the Organization for Economic Cooperation and Development (OECD) and the objectives of the National Development Plan (PND, for its acronyms in Spanish) 2018-2022.

The SIECTI gathers indicators built through various sources of information, such as statistical operations and administrative records. However, there is still a need for increasingly accurate, timely, and territorially focused information to guide the structure and implementation of programs and policies aiming to guarantee inclusive, equitable, and quality education and promote lifelong learning opportunities for all.

For this purpose, a pilot project was structured within the framework of the Data4Now[1] initiative for the calculation of new indicators based on the analysis of distances between the homes of children and adolescents enrolled in official and non-official educational centers in Colombia and their places of learning. This approach recognizes the challenges in terms of geolocation presented by statistical operations and administrative records related to the sector, and seeks to enhance the capacity to generate information with a high public value from different types of sources through standardized integration processes.

This exercise was based on integration of sources such as the 2018 National Population and Housing Census (CNPV 2018, for its acronyms in Spanish), DANE's Statistical Directory of Educational Venues (DIREDU, for its acronyms in Spanish) and the administrative record of the Integrated Enrollment System (SIMAT, for its acronyms in Spanish), in order to analyze possible regional differences in terms of accessibility and educational dropout. Using these sources presents many challenges in coverage, quality, variables standardization,

[1] This initiative is a joint effort between the Global Partnership for Sustainable Development Data, the United Nations Statistics Division, the World Bank, the Thematic Research Network on Data and Statistics (TReNDS) of the Sustainable Development Solutions Network, governments, and partners of the United Nations System.

and georeferencing capacity. This last point is particularly relevant, as the analysis and visualization of geospatial data in the SDG indicators improve the ability of policymakers and the general public to understand and respond to local circumstances and needs across geographic space over time.

On the other hand, the distance between home and educational centers in the literature has been recognized as a relevant factor for primary education processes in children and adolescents. Additionally, this factor is related to school absenteeism and student dropout at primary, secondary and middle school levels. Authors such as Boterman et al. (2019), Frankenberg (2013), and Hamnett & Butler (2013), analyze the relationships between residential segregation patterns and segregation in the educational sphere. Shehu (2018) shows, among other results, that children located in rural areas and those living farther away from the nearest elementary school are more likely to not attend or drop out.

In the same way, Chica-Olmo et al. (2018); Chillón et al. (2017); Fjellman et al. (2019) analyze the effects on children of home-to-schools distance, spatial dependence and the effects on the choice of means of transport. Alvarez-Pedrerol et al. (2017) analyze the effect of exposure to air pollutants in home-to-school travel on the cognitive development of children and adolescents. Finally, authors such as Mantovani et al. (2021); Andersson et al. (2012) analyze how the distance between residence and educational centers affects students' educational process and experience, making a particular emphasis on vulnerable populations and the effects on inequality.

Regarding the above, this Statistical brief presents the results of the home-to-school distance and school dropout analysis project. For this purpose, the document is organized as follows: after this introduction, a compendium of indicators is presented to show the progress and challenges that the country is currently experiencing, in terms of the coverage of the educational system in primary, elementary and middle school, as an approximation to the functioning and structure of the Colombian educational system that is deepened in other reports[2]. Section 2, in its first part, describes the data, the integration taxonomy and the methodologies used for the calculation of distances, while the second part presents the analysis of the results of the project on distance analysis and accessibility to education. The third section presents the conclusions, followed by a section summarizing the main concepts used in the construction of the analysis, and a fifth and final section containing the annexes.

[2] In this regard, see, for example, the first statistical report on Education, Science, Technology - ECTel, prepared in 2021 within the framework of the statistical roundtable on the subject. Available for consultation at: https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/poblacion-escolarizada/educacion-formal

1

Education coverage rates at preschool, elementary and middle school levels

The fourth of the seventeen Sustainable Development Goals (SDGs) is to "ensure inclusive and equitable, quality education and promote lifelong learning opportunities for all". This is part of the 2030 Agenda adopted in 2015 by the United Nations Member States as a commitment to the dignity and equality of people, shared economic prosperity and environmental protection[3].

Education is a fundamental right that plays an important role as a driver of development. Its provision as a public service constitutes one of the primary goals in a social state based on the rule of law such as Colombia (Cortés 2012). One of the most widely used indicators to measure the social demand for education are coverage rates, defined as the percentage ratio between the number of students enrolled with concerning the population theoretically old enough to attend a given level.

At the secondary education level[4], the national gross coverage rate has had a sustained increase in recent years, going from 58.4% in 2000 to rates above 80.0% since 2017 (86.1% in 2020) (Graph 1). However, this is the first step since developing countries have achieved these significant improvements. So now the challenge, in terms of educational public policy, is to reduce inequities in access with quality, especially in the most vulnerable groups (Gamboa, 2010; Rodríguez et al, 2013; OECD, 2021).

Graph 1.

SGD Indicator ODS 4.1.5.C. Gross coverage rate in secondary education. National total. 2000-2019. (Percentage)

Source: National Education Ministry (MEN, for its Spanish acronym) – Integrated Enrollment System (SIMAT, for its Spanish acronym) (2000-2020).

Note:
[1] The gross coverage rate measures the percentage ratio between the total number of students enrolled (regardless of their age) over the total population of the theoretical age to attend the educational level, which regarding secondary education corresponds to the range between 15 and 16 years old.

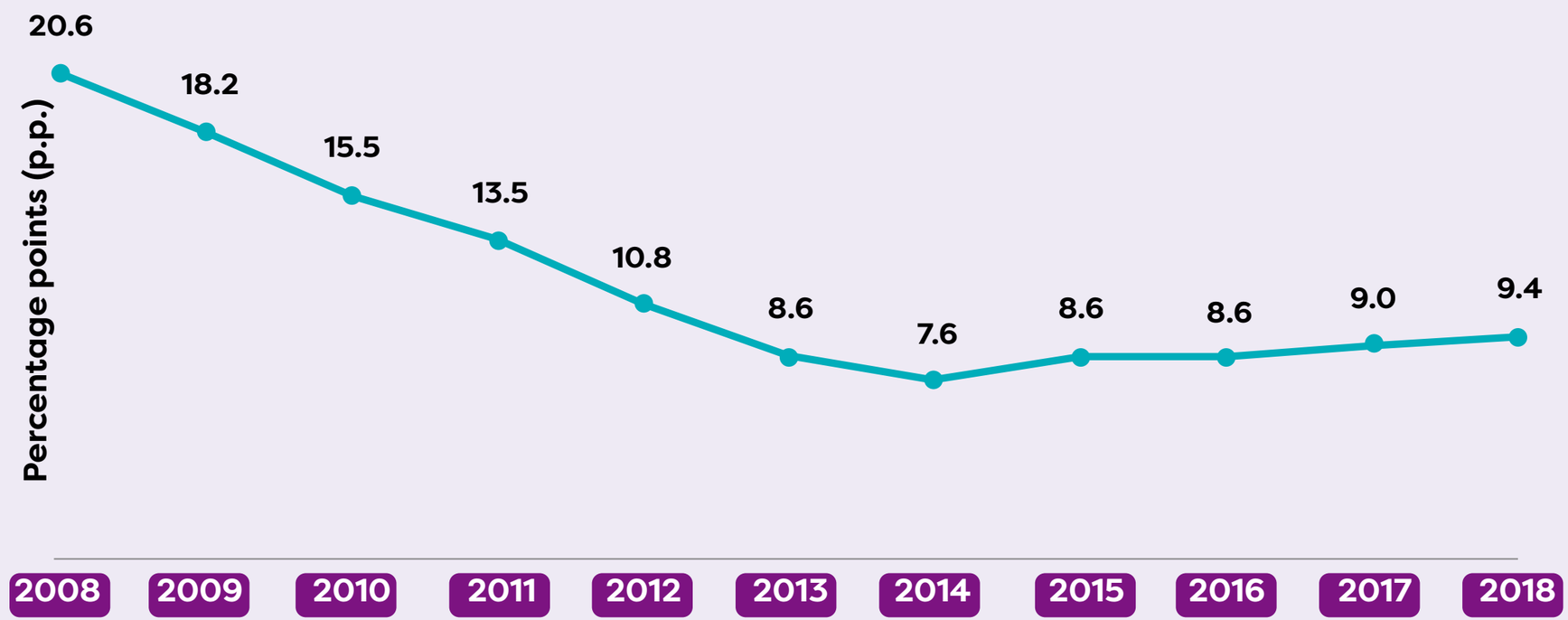
[3] United Nations Resolution A/RES/70/1. Transforming our World: The 2030 Agenda for Sustainable Development
[4] The Colombian education system is made up of: early education, preschool education, basic education (primary five grades and secondary four grades), middle education (two grades and culminates with the bachelor's degree), and higher education (MEN 2018).

When the analysis is extended to preschool, elementary and middle school education levels, the national net coverage gap between urban and rural areas was 9.4 percentage points (p.p.) in 2018, with a decrease of 11.2 p.p. compared to 2008 when the gap was 20.6 p.p. (Graph 2); This reflects a downward trend in the gap between urban-rural net coverage in preschool, elementary and middle school education.

Nevertheless, **the increase in the gap between 2014 and 2018 (from 7.6 to 9.4 p.p.)** evidences a setback in the reduction of inequalities between urban and rural areas in access to preschool, basic, and middle school education, which contributes to reinforce the trend identified in the literature in which the dropout rate is generally higher in rural than in urban areas (Ejaz et al 2019).

Graph 2.

SDG Indicator 4.5.1.C. Gap between urban-rural net coverage in preschool, elementary, and secondary education. National total. 2008-2018. (Percentage points)



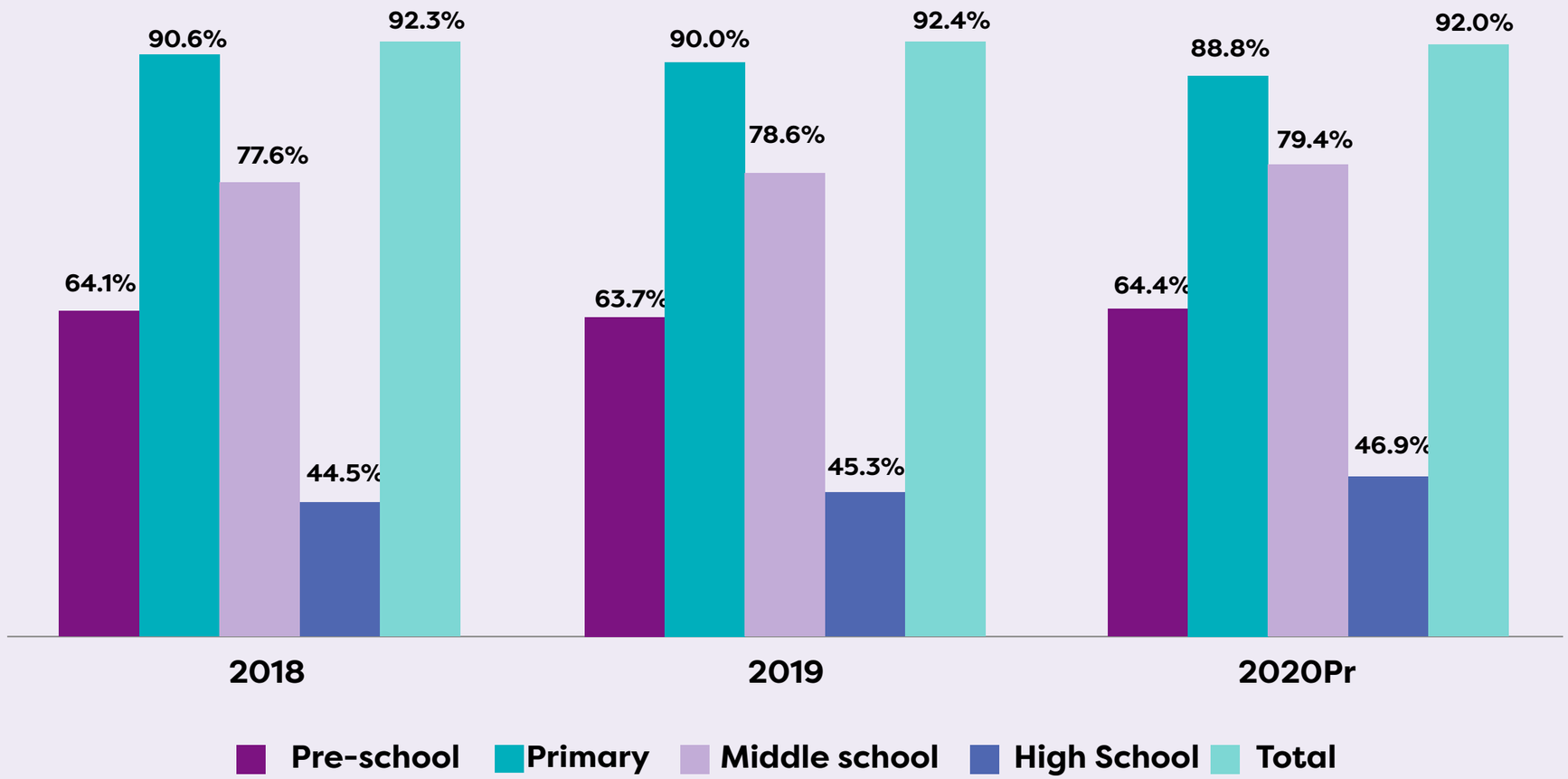
Source: National Education Ministry (MEN, for its Spanish acronyms) –– Integrated Enrollment System (SIMAT, for its Spanish acronym) (2000-2020).

Note:
[1] The gap between net coverage for each zone is calculated as the difference between the urban and rural net coverage rates, that is, the ratio between the total number of enrolled students (of the official age to attend the educational level) residing in urban (rural) residence zones and the urban (rural) population of the theoretical age to attend the educational level; which in the case of preschool, elementary and middle school education corresponds to the range between 5 and 16 years of age.

Between 2018 and 2020Pr, the percentage of students enrolled in preschool, basic and middle school education at the national level on average exceeded 90.0% in relation to the population aged between 5 and 16 years old (Graph 3). When comparing the national net coverage rates in these three educational levels (preschool, elementary and middle school) during 2020Pr, a lower net coverage in middle school education is observed with a rate of 46.9%. Thus, on average, around 53 out of every 100 people in Colombia who were between 15 and 16 years old in 2020, did not have access to middle school education.

Graph 3.

