

RESEARCH ARTICLE

Reasons Why LGBTI People in Bogotá Drop Out of School

Factores que inciden en la deserción estudiantil de las personas de los sectores LGBTI en Bogotá

Motivos de evasão escolar da população LGBTI no Bogotá

SANDRA PATRICIA BARRAGÁN*  LEANDRO GONZÁLEZ** 

* Degree in Mathematics, Master in Mathematical Sciences, PhD in Modeling for Policy and Public Management. Foundations and Didactics of Science Group, Universidad Jorge Tadeo Lozano. Bogotá, Colombia. <http://orcid.org/0000-0001-6503-4445>

** Bachelor's Degree in Mathematics, Specialist in Mathematics Teaching, Master's Degree in Modeling and Simulation, Master's Degree in Statistics. Universidad Jorge Tadeo Lozano. Bogotá, Colombia. Leandro.gonzalez@utadeo.edu.co <https://orcid.org/0000-0002-9870-2312>

OPEN ACCESS 

DOI: <http://dx.doi.org/10.18634/sophiaj.17v.1i.994>

Article information

Received: February 24, 2020
Revised: August 16, 2020
Accepted: January 28, 2021
Published: April 26, 2021

Keywords: sample survey; demographics; dropping out of school; statistical model.

Palabras clave: encuesta por muestreo; demografía; deserción escolar; modelo estadístico.

Palavras-chave: demografia; abandono; Inquérito por amostragem; modelos estatísticos.

How to cite:

Barragán, S.P., González, L. (2021) Reasons Why LGBTI People in Bogotá Drop Out of School. *Sophia* 17 (1), e994

Sophia-Education, volume 17, number 2. January/June 2021. English version

ABSTRACT

The objective of this paper is to establish the academic traits and the reasons why people from LGBTI sectors in Bogotá have abandoned their studies based on the Multipurpose Survey of the District Planning Secretariat of Bogotá and the National Department of Statistics. The complete sample of the survey is 61,725 people representing a universe of 7,794,463 citizens in urban areas. The questions on gender and sexual orientation were applied to people over 18 years of age, so the sample for this work is 172,900 people with an expansion of 6,127,120. With this sample we analyze only the questions that provide information on the traits and academic risk profiles of people from the LGBTI sectors using decision trees as the main technique. It is found that for this population group: the literacy rate is 98.8%; the most influential variables, in descending order, when deciding whether or not to study are: age, type of housing tenure, marital status and sex, while the reasons for not studying are considered to be those who have already finished studying, educational costs and the need to work. In addition, men and women from LGBTI sectors, with similar intensity, remain in the educational system, while among heterosexuals it is women who remain in higher percentages.

Copyright 2021 Universidad La Gran Colombia



Conflict of interest:

The authors declare that they have no conflict of interest.

Author correspondence:

sandra.barragan@utadeo.edu.co
leandro.gonzalez@utadeo.edu.co

RESUMO

O objetivo deste documento é estabelecer as características acadêmicas e as razões pelas quais as «pessoas das comunidades LGBTI» de Bogotá abandonaram seus estudos com base na Pesquisa Multiuso da Secretaria de Planejamento Distrital de Bogotá e do Departamento Nacional de Estatística. A amostra completa da pesquisa é de 61.725 pessoas, representando um universo de 7.794.463 cidadãos de Bogotá das áreas urbanas. As questões de gênero e orientação sexual foram aplicadas a pessoas maiores de 18 anos, portanto a amostra para este trabalho é de 172.900 pessoas com uma expansão de 6.127.120. Com esta amostra, apenas são analisadas as questões que fornecem informação sobre as características e perfis de risco acadêmico de «pessoas das comunidades LGBTI», tendo como técnica principal as árvores de decisão. Verifica-se que, para este grupo populacional: a taxa de alfabetização é de 98,8%; As variáveis que mais influenciam, em ordem decrescente, na decisão de estudar ou não são: idade, tipo de casa própria, estado civil e sexo, enquanto os motivos para deixar de estudar são o fato de ter concluído os estudos, os custos educacionais e a necessidade de trabalhos. Além disso, homens e mulheres dos setores LGBTI, com intensidade semelhante, permanecem no sistema educacional, enquanto nos heterossexuais são as mulheres que permanecem no maior percentual.