Net coverage rate in preschool, elementary and middle school education. National total. 2018-2020Pr. (Percentage)



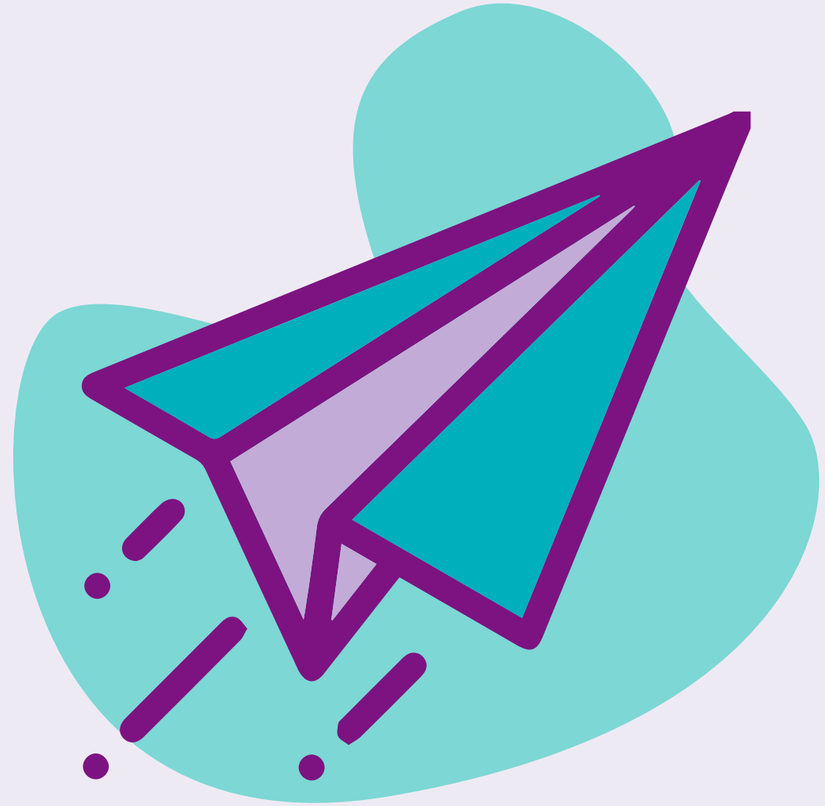
Source: National Education Ministry (MEN, for its spanish acronyms) ––– Integrated Enrollment System (SIMAT, for its Spanish acronym) (2000-2020Pr). DANE, population forecast base on the 2018 National Population and Housing Census (CNPV 2018, for its acronyms in Spanish)

Nota:
Pr Preliminary.
[1] The net coverage rate in preschool, elementary and middle school education measures the percentage ratio between enrollment in education level n that has the official age e to attend this level and the population with the theoretical age e of education level n; where n = preschool, elementary and high school; e = 5 years (preschool), 6-10 years (elementary), 11-14 years (middle school), 15-16 years (high school).

Considering the coverage analysis of the **Certified Territorial Entities** (ETC, for its Spanish acronyms)[5] at the territorial level, **in 2020Pr, 24 of the 96 entities (25.0%) had a total net coverage in preschool, elementary, and middle school education higher than 100%** (Annex 2). Barrancabermeja, Sogamoso, and Riohacha present the highest net coverage rates in these educational levels with 117.9%, 113.2% and 111.9%, respectively; **while San Andrés de Tumaco (68.5%), Vichada (60.2%) and Vaupés (53.5%), have the lowest total net coverage rates in education (preschool, basic, and secondary) at the level of the Certified Territorial Entities** in the country.

Another key element, when analyzing coverage and accessibility to education, is the educational services provision over time. Among the characterization of educational establishments at the departmental level according to status (Annex 3), **in 2021, the departments with the highest proportion of "new active" educational establishments with respect to the departmental total are: Guainía (74.5%), La Guajira (71.5%), and Amazonas (68.1%); in contrast to Caldas (11.8%), Boyacá (4.3%) and Bogotá D.C. (1.5%), which are the departments with the lowest rates of new (active) educational establishments in that year.**
Besides that, in 2021, the departments **with the highest prevalence of educational centers with "temporary" and "definitive" closures -with respect to the total number of educational establishments- are: Vichada (22.5%) and the Archipelago of San Andrés, Providencia and Santa Catalina (26.5%).** These departments are followed by Atlántico (24.9%) and Bogotá D.C. (16.9%). (16.9%); while in the "temporary closure" category, departments such as Arauca (18.3%), Casanare (16.4%) and Putumayo (13.4%) are at the top of the distribution.

After presenting the information related to preschool, elementary and secondary education coverage rates at the national level, based on the information reported in the Integrated Enrollment System - SIMAT and the national framework for monitoring the implementation of the 2030 Agenda and its SDGs in Colombia, the following sections present the development of the pilot exercise conducted by DANE, with support from the Data4Now initiative, on the analysis of accessibility measures between the home and educational institutions; all in line with the contribution to the closing of data gaps and the achievement of SDG 4 "Quality Education".



[5] According to the Law 715 of 2001, certified territorial entities are those departments, districts or municipalities with more than 100,000 inhabitants, in addition to some municipalities with less than 100,000 inhabitants that were certified before the competent authorities, after meeting the requirements set forth in the regulations in technical, administrative and financial matters to autonomously administer the educational service in their jurisdiction.

2

Analysis of educational centers accessibility

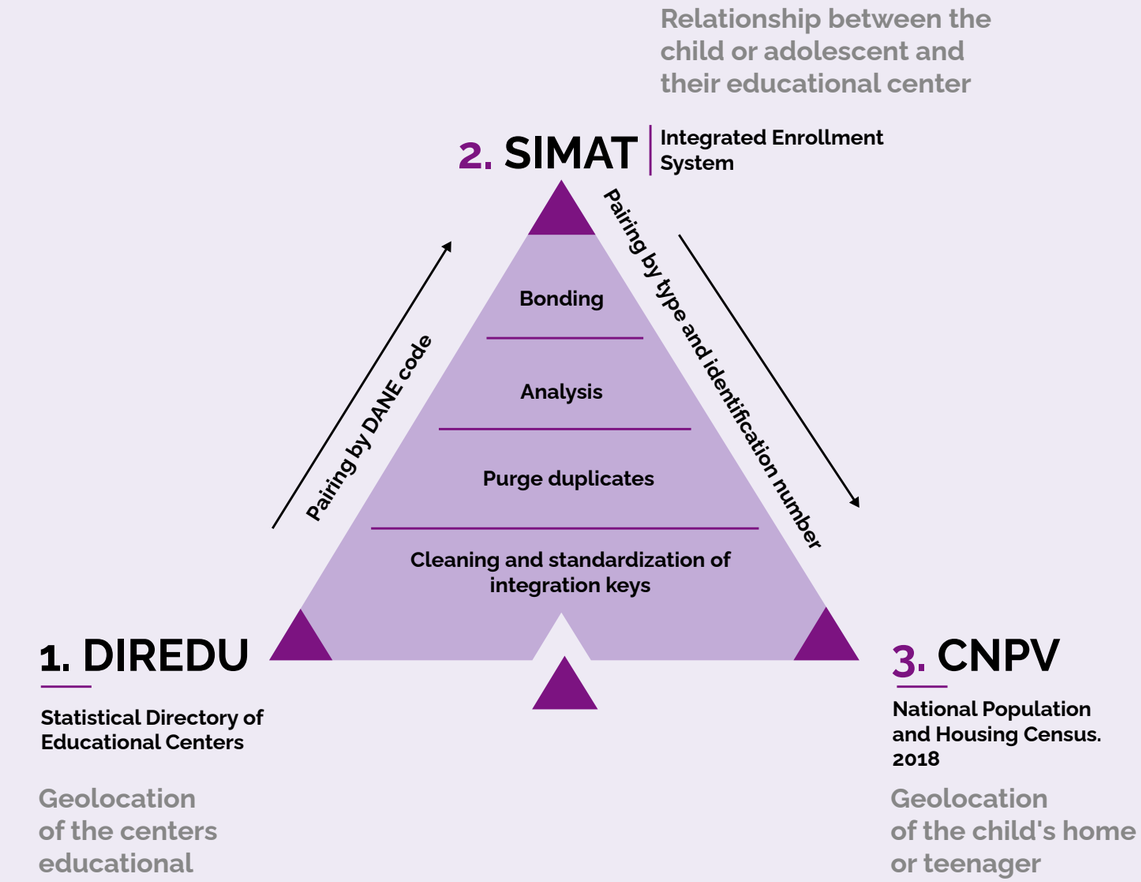
2.1. Methodology

Data and information integration methodology

Three information sources were used in the results presented in this Statistical Brief: the National Population and Housing Census (CNPV 2018, for its acronyms in Spanish), the Statistical Directory of Educational Venues (DIREDU, for its acronyms in Spanish) of DANE and the administrative record of the Integrated Enrollment System (SIMAT, for its acronyms in Spanish). A triangulation process was carried out, recognizing the strengths and weaknesses of each of these sources, as shown in the Illustration 1 for a standardized integration based on unique keys such as DANE codes of educational centers and types and identification numbers of children and adolescents.

Illustration 1.

Integration process of information sources.



Source: DANE 2022. SIMAT - CNPV 2018 - DIREDU.

First, the National Population and Housing Census (CNPV 2018) had the purpose of counting the resident population in the national territory and obtaining sociodemographic information aiming to plan, manage and make decision on the national, territorial, and local levels (DANE, 2018). In this sense, although the CNPV 2018 is a source with a significantly high level of coverage in terms of the location of households, it is a time invariant source, since it does not capture population movements and, therefore, it could be assigned an erroneous location for the housing of some children and adolescents, mainly for the years 2019 and 2020.

Secondly, DANE's Statistical Directory of Educational Centers (DIREDU)[6] includes a list of primary and secondary education centers active in the national territory and incorporates periodically updated information on identification data, location, contact and classification by academic calendar and sector: official (public) and non-official (private)[7]. This directory is integrated by information sources such as the Educational Centers Information System (SISE, for its acronyms in Spanish) operated by DANE[8], the Single Directory of Educational Centers (DUE, for its acronyms in Spanish), administered by the Ministry of National Education (MEN[9], for its acronyms in Spanish), the C-600[10] form, the information of the National Educational Information System (SIEF, for its acronyms in Spanish), headed by the MEN, and the administrative record of the SIMAT.

Finally, the Integrated Enrollment System (SIMAT) is a tool that allows to organize and control the enrollment process in all its stages, as well as to have a reliable and available source of information for decision making. It is a system for managing the enrollment of students in official institutions that facilitates the registration of new students, the registration and updating of existing student data, the consultation of the student by institution and the transfer to another institution, among others (MEN, n.d.).

[6] For more information on DIREDU's statistical strengthening, see Annex 1.
[7] According to the National Education Ministry, the sector is an identifier of the educational establishment that indicates whether it is private or not (official or non-official) (MEN, n.d.-c).
[8] The SISE is the tool where DANE codes are assigned for the individual identification of legally constituted educational centers in the country, facilitating the interoperability of the information systems of the education sector in Colombia (DANE, 2020). For more information, consult the SISE geo-viewer of educational centers at:
https://geoportal.dane.gov.co/geovisores/sociedad/consulta-geosise/?lt=4.456007353293281&lg=-73.2781601239999&z=5
[9] The DUE is the tool that each Education Secretary can manage its educational establishments and report news to the National Education Ministry (MEN, n.d.-b).
[10] The single census form C-600 collects information from educational centers in the formal education sector. Available for consultation at: https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/poblacion-escolarizada/educacion-formal

For the proposed exercise, two main inputs were required, among others: the coordinates of the student's residence and the coordinates of the educational center in which he/she was enrolled during the 2018-2020 period. As a starting point, the exercise takes the information consolidated in CNPV 2018, which includes the address and residence coordinates of each household and the identification information of each of its members. This information was integrated with the SIMAT for the years 2018, 2019 and 2020 to obtain the universe of students enrolled in the different educational centers in the country and their respective geolocation according to the location of the household registered in the census.

Following, after the integration between SIMAT and CNPV 2018, the coordinates of the educational centers were incorporated using DIREDU whose key integration variable is DANE code of the educational centers. It is important to mention that, due to the geographical particularities of the country, DIREDU has a level of completeness of its coordinates of 96.3%, where the remaining 3.7% is made up of educational sites located in rural areas. It should be noted that there were some challenges in the use of the DANE code of the educational institutions as a key to integration, mainly due to the existence of some temporary codes.

After carrying out the integration process, quality checks were performed for the consolidation of the final database, from which 0.012% of the records were eliminated due to duplicities and inconsistencies of information. Regarding the universe of enrolled students identified in SIMAT for the years 2018-2020, the following considerations were made:

- Initially, the exercise selected the population between 5 and 17 years of old since, according to the General Education Law, education in Colombia is compulsory for children between 5 and 15 years old, covering the transition grades (i.e., grade 0) up to ninth grade (Congreso de la República de Colombia, 1994). Additionally, the 16-year-old population was included, since, for 2019, the linkage rate in full-time and part-time secondary education programs was 75%. For the 17-year-old population, this rate drops to 42% in secondary education and 13% in tertiary education, which shows an inflection point towards a reduction in linkage rates (OECD, 2021).This population corresponds to 89.7% of the records.

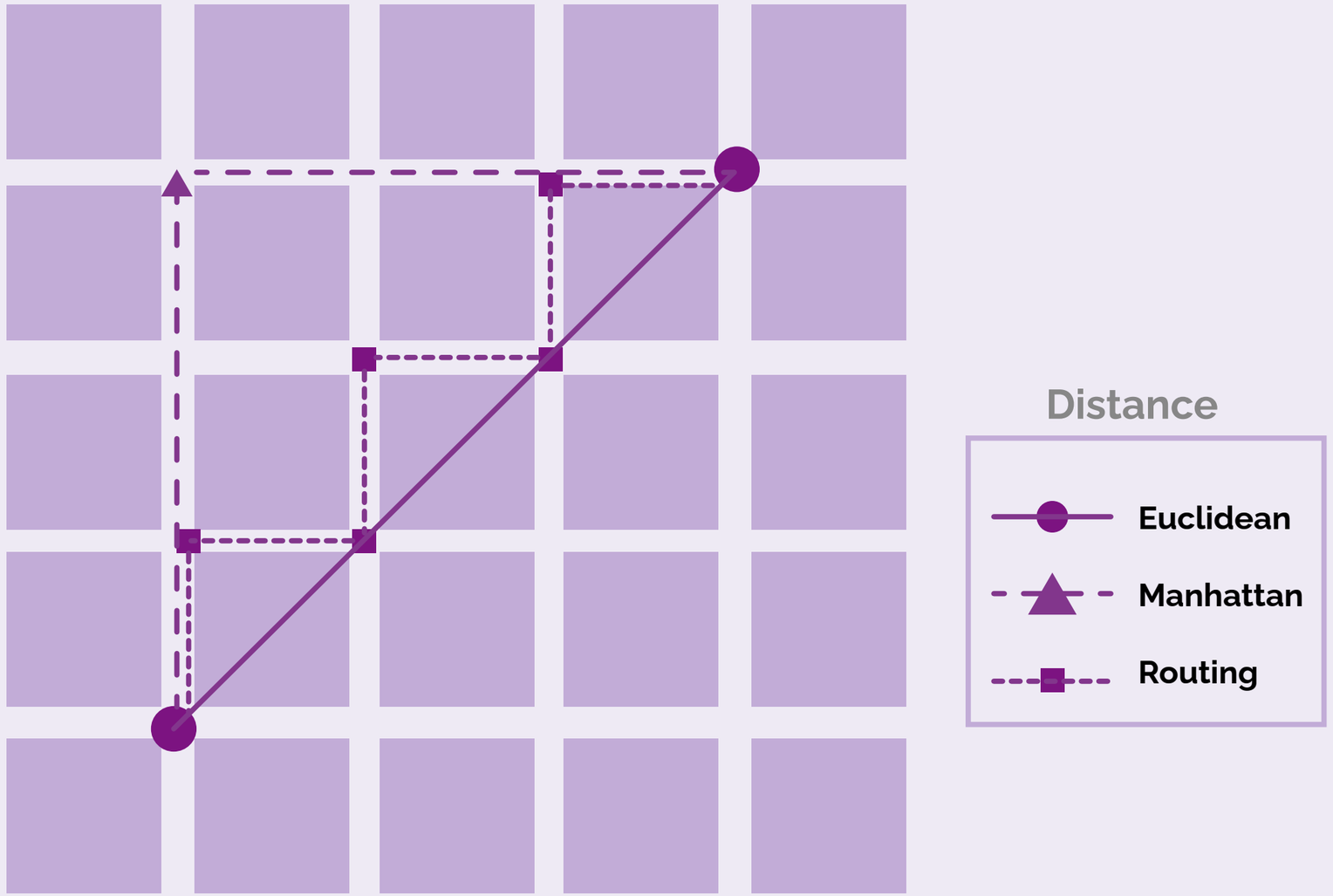
- On the other hand, this study focused on those students enrolled in educational centers whose methodology is classified as traditional, i.e., that which follows pedagogical and didactic principles focused on the transmission of knowledge. This model includes educational levels from preschool through high school. The population enrolled in educational centers that offer this educational model corresponds to 81.7% of the records. Focusing on the 5- to 17-year-old range, this percentage of coverage is 76.4%. Finally, those students enrolled in grades from transition to 11 were selected, which leads to a final sample that corresponds to 76.1% of the records.

Once the inputs were incorporated into the SIMAT bases for the years 2018-2020 and the processes described above were carried out, the points that allowed calculating the distance from the student's place of residence to the educational center in which he/she was enrolled in each year were completed. The following is a description of each of the methodologies used in the approximation of accessibility levels between homes and educational centers based on three distance metrics (Euclidean, Manhattan, routing).

Distance Metrics

Geospatial Information Systems (GIS) use different methodologies and approaches to measure accessibility. The level of access that different types of agents can have between two locations comprises aspects such as time, mode of transport and physical distance (Jang & Kang, 2015). For the exercise described in this Statistical Brief, several distance metrics were implemented to make a first approximation to a measure of the accessibility of students between 5 and 17 years old enrolled between transition and grade eleven (11) in educational centers that provide traditional education. For this purpose, we have three distance metrics (Euclidean, Manhattan, routing) presented in Illustration 2.

Illustration 2. Graphical representation of the three distance metrics (Euclidean, Manhattan, routing).



Source: DANE 2022.

Euclidean distance metric

The Euclidean distance metric is a spatial proximity tool that determines the closest elements to each other by calculating the straight-line distance between them. **The Euclidean distance between two points is defined as the length of a segment connecting them,** which is calculated from Cartesian coordinates. This calculation was performed for all the records of the SIMAT-CNPV 2018-DIREDU integrated databases for the period 2018-2020, which had the two pairs of coordinates (location of students and location of educational sites), obtaining measurements for about 70% of the records in those periods, as shown in Table 1.

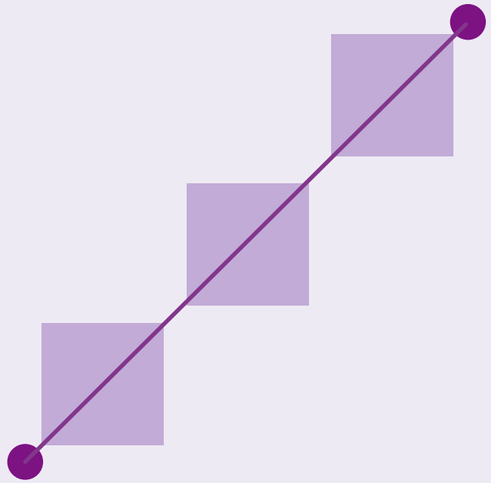


Table 1.

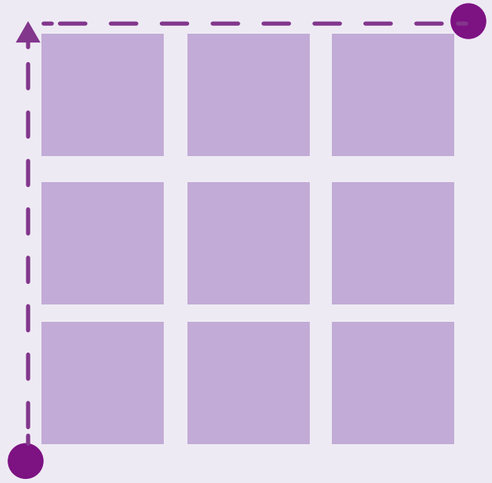
Implementation rate of the Euclidean distance calculation methodology on the SIMAT - CNPV 2018 - DIREDU integrated database records. National total. Urban and rural areas. 2018-2020.
(Data in absolute values and percentages)

Year	Total		Location of student´s home			
			Urban		Rural	
	No covered	Covered	No covered	Covered	No covered	Covered
2018	2,834,708	4,800,552	1,636,837	4,463,561	1,197,871	336,991
	37.1%	62.9%	26.8%	73.2%	78.0%	22.0%
2019	3,347,472	4,405,110	2,070,599	4,118,819	1,276,873	286,291
	42.2%	56.8%	33.5%	66.6%	81.7%	18.3%
2020	2,535,867	5,224,430	2,059,580	4,140,626	476,287	1,083,804
	32.7%	67.3%	33.2%	66.8%	30.5%	69.5%

Source: DANE 2022.

Manhattan distance metric

The Manhattan distance between two points is the sum of the absolute differences in their coordinates. It is known as the cab drivers' geometry, and its name is based on the regular grid design with which the island of Manhattan was built. Compared to Euclidean distance, this metric seeks to analyze minimum routes in cities, being more "realistic" than straight-line distance. Due to the characteristics of the metric, it was calculated on the records associated with municipal capitals that had the two pairs of coordinates mentioned above.



Location of student´s home

Table 2.

Implementation rate of the Manhattan distance calculation methodology on the SIMAT - CNPV 2018 - DIREDU integrated database records National total, urban and rural areas. 2018-2020.
(Data in absolute values and percentages)

Year	Total		Urban		Rural	
	No covered	Covered	No covered	Covered	No covered	Covered
2018	2,834,708	4,800,552	1,636,837	4,463,561	1,197,871	336,991
	37.1%	62.9%	26.8%	73.2%	78.0%	22.0%
2019	3,347,472	4,405,110	2,070,599	4,118,819	1,276,873	286,291
	43.2%	56.8%	33.5%	66.6%	81.7%	18.3%
2020	3,565,840	4,194.457	2,266,328	3,933,878	1,299,512	260,579
	46.0%	54.1%	36.6%	63.5%	83.3%	16.7%

Source: DANE 2022.

Distance per routing

The Dijkstra algorithm was used, also called the minimum paths algorithm, whose purpose is **to determine the shortest route from a vertex of origin and the rest that belong to a graph,** with which the shortest paths between a node and the others are explored. This was done, once again, based on the previously defined inputs (coordinates of students and educational centers), in addition to the use of the Open Street Map (OSM).

Al first, we used records with head town locations, since the requirements of the algorithm, together with the availability/ quality of the road network used (Open Street Map), would guarantee better results in this context. Subsequently, the application was made in rural records where the conditions of the inputs would allow the algorithm to calculate distances. As can be seen in Table 3, the percentage of records obtained under this methodology for rural areas is less than 10.0%. It should also be noted that the above processes were carried out on a PostGIS spatial database engine, which allowed the processing of robust databases efficiently in a spatial context.

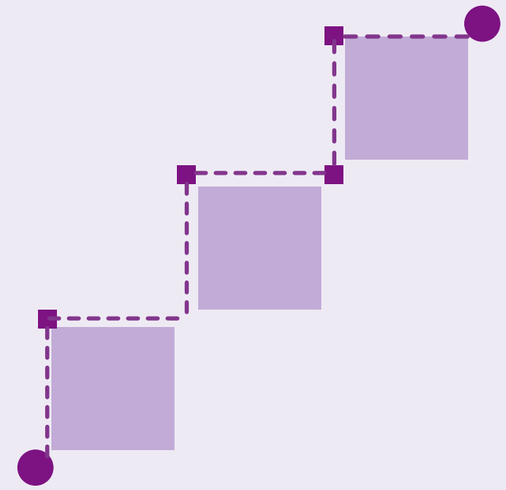


Table 3.

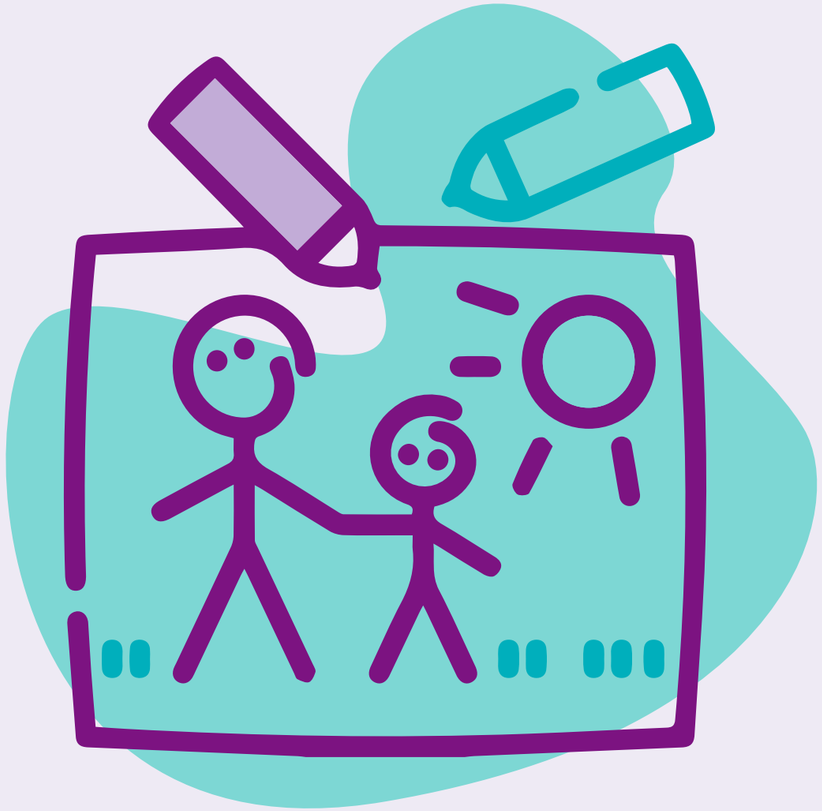
Implementation rate of the routing distance calculation methodology on the SIMAT - CNPV 2018 - DIREDU integrated database records National total. Urban and rural areas. 2018-2020.
(Data in absolute values and percentages)

Year	Location of student´s home					
	Total		Urban		Rural	
	No covered	Covered	No covered	Covered	No covered	Covered
2018	5,562,953	2,072,307	4,164,190	1,936,208	1,398,763	136,099
	72.9%	27.1%	68.3%	31.7%	91.1%	8.9%
2019	5,912,316	1,840,266	4,457,678	1,731,740	1,454,638	108,526
	76.3%	23.7%	72.0%	28.0%	93.1%	6.9%
2020	5,986,339	1,773,958	4,521,961	1,678,245	1,464,378	95,713
	77.1%	22.9%	72.9%	27.1%	93.9%	6.1%

Source: MEN - DANE 2022

Considering the percentage values of records obtained for the calculation of distances presented, **this exercise will concentrate on urban areas**, where, for the period 2018-2020. The maximum crossover percentage is 73.2% of the total number of students in the universe defined for the year 2018 under the Euclidean and Manhattan metrics, and a minimum of 63.5% for the year 2020 for both methodologies (Table 1 and Table 2). Given the low crossover percentages in the rural sector, where a maximum of 22.0% is reached for the year 2018 under the Euclidean and Manhattan methodologies, students residing in this area will be excluded from the distance analyses; although some variables not affected by the implementation rate of the distance calculation methodology will be presented at the beginning of the descriptive analysis. Under this scenario, **the descriptive analysis of the distance metrics was performed on 79.9% of the total records, that is, students between 5 and 17 years old enrolled between transition and grade eleven (11) in educational centers that provide traditional education and live in the urban sector.**

The results for the routing distance metric will be presented only as a contrast, since the crossover rate of records obtained for both rural and urban areas is less than 32.0% of the total number of students (Table 3). Finally, although the rates in terms of the computational capacity of the records for the Euclidean distance metrics are higher than 66.0% for all years (2018, 2019, and 2020), due to possible errors in the geolocation of the housing or educational center of some students, the analysis will be limited to those distances less than 100 kilometers (km), which implies a reduction of 2.24% records of the analysis universe.



2.2. Results

In line with the above, first, statistics of gross coverage rates in education and school dropout rates[11] are presented, with geographic disaggregation at urban and rural areas level, for the years 2018, 2019, and 2020. After that, the main results of the project to analyze the distance to educational centers and its relationship with the accessibility to the Colombian educational system in terms of school dropouts are presented. In this case, the statistical unit for the distance metrics (Euclidean, Manhattan and routing) corresponds to students between the ages of 5 and 17 enrolled between 2018 and 2020 in educational centers that follow a traditional education methodology, and who reside in the urban sector.

In 2020, the gross education coverage rate, for the population between 5 and 17 years old, at the national level was 74.9%. Approximately 25 out of every 100 children and adolescents in this age range did not have access to education in the country.

Table 4.

Education gross coverage rate for the population between 5 and 17 years old
National total. Urban and rural areas.
2018-2020. (Data in absolute values and percentages)

Year	Population CNPV 2018		Enrolled in SIMAT		Gross coverage ratio			Gap (%)
	Rural	Urban	Rural	Urban	Total (%)	Rural (%)	Urban (%)	
2018	2,861,006	7,387,750	1,534,862	6,100,398	74.5	53.7	82.6	28.9
2019[1]	2,886,336	7,432,035	1,563,164	6,189,418	75.1	54.2	83.3	29.1
2020[1]	2,904,501	7,462,419	1,560.091	6.200,206	74.9	53.7	83.1	29.4

Source MEN - DANE 2022
Note:
[1] Population forecast based on National Population and Housing Census (CNPV, for its Spanish acronym) 2018.

[11] The gross coverage rate measures the percentage ratio between the total number of students enrolled (regardless of their age) over the total population of the theoretical age to attend the educational level (5-year-old (transition), 6-10 years-old (primary), 11-14 years-old (secondary), 15-16 years-old (middle).

The Integrated Enrollment System administrative record (SIMAT, for its acronyms in Spanish) **shows an education gross coverage rate of the population between 5 and 17 years old in a range between 74.5% and 75.1% for the years 2018, 2019, and 2020, at the national level** (Table 4). This coverage is higher in urban areas, reaching a range of 82.6% to 83.1% for the period, while for the rural sector the education gross coverage rate lied between 53.7% and 54.2% during the years 2018-2020.

The results of gross coverage rates in education for the years 2018-2020 evidence the information needs regarding the generation of timely and relevant statistics on the main education variables in the rural sector, and together with the coverage indicators at the preschool, basic, and secondary education levels presented in the first section of the Statistical Brief, reinforce the need to have quality georeferenced information for decision making and enrich the supply of measurements on the education sector in Colombia.

In urban areas, the distribution of coverage between males and females is relatively homogeneous across years (2018-2020) and grades (preschool (0) to eleven (11)). **In 2018, the proportion of females enrolled was 50.3%, while the proportion of males was 49.7%. These proportions remained almost unchanged for 2020, where the proportion of females enrolled was 50.2%, and the proportion of males was 48.9%.** Figure 6 presents this distribution in absolute terms by educational levels for the years 2018, 2019, and 2020 in urban areas.

Graph 4.

Population between 5 and 17 years old enrolled in the country's educational centers, by sex and grade level 0 a 11). National total. Urban areas. 2018-2022. (Data in absolute values)

Source: DANE 2022.

Nota:

[1] According with the Law 115 of 1994, formal education in Colombia is organized according to three levels: preschool, which includes a compulsory grade (0), basic education (primary and secondary) of nine (9) grades, and middle school with a duration of two (2) grades.