Introduction

In Colombia, the law and the Political Constitution (1991) ensure that the right to education enables all people to develop in their life cycle all their dimensions: intellectual, physical, affective, artistic, psychic, social, artistic, spiritual, political and their interaction with the environment (Congress of the Republic of Colombia, 1994), (Secretaría Distrital de Planeación, 2015). Therefore, in terms of public policies, it is stated that the Colombian State will guarantee access and permanence, without any type of discrimination, to formal education under conditions of equality, equity and obligatory nature (Secretaría Distrital de Planeación, 2015: 63), (Secretaría de Integración Social, 2009: 49) and therefore, the State considers intolerable the exclusionary education, that freedoms are transgressed or that people cannot remain in education due to economic situations, discrimination, condition, expression or gender identity (Secretaría Distrital de Planeación, 2015: 63), (Secretaría de integración social, 2009: 52), (Congreso de la República de Colombia, 1994: 4). Particularly, it is important to analyze the progress in guaranteeing the right to education of people from LGBTI sectors (Secretaría Distrital de Planeación, 2015) from different perspectives, including school permanence and dropout, as well as the reasons why people from LGBTI sectors abandon their studies. All this by virtue of the fact that the exercise and guarantee of citizens' rights benefits society in general society in general.

Thus, observing both the constitutional parameter and public policies, the research question arose: what are the academic traits and the reasons that people from the LGBTI sectors in Bogotá state for dropping out of school? Hence, the objective of this paper is to analyze the general academic traits of people from LGBTI sectors in Bogotá, including the reasons why people from these sectors dropped out of school. The Multipurpose Survey (MS) database of the District Planning Secretariat (SDP) and the National Statistics Department (DANE) was used to carry out the objective. This sample survey was chosen for the present study considering its coverage, the improvements it has had between applications and the availability of its most recent application, all of which provide an important amount of updated information on Bogotá. There is no large survey in Colombia specifically aimed at analyzing the living conditions of LGBTI people, there are only some small surveys, case studies or with limited techniques that do not allow generalizations. Econometrics in 2010 and the Planning Secretariat in 2013 conducted surveys of social representations. These are not probabilistic, so their representativeness is debatable, since, although the results obtained are interesting as a reference point of the knowledge of the people of the LGBTI sectors, they are only an approach to their reality.

In the section following the introduction, reference will be made to the categories of Colombian public policy on sex, gender, gender identities and sexual orientation, since the questions and items of the MS were formulated with these categories in mind.

For the writing of this article, it was decided that, in order to make reference to plurals, the clarification of the Real Academia de la Lengua Española (RAE) will be followed, and therefore no splitting will be made. The RAE states that, linguistically, the generic use of the masculine is applied to all individuals of the species, without distinction of sex (Real Academia de la Lengua Española, 2005). Consequently, *bogotanos* and *bogotanas* or "*bogotanes*" will not be used, only "people from Bogota", without detriment to Spanish inclusive language. To refer to LGBTI people from Bogotá, the expression "people from LGBTI sectors" will be used in consideration of the terminology used by public policy in Bogotá (Secretaría Distrital de Planeación, 2015: 8).

Definitions of the Categories of Sexual Diversity in District Public Policies

The Technical Secretariat for Sexual Diversity of the SDP has stipulated the categories for sex, gender, gender identity and sexual orientation. It may be that this categorization is debatable in other contexts, but the theoretical explanation of such labels and how they relate to MS reagents is beyond the scope of the proposed objective of this article. The categories proposed by the aforementioned secretariat are:

1) Sex for the cultural idea of man as a person with male primary and secondary sex characteristics; and woman with female primary and sexual characteristics (Technical Secretary Directorate of Sexual Diversity, 2015: 24).