In terms of **school dropouts**, understood as the abandonment of the school system by students due to a combination of factors generated both within the system and in the social, family, individual and environmental contexts (MEN, n.d.-c), based on the SIMAT registry, **between 2018 and 2020, 150,099 students located in urban areas abandoned their studies[12].**

- This represents 2.47% of the annual average of students enrolled at the beginning of each year.
- In the case of the **rural sector**, this average proportion is 3.13%, which represents 48,529 children or adolescents.
- These proportions are maintained when differentiating educational centers belonging to **the official and non-official sectors.** Although the dropout rate is higher for official schools, except for the year 2020, when the dropout rate in non-official schools was 0.64 percentage points (p.p.) higher than that of official schools.
- In the analysis for only **urban areas**, the annual average is 2.47%, and the difference between official and non-official schools ranges between 0.54 and 1.05 percentage points (p.p.) for the years 2018 and 2019 in favor of official schools. In 2020, this trend was reversed, and dropout in non-official schools was 0.74 p.p. higher than in official schools.
- From a gender perspective, the dropout **gap between males and females** was 0.58 percentage points (p.p.) for 2018 and drops to 0.47 p.p. in 2020, with males presenting higher dropout rates. When analyzing only urban areas, this pattern is maintained, and the dropout gap varies between 0.4 p.p. and 0.5 p.p. between 2018 and 2019, with males presenting higher dropout rates.

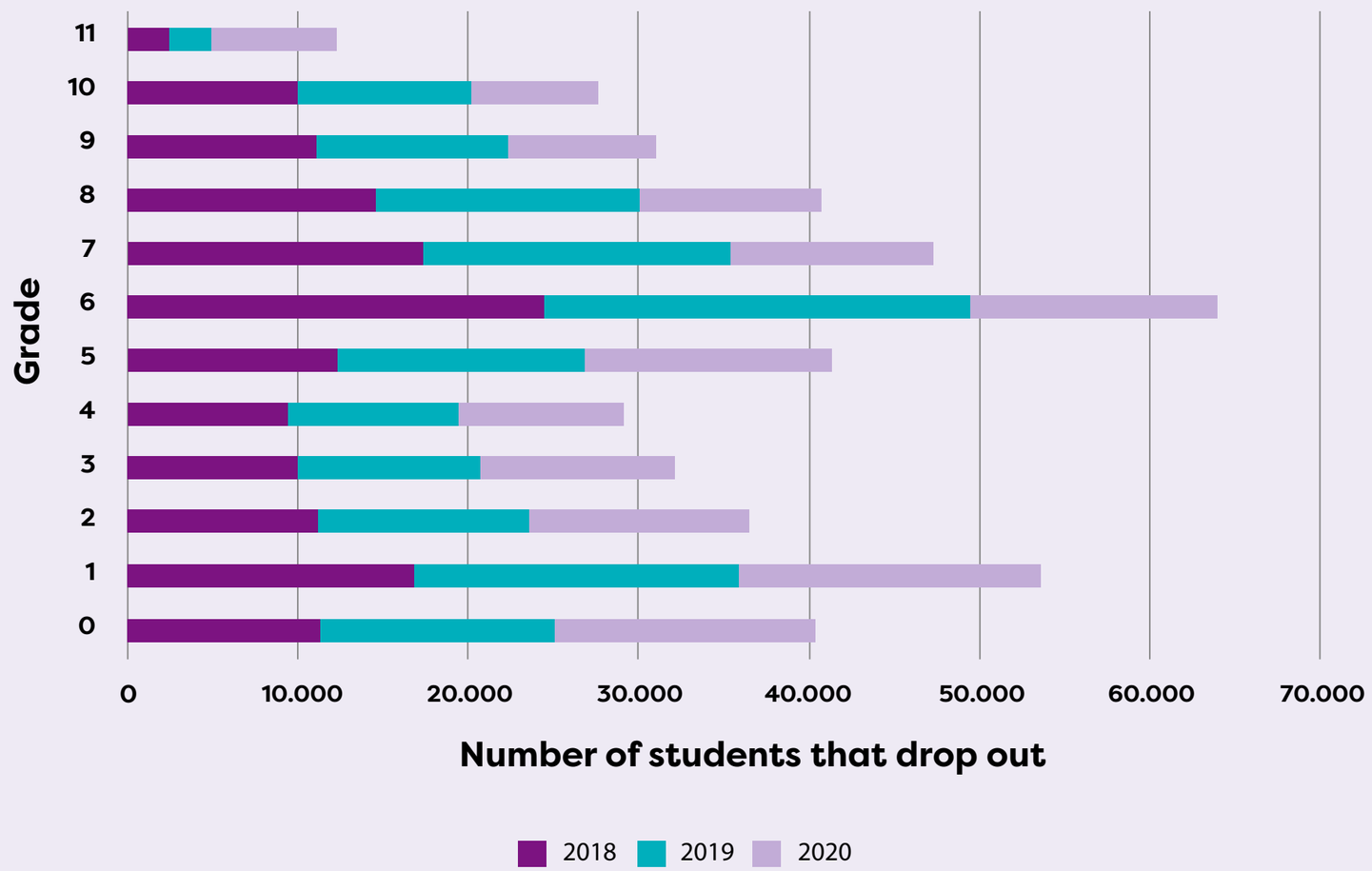
Between 2018 and 2020, sixth grade (6) was the educational level where there was the highest dropout in absolute terms from formal education in urban areas of Colombia, with 63,953 students out of the 456,297 who dropped out of preschool, elementary and middle school during the period of analysis, corresponding to 14.0%.

[12] SIMAT has two variables that were used for the construction of the count of students who dropped out: the student's condition at the end of the previous year, one of which is dropout, and the efficiency variable that reports the dropout in the current year.

Graph 5 presents the number of students in urban areas, by year, who dropped out between 2018-2020, in the levels of transition (0) to eleven (11). A strong increase in dropout levels is observed when students reach sixth grade (6) and, on the contrary, the lowest dropout levels between 2018 and 2020 were presented at the middle school level (grade eleven (11)), there is a dropout of 17,400 students in the referenced period.

Graph 5.

Population between 5 and 17 years old who dropped out of the school system, by grade (0 to 11). National total. Urban areas. 2018-2020. (Data in absolute values)



Source: DANE 2022.

Nota:

[1] In accordance with Law 115 of 1994, formal education in Colombia is organized according to three levels: preschool, which will comprise a mandatory grade (0), basic education (primary and secondary) of nine (9) grades, and secondary education with a duration of two (2) grades.

Under this landscape, **the project developed under the Data4Now initiative seeks to analyze the existing relationship between the distance from home in urban areas for students between 5 and 17 years old and the educational center in which they were enrolled for the years 2018-2020,** as an approximation to accessibility and its corresponding correlation with student dropout; all in line with the Sustainable Development Goal (SDG) 4 "Quality Education" and its premise of leaving no one behind.

Table 5 presents the general results of the distance metrics (Euclidean, Manhattan, and routing) between homes and educational centers of children and adolescents between 5 and 17 old enrolled in educational centers in urban areas of the country in grades transition (0) to eleven (11):

On average, **under the Euclidean metric, the distance is 2.6 kilometers (km) for 2018, 3.2 km for 2019, and 4.0 km for 2020.** Additionally, significant dispersion is observed in the data under this spatial proximity methodology, as the standard deviation ranges from 7.3 to 10.9 kilometers during the period 2018-2020[13].

The **Manhattan distance metric**, like the Euclidean distance, uses the coordinates of the educational centers and homes of the children and adolescents (equal number of observations for each of the years), **presents an average of 3.3 km in 2018, 4.0 km for 2019 and 4.5 km for 2020**. In the same way, the standard deviation is greater than that presented under the Euclidean methodology, being in a range between 9.4 and 12.8 km during the period 2018-2020.

[13] As mentioned in the methodology section of this Statistical Brief, the coordinates of the students' homes for the calculation of the distances to the educational centers where they were enrolled had as a source of information the National Population and Housing Census (CNPV 2018). As this is a time-invariant source, it does not capture population location movements and, therefore, could be assigning erroneous values for the housing of some children and adolescents, particularly for the years 2019 and 2020.

Table 5.

Descriptive statistics of distance metrics (Euclidean, Manhattan, routing) for urban areas.. 2018-2020.
(Values in meters (m))

Year	Metrics	Half	Median	Percentile 90	Standard deviation	Minimum	Maximum	Obs
2018	Euclidean	2,621.5	892.1	5,063.3	7,332.1	0.0	99,999.1	4,322,248
	Manhattan	3,321.6	1,128.3	6,429.5	9,435.0	0.0	141,023.0	4,322,248
	Routing[1]	1,502.5	848.0	3,228.6	2,270.6	1.4	87,694.0	1,936,206
2019	Euclidean	3,147.5	928.1	5,822.0	9,001.8	0.0	99,998.7	3,947,084
	Manhattan	3,998.9	1,174.3	7,364.2	11,619.3	0.0	141,035.7	3,947,084
	Routing[1]	1,522.0	857.5	3,263.6	2,314.7	1.4	126,705.9	1,731,739
2020	Euclidean	4,010.4	1,011.2	7,601.7	10,928.3	0.0	99,998.7	3,935,077
	Manhattan	4,451.8	1,217.9	8,140.5	12,793.5	0.0	141,061.0	3,761,642
	Routing[1]	1,546.5	869.2	3,322.8	2,348.4	1.1	126,942.2	1,621,204

Source: DANE

Nota:

[1] Considering the low coverage rates of the routing distance, this methodology presents significantly lower results than the other methodologies. However, by restricting the sample to only those observations with routing distance information (i.e., between 1,621,204 and 1,936,206 observations depending on the year of analysis), the mean of this methodology becomes the largest of the three and the mean Euclidean distance drops to 814.8 in 2018, 827.3 in 2019, and 843.8 in 2020

The analysis of these approximations according to the sector to which the educational center belongs (official and non-official) observed that **the distances for students enrolled in official educational centers are shorter than the distances for students in non-official schools located in urban areas.**

Table 6 presents the general results of the distance metrics (Euclidean, Manhattan, and routing) between homes and educational centers of children and adolescents between 5 and 17 years old enrolled in educational centers in the country's urban areas in grades transition (0) to eleven (11), **according to the sector to which the educational centers belong (official and non-official):**

For the Euclidean metric, the difference in average distances ranges from 210.1 meters (m) to 1.1 kilometers (km) for 2018, 2019, and 2020. This difference is reduced when moving from 2018 to 2020, which may be due in part to the accuracy of the geolocation of children and adolescents' housing when using the 2018 CNPV, given that it is a time invariant source.

This fact might be related to processes of residential segregation and segregation in the school sphere, as well as the household's ability to pay. This last point is particularly relevant for the household's decision on how to make the journey between housing and the educational center. Rodríguez-López et al. (2017) shows that in Spain for children that the urban sector after 1,250 meters and for adolescents after 1,350 meters, students seek alternative means of transport other than walking to make these journeys. Additionally, people who attend schools in the informal sector may be willing to travel longer distances between their homes and the educational center using means of transportation such as private vehicles or school routes.

Table 6.

Descriptive statistics of distance metrics (Euclidean, Manhattan, routing) for urban areas, by school sector.
2018-2020.
(Values in meters (m))

Year	Metrics	Sector	Half	Median	Standard deviation	Minimum	Maximum	Obs
2018	Euclidean	Official	2,311.7	800.3	7,483.5	0.0	99,999.1	3,125,681
		No official	3,430.8	1,360.5	6,855.4	0.1	99,998.7	1,196,567
	Manhattan	Official	2,931.8	1,011.0	9,612.7	0.0	140,942.4	3,125,681
		No official	4,339.9	1,722.7	8,873.8	0.1	141,023.0	1,196,567
	Routing[1]	Official	1,428.5	835.0	2,176.9	1.4	87,694.0	1,511,711
		No official	1,766.0	919.2	2,559.7	1.4	52,238.9	424,495
2019	Euclidean	Official	2,927.8	829.1	9,381.0	0.0	99,998.7	2,937,840
		No official	3,787.2	1.470.5	7,758.4	0.1	99,990.8	1,009,244
	Manhattan	Official	3,725.2	1047.8	12,097.7	0.0	141,035.7	2,937,840
		No official	4,795.6	1.861.3	10,056.3	0.1	140,972.1	1,009,244
	Routing[1]	Official	1,441.1	836.5	2,208.4	1.4	126,705.9	1,389,409
		No official	1,850.4	987.0	2,678.6	1.4	87,022.4	342,330
2020	Euclidean	Official	3,952.3	903.9	11,572.8	0.0	99,998.7	2,846,893
		No official	4,162.4	1,502.8	9,025.1	0.0	99,994.8	1,088,184
	Manhattan	Official	4,243.8	1,084.7	13,412.1	0.0	140,885.4	2,698,471
		No official	4,979.8	1,840.5	11,051.6	0.1	141,061.0	1,063,171
	Routing[1]	Official	1,467.5	849.0	2,256.0	1.4	126,942.2	1,254,235
		No official	1,816.4	954.0	2,622.2	1.1	87,022.4	366,969

Source: DANE 2022.

Nota:
[1] Taking into account the low coverage rates of the routing distance, this methodology presents significantly lower results than the other methodologies. However, by restricting the sample to only those observations with routing distance information (i.e., between 1,621,204 and 1,936,206 observations depending on the year of analysis), the mean of this methodology becomes the highest of the three methodologies.

In 2020, **Vaupés, Amazonas, Putumayo y Meta** departments with the greatest distances from the Euclidean metric in urban areas

From the territorial perspective, Graph 7 and Graph 8 present the median Euclidean distance index[14] for each of the country's departments between 2018 and 2020:

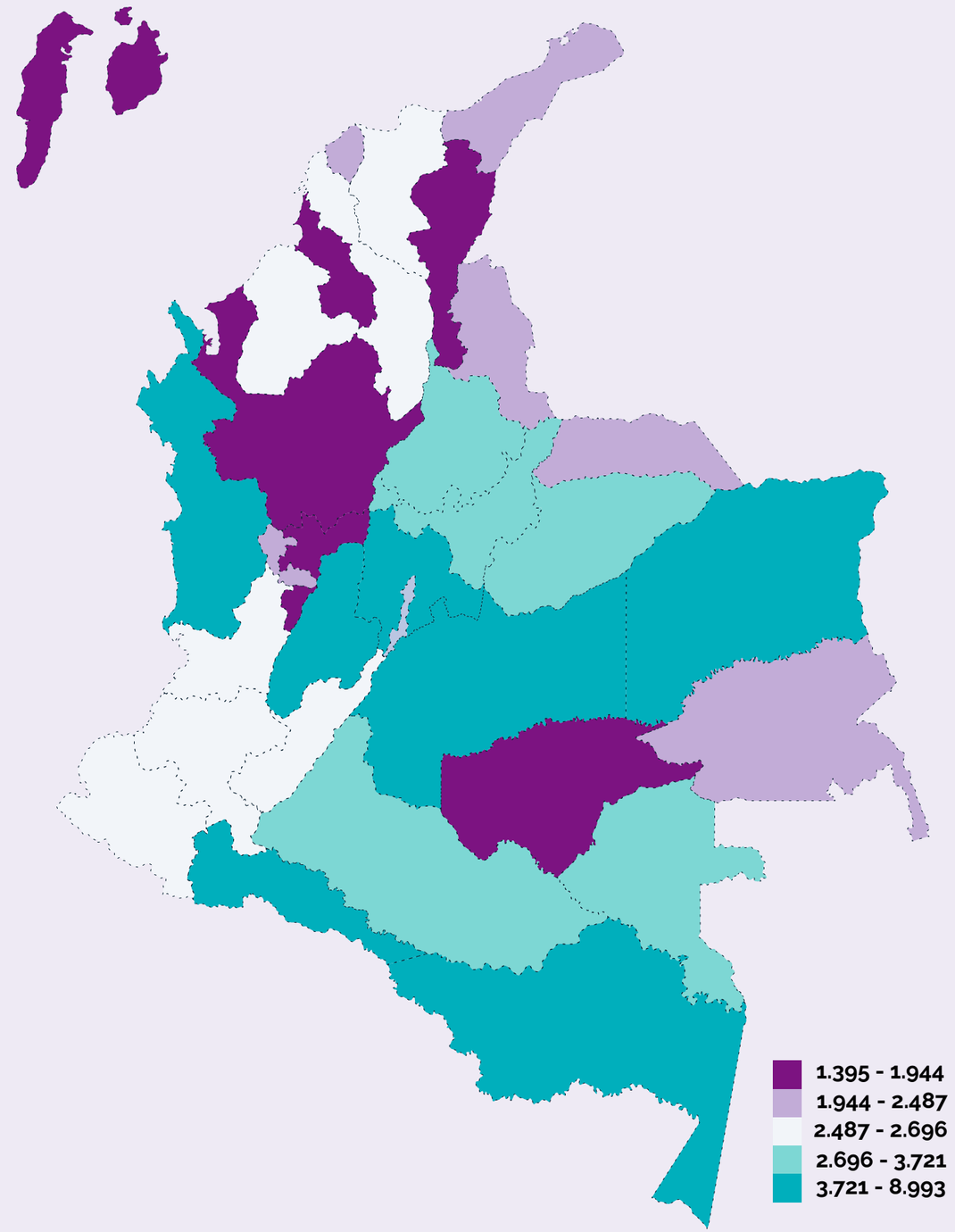
Among the department with the greatest distances in 2018, were Amazonas, Chocó, Putumayo, Cundinamarca, and Vichada **(Illustration 3)**.

In 2020, the five departments with the greatest distances under the Euclidean methodology, that is, the metric of greatest proximity between the coordinates of educational centers and students' homes, **were Vaupés, Amazonas, Chocó, Putumayo, and Meta,** with distances ranging between 6.3 and 11.9 kilometers for that year **(Illustration 4).**

[14] This index takes the Human Development Index methodology to standardize distances at the municipal level and calculates the median of these standardized metrics at the departmental level more information on DIREDU's statistical strengthening, see Annex 1.

Illustration 3.

Mean Euclidean Distance Index1 for urabn areas, by deparment. 2018. (Values in meters (m))

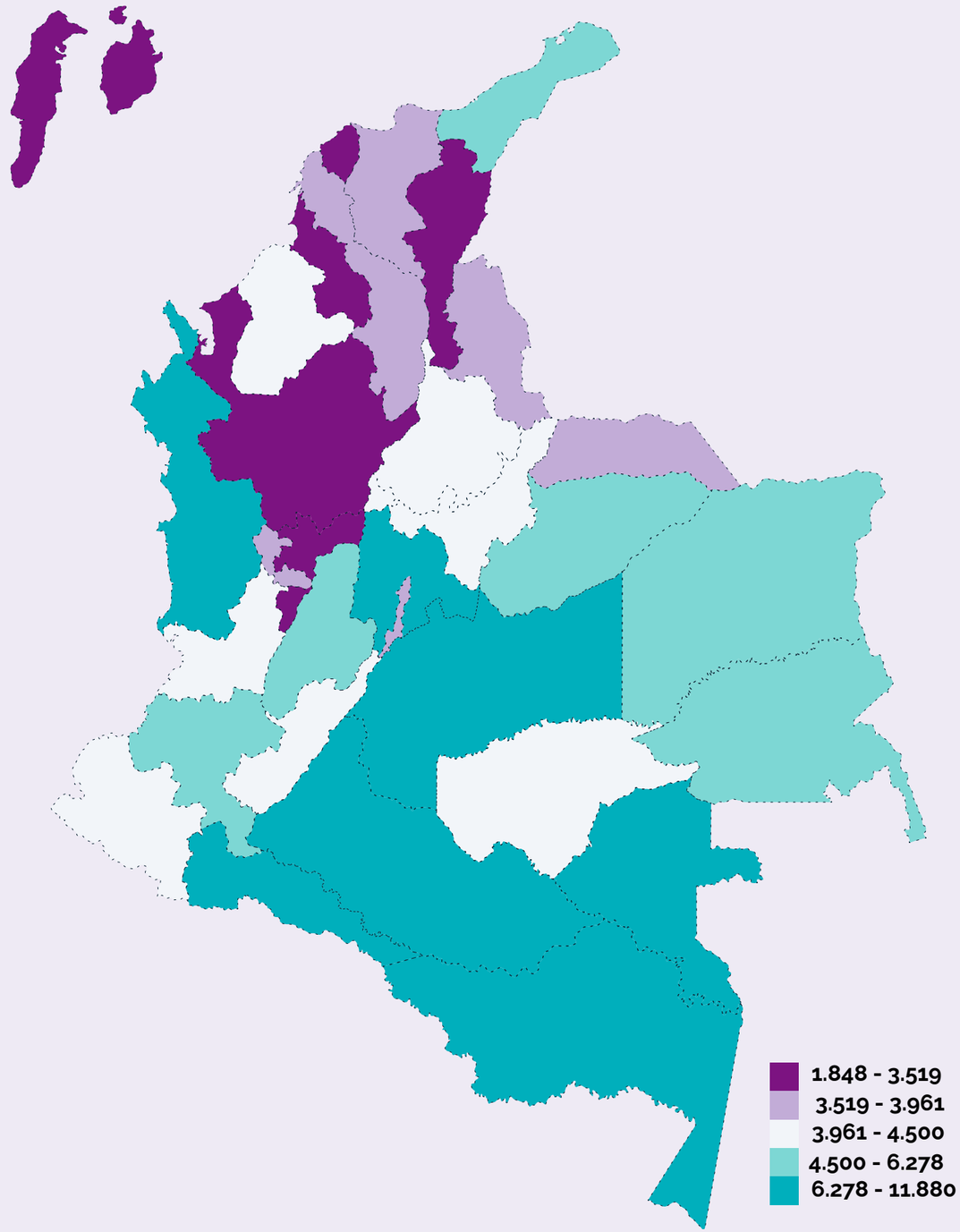


Source: DANE 2022.

Nota:
[1] The construction of the Euclidean distance index consists of standardizing the distances based on the methodology of the Human Development Index, that is, standardizing by the maximum difference of distances within each municipality, and then calculating the departmental average of the standardized values.

Illustration 4.

Mean Euclidean Distance Index[1] for urban areas, by deparment. 2020. (Values in meters (m))



Source: DANE 2022.

Nota:
[1] The construction of the Euclidean distance index consists of standardizing the distances based on the methodology of the Human Development Index, that is, standardizing by the maximum difference of distances within each municipality, and then calculating the departmental average of the standardized values.

For the year 2020, there was a **growth in student dropout rates in educational** centers belonging to the **non-official sector**

Table 7 presents the general results of the distance metrics (Euclidean, Manhattan, and routing) between the homes and educational centers of children and adolescents between 5 and 17 years of age enrolled in educational centers in urban areas of the country in grades transition (0) to eleven (11), and **their relationship with student dropout:**

In 2018, the dropout rate goes from 1.7% to 3.1% between quintiles by distance 1 and 5 for official educational centers, while for non-official educational centers this rate decreases by 0.1 percentage points (p.p.) between quintile 1 and quintile 5 of distances.

This pattern is maintained for official educational centers in the years 2019 and 2020, the latter with a growth in the dropout rate between quintile 1 and quintile 5 of 0.7 p.p. Additionally, **in the year 2020 there was a significant growth in dropout rates for non-official educational centers (0.6 p. p. above the 2019 values).** However, this is not necessarily explained by the distance between the home and the educational center and may be related to the effect of the COVID-19 pandemic and the preventive confinement and isolation measures.

Table 7.

Relationship between deciles of the Euclidean distance metric and dropout rates for urban areas, by gender and school sector. 2018-2020. (Values in absolute numbers, meters (m) y percentages)

Euclidean distance quintiles (meters (m))			Dropout rate			Obs
			Women (%)	Men (%)	Total (%)	
2018	1 Mean: 237.2	Official	1.5	2.0	1.7	701,727
		No official	1.5	1.6	1.5	203,110
	2 Mean: 543.3	Official	1.6	2.2	1.9	714,091
		No official	1.3	1.4	1.4	180,651
	3 Mean: 954.3	Official	1.8	2.4	2.1	696,396
		No official	1.2	1.6	1.4	188,461
	4 Mean: 1.879.5	Official	1.9	2.6	2.2	602,642
		No official	1.2	1.5	1.4	262,590
	5 Mean: 10.561.3	Official	2.7	3.5	3.1	410,825
		No official	1.5	1.7	1.6	361,755
	Total	Official	1.8	2.4	2.1	3,125,681
		No official	1.4	1.6	1.5	1,196,567

Source: DANE 2022.

Table 7.

Relationship between deciles of the Euclidean distance metric and dropout rates for urban areas, by gender and school sector. 2018-2020. (Values in absolute numbers, meters (m) y percentages)

Euclidean distance quintiles (meters (m))			Dropout rate			Obs
			Women (%)	Men (%)	Total (%)	
2019	1 Mean: 237.5	Official	1.6	2.1	1.8	642,407
		No official	2.4	2.6	2.5	154,293
	2 Mean: 543.3	Official	1.8	2.3	2.0	650,816
		No official	2.0	2.2	2.1	145,438
	3 Mean:: 954.9	Official	1.9	2.6	2.2	635,486
		No official	1.8	2.1	2.0	159,505
	4 Mean: 1.883.0	Official	1.9	2.6	2.3	560,182
		No official	1.8	2.0	1.9	228,520
	5 Mean: 12.404.9	Official	2.5	3.1	2.8	448,949
		No official	1.9	2.0	2.0	321,488
	Total	Official	1.9	2.5	2.2	2.937,840
		No official	2.0	2.1	2.0	1.009,244

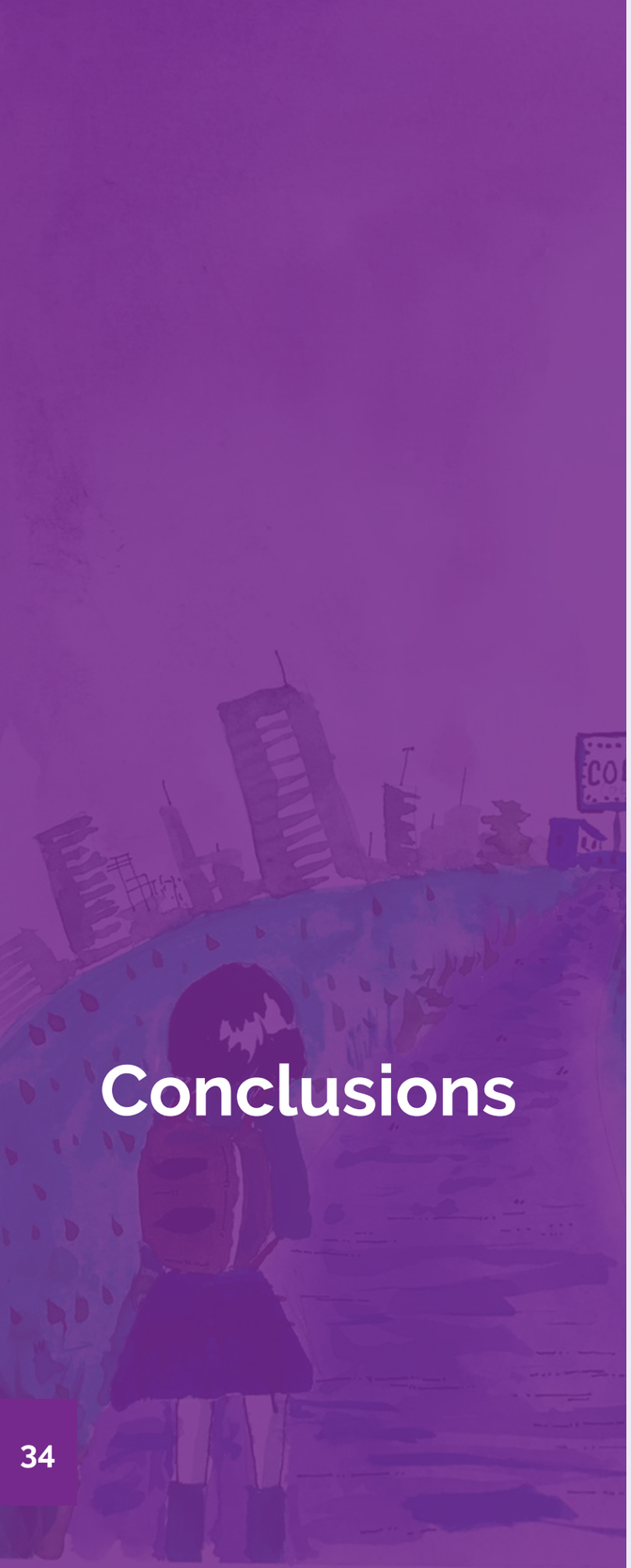
Source: DANE 2022.

Table 7.

Relationship between deciles of the Euclidean distance metric and dropout rates for urban areas, by gender and school sector. 2018-2020. (Values in absolute numbers, meters (m) y percentages)

Euclidean distance quintiles (meters (m))			Dropout rate			Obs
			Women (%)	Men (%)	Total (%)	
2020	1 Mean: 238.2	Official	1.2	1.5	1.4	571,548
		No official	2.9	3.4	3.2	167.797
	2 Mean: 544.0	Official	1.4	1.8	1.6	594,215
		No official	2.7	3.0	2.8	155,671
	3 Mean: 956.0	Official	1.5	1.9	1.7	594,141
		No official	2.4	2.7	2.5	166,893
	4 Mean: 1.892.3	Official	1.5	1.8	1.6	546,888
		No official	2.2	2.6	2.4	240,062
	5 Mean:14.457.1	Official	1.9	2.4	2.1	540,101
		No official	2.1	2.5	2.3	357,761
	Total	Official	1.5	1.9	1.7	2,846,893
		No official	2.4	2.7	2.6	1,088,184

Source: DANE 2022.



Conclusions

Educational coverage rates, defined as the percentage ratio between the number of students enrolled with respect to the population theoretically old enough to attend a given level, together with s**chool dropout rates**, are two of the main indicators that measure access to education supply in Colombia:

- In 2020, the gross coverage rate in education at the national level was 74.9% for the population between 5 and 17 years of age, in the educational levels of transition (grade 0) to grade eleven (11). For the reference year, this coverage was higher in urban areas than at the national level (83.1%), representing a decrease of 0.2 percentage points (p.p.) with respect to 2019.

- In terms of school dropout, between 2018 and 2020, 2.47% of students located in urban areas dropped out, and the difference between official and non-official schools ranged between 0.54 and 1.05 percentage points (p.p.) in favor of official schools, during 2018 and 2019. In 2020, this trend was reversed in urban areas, and dropout in non-official schools was 0.74 p.p. higher than in official schools. For the reference period (2018-2020), in terms of educational levels, the highest percentage of dropouts occurred in sixth grade (6) (14.0%).

In line with the need for increasingly accurate, timely, and georeferenced information, this Statistical Brief presents the results of the project to approximate **education accessibility metrics based on the analysis of distances and their correlation with dropout rates, for the years 2018, 2019 and 2020.**

This exercise consisted of an **integration process** between various sources of information from the education sector in Colombia, such as the Integrated Enrollment System (SIMAT, Spanish acronyms) and the Statistical Directory of Educational Centers (DIREDU, Spanish acronyms), together with the geolocation of the households of children and adolescents between 5 and 17 years of age who were enrolled in educational centers that provide traditional education in the country, which was obtained through the National Population and Housing Census (CNPV 2018, Spanish acronyms).

The three distance metrics are: Euclidean, Manhattan and routing, where the first two take as inputs for the spatial analysis the coordinates of the students' educational center and their homes. On the one hand, the Euclidean metric calculates the distance in a straight line between both points (proximity analysis), while the Manhattan metric corresponds to the sum of the absolute differences of their coordinates, being on average greater, but more "realistic", than the distance in a straight line. In addition to the coordinates, the routing distance incorporates the road grid element.

The results of the routing distance metric are presented only as a contrast since the percentage of crossover records for both rural and urban areas is less than 32.0% of the total number of students; while for Euclidean and Manhattan distances this percentage is higher at 66.0% and 63.5%, respectively.

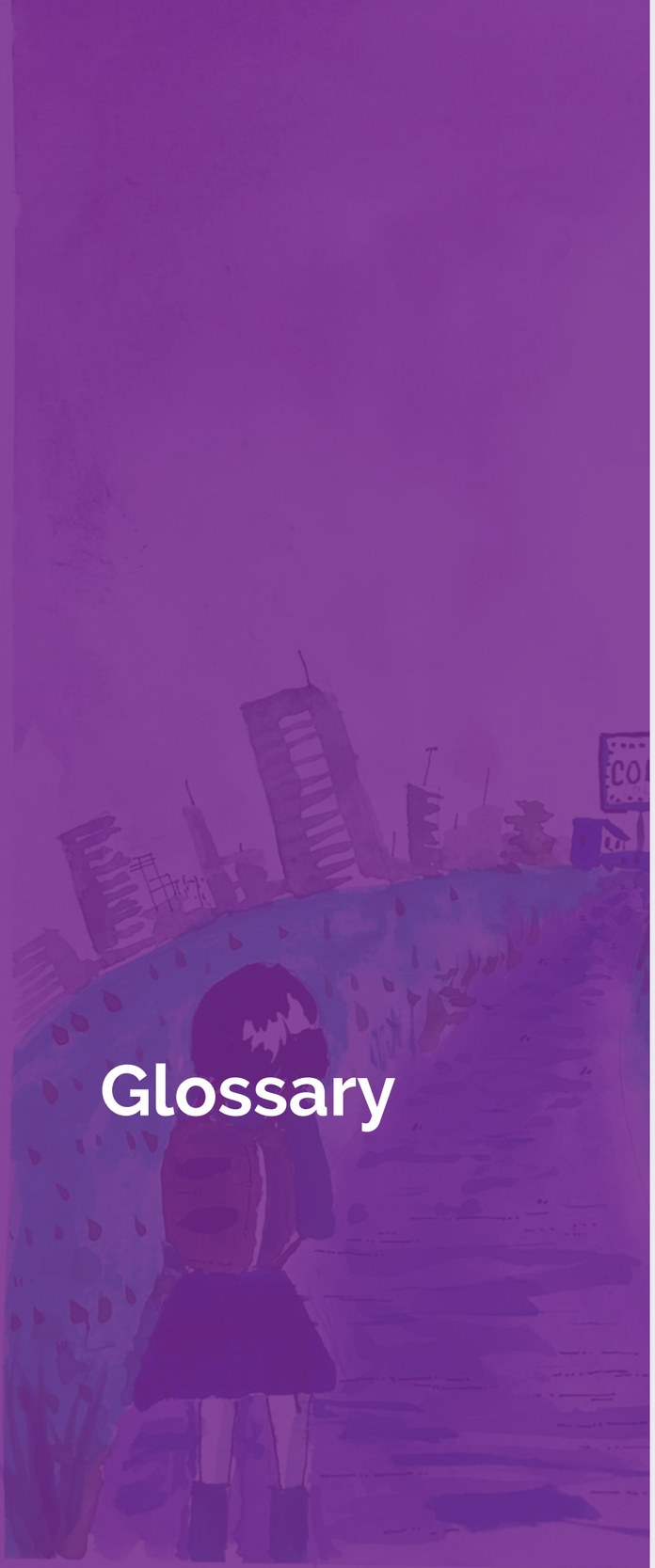
Given the implementation rates of the distance methodologies presented, the exercise focused on urban areas, where, for this period 2018-2020, the maximum percentage of crossover of the records obtained is 73.2% of the total number of students in the universe defined for the year 2018, under the Euclidean and Manhattan metrics, and a minimum of 63.5% for the year 2020 for both methodologies.

- **This distance is 2.6 kilometers (km) under the Euclidean metric for 2018, 3.2 km for 2019, and 4.0 km for 2020. The Manhattan distance presents an average of 3.3 km for 2018, 4.0 km for 2019, and 4.5 km for 2020,** with a standard deviation greater than that presented under the Euclidean methodology, lying in a range between 9.4 and 12.8 km during the period 2018-2020.

- When analyzing these approximations according to the sector to which the educational center belongs (official and non-official), it is observed that **the distances for students enrolled in official educational centers are shorter than the distances for students in non-official schools.**

- **In 2020, the five departments with the greatest distances under the Euclidean methodology were Vaupés, Amazonas, Chocó, Putumayo, and Meta,** with distances ranging between 6.3 and 11.9 kilometers on average for that year.

- **In 2018, the dropout rate increased from 1.7% to 3.1% between distance quintiles 1 and 5 for official educational centers, while for non-official educational centers this rate decreased by 0.1 percentage points (p.p.) between distance quintile 1 and 5.** Additionally, in 2020 there is a significant growth in dropout rates for non-official schools (0.6 p.p. above the figure for 2019).



Glossary

Academic secondary education: is the level of education that deepens in a specific field of science, arts or humanities, and allows the student to access higher education upon completion of secondary education (grades 10 to 11). In academic secondary education, the same areas of basic education are compulsory and fundamental at a more advanced level, in addition to economic and political sciences and philosophy (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Basic education: is the education process, which lasts nine grades. It is structured around a common curriculum made up of the fundamental areas of knowledge and human activity and whose development is carried out in two cycles: basic primary education in five grades and basic secondary education in four grades (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Basic primary education is the training process that lasts for five grades of basic elementary education. It comprises the first five grades of basic education: first, second, third, fourth and fifth (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Basic secondary education: is the training process that lasts for four grades of basic education instruction. It comprises the following four grades of basic education: sixth, seventh, eighth and ninth (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Classification of Higher Education Institutions – (IES, Spanish acronyms): according to their academic nature, Higher Education Institutions (IES, Spanish acronyms) are classified as follows: Professional Technical Institutions; Technological Institutions; University Institutions or Technological Schools; Universities.

Dispersed rural: geographic delimitation defined by DANE for statistical purposes, comprised between the census perimeter of the municipal capitals and population centers, and the municipal boundary. It is characterized by the dispersed arrangement of dwellings and agricultural and livestock farms existing in it. It corresponds to the territory that is neither part of the municipal capital (class 1) nor of the population centers (class 2). It is characterized by objects and elements related to agriculture, rest or recreation, mining or extractive uses. The number of residential units per area is lower than in urban zones (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational center: an educational establishment that does not offer a full course of education (at least one year of preschool education and at least nine grades of basic education and high school) and that must partner with other educational institutions or centers in order to offer the full cycle of basic education to students (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational establishments of the official sector: these are state-owned establishments, financed with public resources. Within these establishments are also those of special regime that are financed with public resources other than those of the General System of Participations (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational institution: a group of persons and assets promoted by public authorities or private individuals, whose purpose is to provide one year of preschool education and at least nine grades of basic education, and secondary education. Those that do not offer all these grades will be called educational centers and must be associated with other institutions in order to offer the complete cycle of basic education to students. They must have an operating license or official recognition, administrative infrastructure, pedagogical support, physical plant and adequate educational resources (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational level: refers to the consecutive stages of education that exist in the Colombian educational system. It includes preschool, elementary school, elementary school, middle school, undergraduate (professional technical, technological, university) and postgraduate (specialization, master's, doctorate) education (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational model: set of structured strategies to meet with quality and relevance the educational demands of a target population with specific characteristics. A model has explicit pedagogical and didactic principles, as well as its articulation pathways in an educational site. It also has a basket of educational materials and processes (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational service: includes the set of legal norms, curricular programs, education by levels and grades, non-formal education, informal education, educational establishments, social institutions (state or private) with educational, cultural and recreational functions, human, technological, methodological, material, administrative and financial resources, articulated in processes and structures to achieve the objectives of education (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Educational site: basic unit in the organization of the public education service. Strictly speaking, it is an economic unit that has been legally constituted for the provision of the public education service and for this purpose has an operating license (applicable to non-official sector establishments) or an act of recognition (for official sector educational centers), a physical plant and an administrative structure. The organization or association between sites may lead to the creation or recognition of educational establishments, institutions or centers (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Formal education: is the one provided in approved educational institutions, in a regular sequence of school cycles, subject to progressive curricular guidelines and leading to degrees and diplomas. It is divided into three levels: preschool, elementary and middle school (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Gross coverage rate in education: Corresponds to the percentage ratio between students enrolled in a specific level of education (regardless of their age) and the school population that is of the appropriate age to attend that level. It is feasible that in the calculation of this indicator, coverage results greater than 100% are obtained, because all or most of the population in this age group is covered by the education system and additionally, there are students enrolled in extra-curricular programs (https://www.mineducacion.gov.co/1621/article-123926.html).

Municipal capital: geographic delimitation defined by DANE for statistical purposes, alluding to the geographic area delimited by the census perimeter. Within it is located the administrative headquarters of the municipality, i.e. the mayor's office.

Net coverage rate in education: This is the ratio of students enrolled in an educational level who are of the appropriate age to attend it to the total population in the appropriate age range for that level. In the same way as the gross coverage indicator, it is possible to disaggregate the net coverage rate indicator by departments, zones (rural, urban), gender (men and women), and educational levels (https://www.mineducacion.gov.co/1621/article-123926.html).

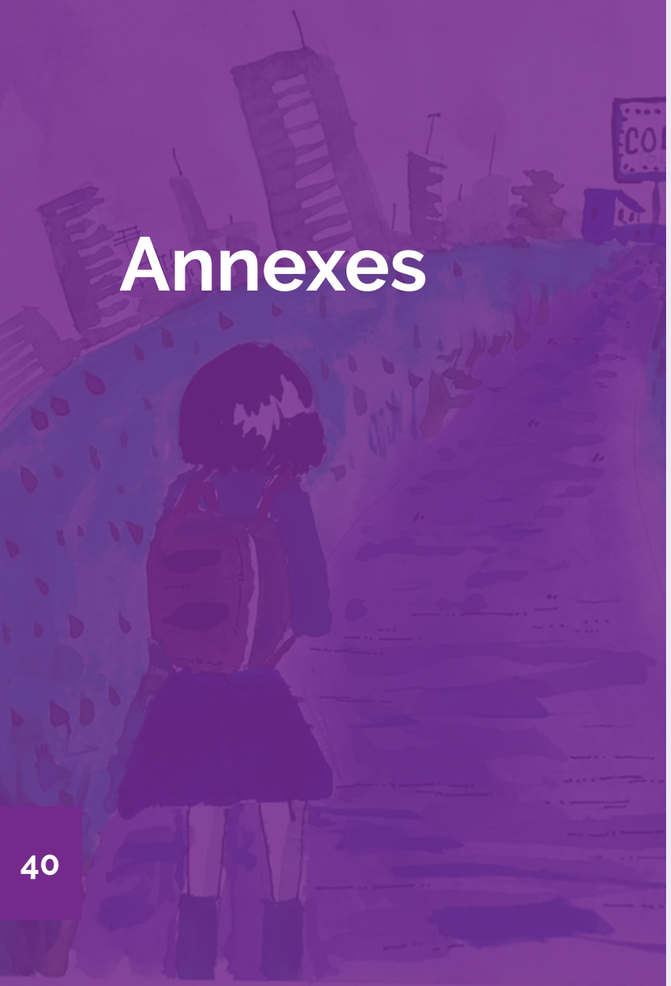
Non-official sector educational establishments: these are establishments under the administration and management of private (non-official) natural or legal persons, such as religious or confessional communities, cooperatives, foundations, federations or corporations, compensation funds, etc. According to the National Constitution, the government does not finance or subsidize private educational institutions, although it does regulate the provision of its services (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Population center: concept created by DANE for statistical purposes, for the identification and geographic location of population nuclei or settlements. It is defined as a minimum concentration of twenty contiguous dwellings, neighboring or attached to each other, located in the rest of the municipal area or in a non-municipalized area (departmental township). It includes the population centers of the municipal townships, police inspections and hamlets. This concentration has characteristics such as the delimitation of vehicular and pedestrian roads (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Public educational service: the one provided under the regulated scheme and modality so that both public and private operators may offer and provide the public education service. The departments, districts and certified municipalities shall provide the public education service through official educational institutions. They may, when the insufficiency of the State educational institutions is demonstrated, contract the provision of the service with State or non-State entities that provide educational services, of recognized trajectory and suitability, prior accreditation, with resources from the General System of Participations (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Secondary education: is the culmination, consolidation and progress in the achievement of the previous levels (preschool and elementary school), whose character is academic or technical, with a duration of two grades of educational instruction, the tenth (10th) and eleventh (11th). Its purpose is the understanding of ideas, the development of universal values and the preparation of the learner for entry into higher education and employment (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).

Urban sector: is a cartographic division created by DANE for statistical purposes, made up of urban sections. It is the largest division or maximum level of aggregation defined within the census perimeter of municipalities and population centers (https://conceptos.dane.gov.co/conceptos/conceptos_catalogo).



Annexes

Annex 1

Statistical Directory of Educational Centers

DANE's Statistical Directory of Educational Centers (DIREDU) includes a list of primary and secondary educational centers active in the national territory and incorporates periodically updated information on identification data, location, contact and classification by academic calendar and sector (official and non-official).

The sources of information that make up this registry are mainly secondary. On the one hand, there is the Educational Centers Information System (SISE) operated by DANE, which centralizes the process of assigning DANE codes and new developments in educational centers, at the request of each of the 96 Education Secretariats in the country. The SISE was conceived for the assignment of the unique code and identifier of educational centers. This code is a unique identification number made up of 12 digits, unchangeable and unrepeatable, which is created and assigned only once to legally constituted educational establishments [12].

The SISE interoperates with the Single Directory of Educational Establishments (DUE, Spanish acronyms), administered by the Ministry of National Education (MEN), in which each Ministry of Education can manage its educational establishments and report centralized news to the Ministry (MEN, n.d.-b).

On the other hand, DIREDU integrates information from the Integrated Enrollment System (SIMAT, Spanish acronyms), administered by the MEN, which is a tool that allows organizing and controlling the student enrollment process at all educational levels (MEN, n.d.-b). Finally, DANE, as part of its statistical operations, has a census of formal educational centers in the official and non-official sector (C-600), which compiles information on enrollment and teaching staff at the different levels provided by formal educational institutions or centers throughout the country. The completion of this form by the schools feeds the National Educational Information System (SIEF) (DANE, 2005).

[12] DANE Code is a mandatory requirement demanded within the process of recognition of the educational centers of the official sector and within the process of creation and licensing of the educational centers of the unofficial sector.

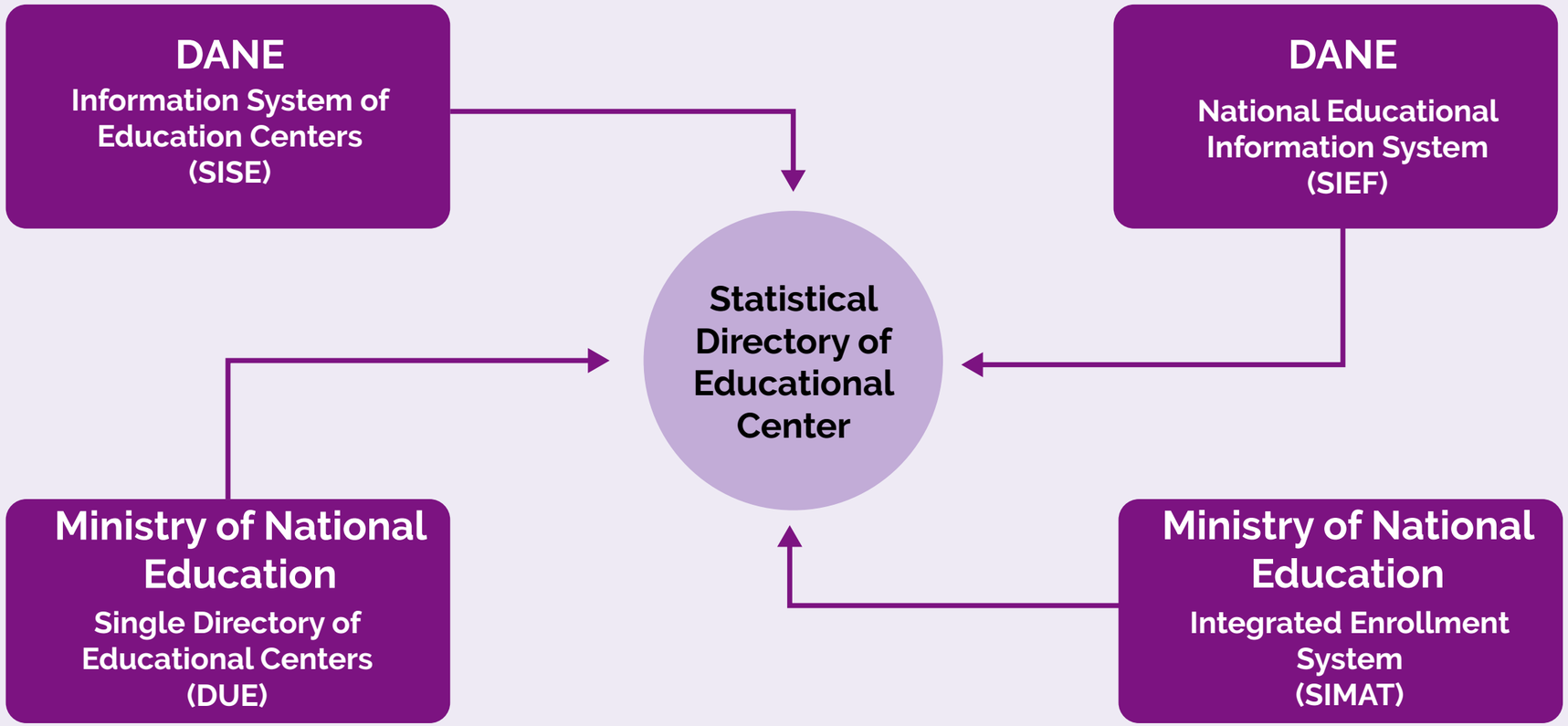
Together, these four sources of information, shown in Illustration 5, contain information on identification, location, contact and variables specific to the educational centers, such as calendar, sector, school days and number of students enrolled.

The master database on which DIREDU updates are made on a monthly basis is the SISE, which integrates on a daily basis the information of new sites created or the incorporation of new developments such as activations, temporary or definitive closures and changes in their basic information such as name, location and contact, among others. This information, which interoperates with the DUE, is used to maintain homologated and updated information in the DANE and the MEN simultaneously. The SIEF systems allow the capture of new information in the field and additional attributes in the case of SIMAT.

The DIREDU information has as its statistical unit the educational centers identified by a unique key, the DANE code. In this integration of information, cross validations are carried out in which the most recent information is taken from the educational centers, the location and contact information is standardized, and quality validations are carried out one by one in monthly exercises with the country's Secretariats of Education. Among the attributes of the location information are the coordinates of the educational centers, which have been strengthened with the support of the United Nations Children's Fund (UNICEF) through a mobile web application and artificial intelligence tools.

Illustration 5.

Conformation of the Statistical Directory of Educational Centers (DIREDU)



Source: DANE – DEST, DIG.

Annex 2.

Net coverage rate by Certified Territorial Entity (ETC Spanish acronyms) in preschool, elementary and middle school education. National total. 2020Pr. (Values in percentages

ETC	Total
Vaupés (ETC)	53.5%
Vichada (ETC)	60.2%
San Andrés de Tumaco	68.5%
Guainía (ETC)	70.6%
Jamundí	76.6%
Chocó (ETC)	77.3%
Valle del Cauca (ETC)	78.0%
Nariño (ETC)	79.0%
Tuluá	79.5%
Cartago	80.1%
Buenaventura	80.2%
Manizales	80.9%
Cauca (ETC)	81.0%
Amazonas (ETC)	81.2%
Palmira	81.5%
Soledad	82.4%
Guaviare (ETC)	83.1%
Guadalajara de Buga	83.2%
Caquetá (ETC)	84.0%
Cali	84.1%

ETC	Pre-school
Vaupés (ETC)	30.6%
San Andrés de Tumaco	30.9%
Guaviare (ETC)	43.3%
Vichada (ETC)	43.5%
Chocó (ETC)	46.8%
Buenaventura	47.5%
Nariño (ETC)	48.4%
Córdoba (ETC)	49.7%
Amazonas (ETC)	50.2%
Putumayo (ETC)	50.7%
Caquetá (ETC)	50.9%
Cauca (ETC)	51.1%
Jamundí	53.5%
Guadalajara de Buga	53.9%
Cesar (ETC)	55.0%
Magdalena (ETC)	55.2%
Malambo	55.5%
Cartago	56.2%
Guainía (ETC)	56.2%
Valle del Cauca (ETC)	57.3%

ETC	Primary
Vaupés (ETC)	55.0%
San Andrés de Tumaco	62.3%
Vichada (ETC)	64.7%
Guainía (ETC)	70.0%
Buenaventura	71.4%
Valle del Cauca (ETC)	73.5%
Jamundí	74.3%
Tuluá	75.0%
Nariño (ETC)	75.3%
Cartago	76.0%
Guaviare (ETC)	76.5%
Palmira	76.6%
Manizales	76.7%
Guadalajara de Buga	77.3%
Amazonas (ETC)	79.6%
Chocó (ETC)	79.7%
Cali	79.9%
Caldas (ETC)	80.0%
Putumayo (ETC)	80.6%
Soledad	81.7%

ETC	Middle School
Vichada (ETC)	30.6%
Vaupés (ETC)	30.9%
Uribia	43.3%
Guainía (ETC)	43.5%
Chocó (ETC)	46.8%
San Andrés de Tumaco	47.5%
La Guajira (ETC)	48.4%
Amazonas (ETC)	49.7%
Buenaventura	50.2%
Nariño (ETC)	50.7%
Caquetá (ETC)	50.9%
Cauca (ETC)	51.1%
Maicao	53.5%
Magdalena (ETC)	53.9%
Jamundí	55.0%
Guaviare (ETC)	55.2%
Malambo	55.5%
Soledad	56.2%
Cesar (ETC)	56.2%
Valle del Cauca (ETC)	57.3%

ETC	High school
Uribia	7.8%
Vichada (ETC)	9.9%
Guainía (ETC)	10.7%
Vaupés (ETC)	13.3%
Chocó (ETC)	20.7%
San Andrés de Tumaco	24.8%
La Guajira (ETC)	27.0%
Caquetá (ETC)	28.4%
Amazonas (ETC)	30.4%
Buenaventura	33.9%
Cauca (ETC)	34,1%
Guaviare (ETC)	34.3%
Norte de Santander (ETC)	35.3%
Maicao	35.5%
Cesar (ETC)	35.9%
Nariño (ETC)	36.9%
Magdalena (ETC)	37.2%
Arauca (ETC)	37.3%
Jamundí	38.3%
Ciénaga	38.7%

ETC	Total
Putumayo (ETC)	84.4%
Malambo	84.5%
Norte de Santander (ETC)	85.9%
Soacha	86.5%
Caldas (ETC)	86.9%
Bello	87.2%
Uribia	87.3%
Atlántico (ETC)	87.8%
La Guajira (ETC)	88.1%
Facatativá	88.7%
Dosquebradas	89.0%
Boyacá (ETC)	89.0%
Santander (ETC)	89.0%
Floridablanca	89.3%
Antioquia (ETC)	89.5%
Cesar (ETC)	89.7%
Valledupar	89.8%
Casanare (ETC)	89.9%
Huila (ETC)	90.3%
Córdoba (ETC)	90.4%
Risaralda (ETC)	90.4%
Itagüí	90.6%
Magangué	90.8%
Zipaquirá	91.0%

ETC	Pre-school
Palmira	57.7%
Soledad	57.7%
Arauca (ETC)	59.0%
Magangué	59.1%
Valledupar	59.2%
Santander (ETC)	59.2%
Huila (ETC)	59.6%
Cali	60.2%
Risaralda (ETC)	60.4%
Cundinamarca (ETC)	60.5%
Atlántico (ETC)	60.5%
Sahagún	60.6%
Turbo	60.7%
Antioquia (ETC)	60.8%
Bolívar (ETC)	61.0%
Montería	61.5%
Tuluá	61.5%
Norte de Santander (ETC)	62.3%
Caldas (ETC)	62.5%
Lorica	63.4%
Sucre (ETC)	63,4%
Boyacá (ETC)	63,5%
Tolima (ETC)	64.0%
Quindío (ETC)	64.1%

ETC	Primary
Cauca (ETC)	82.0%
Bello	83.2%
Caquetá (ETC)	83.6%
Itagüí	83.6%
Córdoba (ETC)	84.0%
Antioquia (ETC)	84.1%
Malambo	84.1%
Boyacá (ETC)	84.4%
Santander (ETC)	85.2%
Risaralda (ETC)	85.3%
Atlántico (ETC)	85.7%
Magdalena (ETC)	85.8%
Facatativá	86.0%
Soacha	86.1%
Valledupar	86.1%
La Guajira (ETC)	86.1%
Norte de Santander (ETC)	87.3%
Dosquebradas	87.3%
Cesar (ETC)	87.5%
Magangué	87.7%
Floridablanca	88.2%
Zipaquirá	88.7%
Casanare (ETC)	88.8%
Apartado	88.9%

ETC	Middle School
Norte de Santander (ETC)	57,7%
Bolívar (ETC)	57,7%
Atlántico (ETC)	59,0%
Cartago	59,1%
Sucre (ETC)	59,2%
Magangué	59,2%
Putumayo (ETC)	59,6%
Ciénaga	60,2%
Palmira	60,4%
Tuluá	60,5%
Arauca (ETC)	60,5%
Manizales	60,6%
Guadalajara de Buga	60,7%
Risaralda (ETC)	60,8%
Valledupar	61,0%
Cali	61,5%
Córdoba (ETC)	61,5%
Antioquia (ETC)	62,3%
Bello	62,5%
Caldas (ETC)	63,4%
Casanare (ETC)	63,4%
Meta (ETC)	63,5%
Soacha	64,0%
Santander (ETC)	64,1%

ETC	High school
Putumayo (ETC)	39,5%
Sucre (ETC)	39,8%
Malambo	40,2%
Magangué	40,4%
Riohacha	40,5%
Bolívar (ETC)	40,6%
Bello	41,5%
Florencia	41,8%
Palmira	42,3%
Valle del Cauca (ETC)	42,7%
Antioquia (ETC)	43,0%
Casanare (ETC)	43,1%
Córdoba (ETC)	43,1%
Cartago	43,2%
Meta (ETC)	43,4%
Pitalito	44,0%
Atlántico (ETC)	44,4%
Soledad	44,6%
Quibdó	44,7%
Soacha	45,2%
Girón	45,3%
Rionegro	45,6%
Dosquebradas	45,6%
Huila (ETC)	45,9%

ETC	Total
Magdalena (ETC)	91.7%
Bolívar (ETC)	92.1%
Sucre (ETC)	92.6%
San Andrés (ETC)	92.7%
Arauca (ETC)	92.8%
Mosquera	92.9%
Meta (ETC)	93.0%
Girón	94.3%
Quindío (ETC)	94.7%
Sahagún	94.7%
Villavicencio	94.8%
Cundinamarca (ETC)	95.1%
Tolima (ETC)	95.1%
Piedecuesta	95.2%
Ciénaga	95.3%
Cúcuta	95.4%
Fusagasugá	95.8%
Bogotá, D.C. (ETC)	95.9%
Maicao	96.0%
Yumbo	96.2%
Neiva	96.3%
Envigado	96.9%
Armenia	97.1%
Florencia	97.3%

ETC	Pre-school
Soacha	64.5%
Casanare (ETC)	64.7%
Manizales	64.7%
Meta (ETC)	65.3%
Villavicencio	66.0%
Floridablanca	66.0%
Zipaquirá	66.3%
Ciénaga	66.6%
Apartado	67.3%
Piedecuesta	67.7%
Girardot	67.9%
Facatativá	68.7%
Florencia	69.1%
Dosquebradas	69.1%
Bogotá, D.C. (ETC)	69.2%
Bello	69.6%
Fusagasugá	69.8%
Itagüí	69.9%
Santa Marta	70.4%
Envigado	70.9%
Barranquilla	71.3%
La Guajira (ETC)	71.4%
Yumbo	72.2%
Pitalito	72.3%

ETC	Primary
Sahagún	88.9%
Huila (ETC)	89.3%
Sucre (ETC)	89.7%
Yumbo	89.7%
Bolívar (ETC)	89.7%
Quindío (ETC)	90.0%
Meta (ETC)	90.5%
Cundinamarca (ETC)	90.5%
San Andrés (ETC)	90.7%
Arauca (ETC)	90.7%
Cúcuta	91.1%
Ciénaga	91.2%
Mosquera	91.2%
Villavicencio	91.6%
Tolima (ETC)	91.8%
Envigado	92.4%
Fusagasugá	92.5%
Armenia	92.7%
Sabaneta	93.5%
Bogotá, D.C. (ETC)	93.8%
Rionegro	94.0%
Piedecuesta	94.0%
Medellín	94,0%
Maicao	94,0%

ETC	Middle School
Huila (ETC)	64.5%
Florencia	64.7%
Floridablanca	64.7%
Facatativá	65.3%
Girón	66.0%
Tolima (ETC)	66.0%
Dosquebradas	66.3%
Riohacha	66.6%
San Andrés (ETC)	67.3%
Neiva	67.7%
Sahagún	67.9%
Quibdó	68.7%
Cúcuta	69.1%
Boyacá (ETC)	69,1%
Villavicencio	69,2%
Zipaquirá	69.6%
Montería	69.8%
Quindío (ETC)	69,9%
Santa Marta	70.4%
Pitalito	70.9%
Itagüí	71.3%
Fusagasugá	71.4%
Mosquera	72.2%
Cundinamarca (ETC)	72.3%

ETC	High school
Zipaquirá	46.8%
Cali	47.0%
Valledupar	47.4%
Risaralda (ETC)	47.9%
Neiva	47.9%
Cúcuta	48.2%
Montería	48.5%
Manizales	48.7%
Tuluá	48.8%
Santander (ETC)	49.0%
Tolima (ETC)	49.0%
Caldas (ETC)	49.3%
Turbo	49.4%
Piedecuesta	49.6%
Guadalajara de Buga	49.6%
Mosquera	49.7%
Apartado	50.0%
Floridablanca	50.7%
Facatativá	50.7%
Sincelejo	50.9%
San Andrés (ETC)	50.9%
Cundinamarca (ETC)	51.0%
Villavicencio	51.1%
Yumbo	51.2%

ETC	Total
Montería	97.6%
Medellín	98,2%
Apartado	98.3%
Rionegro	99.4%
Sabaneta	100.1%
Funza	101.2%
Duitama	101.4%
Pitalito	101.5%
Sincelejo	101.5%
Bucaramanga	101.6%
Yopal	101.7%
Barranquilla	101.9%
Santa Marta	103.0%
Girardot	103.1%
Lorica	103.1%
Popayán	103.9%
Tunja	105.7%
Pasto	105.7%
Ibagué	106.6%
Turbo	107.4%
Quibdó	107.5%
Cartagena	107.8%
Ipiales	108.3%
Chía	108.7%

ETC	Pre-school
Cúcuta	72.6%
Maicao	72.9%
Ipiales	72.9%
Girón	73.0%
Cartagena	73.2%
Chía	73.2%
Neiva	73.6%
Popayán	73.7%
Sincelejo	74.0%
Quibdó	74.4%
Sogamoso	76.0%
Bucaramanga	76.2%
Mosquera	76.2%
Pereira	77.1%
San Andrés (ETC)	77.4%
Armenia	78.3%
Sabaneta	78.4%
Funza	78.5%
Duitama	78.9%
Barrancabermeja	79.9%
Uribia	81.4%
Yopal	82.2%
Ibagué	83.1%
Medellín	83.7%

ETC	Primary
Montería	94.2%
Girón	94.3%
Lorica	94.4%
Neiva	94.7%
Florencia	95.9%
Uribia	96.5%
Barranquilla	96.8%
Funza	97.2%
Duitama	97.2%
Girardot	98.0%
Pasto	98.1%
Santa Marta	98.2%
Turbo	99.2%
Yopal	99.5%
Bucaramanga	99.9%
Pitalito	100.1%
Sincelejo	100.5%
Popayán	102.6%
Pereira	104.5%
Tunja	104.5%
Ipiales	104.9%
Cartagena	105.8%
Sogamoso	106.1%
Chía	106.2%

ETC	Middle School
Piedecuesta	72.6%
Yumbo	72.9%
Sincelejo	72.9%
Rionegro	73.0%
Bogotá, D.C. (ETC)	73.2%
Barranquilla	73.2%
Medellín	73.6%
Armenia	73.7%
Apartado	74.0%
Envigado	74.4%
Yopal	76.0%
Bucaramanga	76.2%
Lorica	76.2%
Popayán	77.1%
Girardot	77.4%
Ibagué	78.3%
Cartagena	78.4%
Ipiales	78.5%
Turbo	78.9%
Duitama	79.9%
Funza	81.4%
Chía	82.2%
Pereira	83,1%
Tunja	83.7%

ETC	High school
Sahagún	52.0%
Santa Marta	52.2%
Medellín	52.2%
Fusagasugá	52.2%
Quindío (ETC)	52.3%
Ipiales	52.9%
Yopal	53.5%
Itagüí	53.6%
Tunja	55.2%
Ibagué	55.4%
Bucaramanga	55.6%
Girardot	55.9%
Barrancabermeja	55.9%
Boyacá (ETC)	56.2%
Cartagena	56.9%
Barranquilla	57.1%
Bogotá, D.C. (ETC)	57.3%
Pereira	58.1%
Envigado	58.2%
Popayán	59.5%
Armenia	60.1%
Funza	60.5%
Sabaneta	62.6%
Pasto	62.8%

ETC	Total
Pereira	108.9%
Riohacha	111.9%
Sogamoso	113.2%
Barrancabermeja	117.9%

ETC	Pre-school
Tunja	84.7%
Rionegro	89.0%
Riohacha	92.0%
Pasto	93.3%

ETC	Primary
Ibagué	106.6%
Quibdó	107.1%
Riohacha	107.8%
Barrancabermeja	114.9%

ETC	Middle School
Sabaneta	84.7%
Pasto	89,0%
Barrancabermeja	92.0%
Sogamoso	93.3%

ETC	High school
Chía	63.0%
Lorica	64.7%
Duitama	65.4%
Sogamoso	73.4%

Source: MEN-SIMAT. DANE, Proyecciones de población – CNPV 2018

Annex 3.

Educational establishments by status (old-active, permanently closed, temporarily closed, new-active, not reported (NR)). National total. 2021. (Values in absolute values and percentages))

Department	Old - active		Definitive closure		Temporary closure		New - active		NR	
	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)
Amazonas	42	29.2%	3	2.1%		0.7%	98	68.1%	-	0.0%
Antioquia	4,357	61.7%	600	8.5%	371	5.3%	1,724	24.4%	6	0.1%
Arauca	331	50.8%	15	2.3%	119	18.3%	187	28.7%	-	0.0%
Archipiélago de San Andrés, Providencia y Santa Catalina	27	55.1%	13	26.5%	3	6.1%	6	12.2%	-	0.0%
Atlántico	957	50.1%	476	24.9%	15	0.8%	455	23.8%	6	0.3%
Bogotá D.C.	2,617	74.1%	596	16.9%	267	7.6%	52	1.5%	1	0.0%
Bolívar	1,411	59.1%	213	8.9%	232	9.7%	523	21.9%	7	0.3%
Boyacá	2,376	84.2%	92	3.3%	230	8.1%	121	4.3%	4	0.1%
Caldas	1,107	68.5%	179	11.1%	133	8.2%	191	11.8%	7	0.4%
Caquetá	1,056	62.3%	104	6.1%	148	8.7%	385	22.7%	2	0.1%
Casanare	521	63.2%	26	3.2%	135	16.4%	137	16.6%	5	0.6%
Cauca	2,347	78.7%	40	1.3%	110	3.7%	483	16.2%	2	0.1%
Cesar	685	35.8%	142	7.4%	75	3.9%	1,006	52.6%	3	0.2%
Chocó	678	45.6%	53	3.6%	175	11.8%	579	38.9%	2	0.1%
Córdoba	1,653	65.7%	164	6.5%	180	7.2%	477	19.0%	42	1.7%

Department	Old - active		Definitive closure		Temporary closure		New - active		NR	
	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)	Educational centers	(%)
Cundinamarca	2,835	67.0%	212	5,0%	231	5.5%	939	22,2%	15	0,4%
Guainía	25	25.5%	-	0.0%	-	0.0%	73	74.5%	-	0.0%
Guaviare	159	50.6%	31	9.9%	31	9.9%	93	29.6%	-	0.0%
Huila	1,657	76.0%	170	7.8%	6	0.3%	344	15.8%	3	0.1%
La Guajira	372	21.1%	47	2.7%	80	4.5%	1.263	71.5%	4	0.2%
Magdalena	1,124	59.2%	187	9.9%	122	6.4%	459	24.2%	6	0.3%
Meta	966	58.0%	70	4.2%	176	10.6%	452	27.1%	2	0.1%
Nariño	1,641	51.3%	83	2.6%	110	3.4%	1,354	42.3%	11	0.3%
Norte de Santander	1,793	65.8%	84	3.1%	141	5.2%	704	25.8%	2	0.1%
Putumayo	558	49.0%	23	2.0%	153	13.4%	391	34.3%	14	1.2%
Quindío	261	57.0%	46	10.0%	7	1.5%	143	31.2%	1	0.2%
Risaralda	530	42.8%	93	7.5%	18	1.5%	591	47.7%	6	0.5%
Santander	2,658	72.2%	172	4.7%	221	6.0%	622	16.9%	8	0.2%
Sucre	568	45.7%	134	10.8%	54	4.3%	486	39.1%	1	0.1%
Tolima	1,910	69.2%	127	4.6%	205	7.4%	516	18.7%	4	0.1%
Valle del Cauca	3.174	64.5%	522	10.6%	330	6.7%	856	17.4%	36	0.7%
Vaupés	32	22.1%	4	2.8%	11	7.6%	96	66.2%	2	1.4%
Vichada	41	14.4%	9	3.2%	64	22.5%	169	59.3%	2	0.7%
Total general	**40,469**		**4,730**		**4,154**		**15,975**		**204**	

Source: MEN-DUE.

References

Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., Suades-González, E., Donaire-Gonzalez, D., Cirach, M., de Castro, M., Esnaola, M., Basagaña, X., Dadvand, P., Nieuwenhuijsen, M., & Sunyer, J. (2017). Impact of commuting exposure to traffic-related air pollution on cognitive development in children walking to school. Environmental Pollution, 231, 837–844. https://doi.org/10.1016/j.envpol.2017.08.075

Andersson, E., Malmberg, B., & Östh, J. (2012). Travel-to-school distances in Sweden 2000–2006: Changing school geography with equality implications. Journal of Transport Geography, 23, 35–43. https://doi.org/10.1016/j.jtrangeo.2012.03.022

Boerman, W., Musterd, S., Pacchi, C. & Ranci, C. (2019). School segregation in contemporary cities: Socio-spatial dynamics, institutional context and urban outcomes. Urban Studies; 56(15):3055-3073.

Chica-Olmo, J., Rodríguez-López, C., & Chillón, P. (2018). Effect of distance from home to school and spatial dependence between homes on mode of commuting to school. Journal of Transport Geography, 72, 1–12. https://doi.org/10.1016/j.jtrangeo.2018.07.013

Chillón, P., Herrador-Colmenero, M., Migueles, J. H., Cabanas-Sánchez, V., Fernández-Santos, J. R., Veiga, Ó. L., Castro-Piñero, J., Marcos, A., Marcos, A., Veiga, O. L., Castro-Piñero, J., Bandrés, F., Martínez-Gómez, D., Ruiz, J. R., Carbonell-Baeza, A., Gomez-Martinez, S., Santiago, C., Marcos, A., Gómez-Martínez, S., … Gómez-Gallego, F. (2017). Convergent validation of a questionnaire to assess the mode and frequency of commuting to and from school. Scandinavian Journal of Public Health, 45(6), 612–620. https://doi.org/10.1177/1403494817718905

Congreso de la República (1994). Ley 115 de Febrero 8 de 1994. "Por la cual se expide la Ley General de Educación".

Cortés, F. (2012). El derecho a la educación como derecho social fundamental en sus tres dimensiones: educación primaria, secundaria y superior. Revista Estudios Socio-Jurídicos, pp. 185-205. En línea. https://revistas.urosario.edu.co/index.php/sociojuridicos/article/view/2067

DANE (2005). Manual de Diligenciamiento Educación Formal – EDUC Formulario C600 A Institución Educativa (Sector Oficial). En línea. https://www.dane.gov.co/files/investigaciones/educacion/formularios/Manual_C600A_Oficial.pdf

DANE (2018). COLOMBIA – Censo Nacional de Población y Vivienda – CNPV – 2018. Dirección de Censos y Demografía – CDD.

DANE (2020). Sistema de Identificación de Sedes Educativas – SISE. [Archivo de video]. YouTube. https://www.youtube.com/watch?v=EBrBzjNL8wU.

DANE (2022). Educación formal. En línea. https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/poblacion-escolarizada/educacion-formal

Ejaz, Z., Naz, U., & Khan, N. (2019). Determinants of Dropout and Child School Enrollment: A Case Study from Rural Islamabad. Journal of Quantitative Methods; Vol 3 No 2 (2019).

Fjellman, A.-M., Yang Hansen, K., & Beach, D. (2019). School choice and implications for equity: The new political geography of the Swedish upper secondary school market. Educational Review, 71(4), 518–539. https://doi.org/10.1080/00131911.2018.1457009

Frankenberg, E. (2013). The Role of Residential Segregation in Contemporary School Segregation. Education and Urban Society. ;45(5):548-570.

Gamboa, L., M. Rodríguez-Acosta, & A. García-Suaza (2010). Academic achievement in sciences: the role of preferences and educative assets.

Guarín, A., Medina, C., & Posso, C. (2018). Calidad, cobertura y costos ocultos de la educación secundaria pública y privada en Colombia. Desarrollo y Sociedad, 81, 61–114. En línea. https://doi.org/10.13043/dys.81.2

Hamnett, C. & Butler, T. (2013. Distance, education and inequality. Comparative Education, 49:3, 317-330.

Hanushek, E. A., & Woessmann, L. (2007). The Role of Education Quality for Economic Growth.

Jang, M., & Kang, C.-D. (2015). Retail accessibility and proximity effects on housing prices in Seoul, Korea: A retail type and housing submarket approach. Habitat International, 49, 516–528. https://doi.org/10.1016/j.habitatint.2015.07.004

Mantovani, D., Gasperoni, G., & Santangelo, F. (2021). Home-school distance among native and immigrant-origin lower secondary students in urban Northern Italy. Journal of Ethnic and Migration Studies, 1–27. https://doi.org/10.1080/1369183X.2021.1935659

Ministerio de Educación Nacional – MEN (s.f.-a). Deserción escolar. Disponible en: https://www.mineducacion.gov.co/1621/article-82745.html

Ministerio de Educación Nacional – MEN (s.f.-b). Directorio Único de Establecimientos Educativos – DUE. Disponible en: https://www.mineducacion.gov.co/sistemasdeinformacion/1735/w3-article-296670.html?_noredirect=1

Ministerio de Educación Nacional – MEN (s.f.-c). Glosario – Directorio de Establecimiento Educativo. Disponible en: https://www.mineducacion.gov.co/sistemasdeinformacion/1735/w3-article-296906.html?_noredirect=1

Ministerio de Educación Nacional – MEN (2018). Sistema educativo colombiano. Disponible en: https://www.mineducacion.gov.co/portal/Preescolar-basica-y-media/

Ministerio de Educación Nacional – MEN (2021). Sistema Integrado de Matrícula. Disponible en: https://www.mineducacion.gov.co/portal/micrositios-institucionales/Sistemas-de-Informacion/168883:SIMAT

Ndou, Adam & Tsaurai, Kunofiwa. (2019). Infrastructure, Human Capital Development and Economic Growth in Transitional Countries. Comparative Economic Research, 22(1), 33–52.

OECD (2021), Education at a Glance 2021: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/b35a14e5-en.

Rodríguez Lesmes, P. A., Trujillo Suárez, J. D., & Valderrama González, D. (2013). Más allá de la infraestructura: el impacto de las bibliotecas públicas en la calidad de la educación. Universidad del Rosario.

Shahidul, S., Zehadul Karim, A. (2015). Factors contributing to school dropout among the girls: A review of literature. European Journal of Research and Reflection in Educational Sciences, 3(2), 25–36.

Shehu, H. K. (2018). Factors Influencing Primary School Non-attendance among Children in North West Nigeria. Literacy Information and Computer Education Journal, 9(2), 2916–2922. https://doi.org/10.20533/licej.2040.2589.2018.0383

Sistema Integrado de Matrícula - SIMAT. (s.f.). Ministerio de Educación Nacional. En línea. Consultado el 9 de diciembre de 2021. Disponible en: https://www.mineducacion.gov.co/sistemasinfo/sistemapbm/Informacion-Institucional/297802:Que-es-SIMAT

Sistema de Información de Sedes Educativas. (2020). Departamento Administrativo Nacional de Estadística. En línea. Disponible en: https://www.youtube.com/watch?v=EBrBzjNL8wU

www.dane.gov.co

@DANE_Colombia /DANEColombia /DANEColombia DANE_Colombia