2) Gender for the psychological and social interpretation that organizes and determines the roles of the sex category. Gender refers to the masculine and the feminine (Technical Secretary Directorate of Sexual Diversity, 2015: 23-24).

3) Gender identity defines the attributes of individuals and collectives through subjective, intersubjective and sociocultural constructions. Here we find transgender people and people who are situated between the masculine and the feminine, including transformists (people who eventually use male and female clothing, gestures and roles in different environments), transvestites (people who permanently assume the attributes of the opposite gender to the one socially assigned to them), transsexuals (people who adopt a gender different from the one socially assigned to them) and trans (people build a political identity based on the experience of transit between sexes and gender) (Secretaría Técnica Dirección de Diversidad Sexual, 2015: 26, 51-52).

4) Sexual orientation groups the identities of people according to the erotic and affective desire according to their sex. The identities are: Homosexual (lesbian and gay as political categories), heterosexual and bisexual (Secretaría Técnica Dirección de Diversidad Sexual, 2015: 25).

In the context of the categories described above, the MS for Bogotá was addressed to establish the general academic features of Bogotá from the LGBTI sectors.

Below are the most relevant features, for this article, of MS and its applications.

The Multipurpose Survey

Decree 16 of 1993 "Whereby the internal structure of the District Planning Secretariat is adopted and other provisions are issued" indicates that the SDP must guide both the formulation and the "monitoring of territorial, economic, social and environmental policies and planning", in harmony with the other sectors of the Capital District (Alcaldía Mayor de Bogotá D.C., 1993) as part of its functions at the central level of the organizational chart of the administrative structure of Bogotá as a Capital District (Colombia. Presidency of the Republic, 1993), (Alcaldía Mayor de Bogotá D.C., 1999).

Therefore, the SDP has funded the 2011 (MS2011), 2014 (MS2014) and 2017 (MS2017) versions of the MS in which the field operator has been the National Statistics Department (DANE, 2018), (Secretaría Distrital de Planeación, 2015). The MS contributes to compile, provide and consolidate information, statistics, models and economic, social, cultural, environmental, territorial, productivity and competitiveness indicators for decision-making by the District Administration (Alcaldía Mayor de Bogotá D.C., 1993:1).

DANE and SDP allow access to anonymized microdata¹, forms, sample design documents, reports and bulletins for the 2011, 2014 and 2017 applications to anyone interested. This article was made by consulting this data.

In the MS2011 the interviewer recorded the sex variable for the persons that made up the household with two response options: man and woman (Secretaría Distrital de Planeación, 2012: 49. On the other hand, in the composition of the household of the MS2014 the question sex was contemplated, with three reagents: man, woman and intersex. This last option was formulated only to people over 15 years of age (Secretaría Distrital de Planeación, 2015). With the consideration of these questions and in accordance with the objective of this work, it was chosen to analyze only the MS2017, because as will be seen in the following section it has specific questions on the aforementioned categories of sexual diversity.

Multipurpose Survey application of 2017

The MS2017 was applied in the 20 localities and in the urban area of 37 municipalities of Cundinamarca². The 362 questions dealt, in general, with the same topics as the MS2014 (Departamento Nacional de Estadística, 2017).

In this case, the improvement consisted in the inclusion of two questions on household composition and demographics directed to persons 18 years of age or older. These were Number 26 (What is the sexual

1. Access on http://microdatos.dane.gov.co/index.php/catalog/MICRODATOS/about_collection/15/1S/about_collection/15/1

2. Cundinamarca is a department of the Andean region of Colombia whose capital is Bogotá.

orientation of...?), with reagents: heterosexual, homosexual, and bisexual; and Number 27 (With what gender do you identify...?), with reagents: female, male and transgender (Secretaría Distrital de Planeación, 2018: 13), (Secretaría Distrital de Planeación, 2017). THE MS2017 provides a great opportunity to approach the research question that guides this study.

Materials and Methods

Information for MS2017 was collected between September 1 and November 30, 2017 (Departamento Nacional de Estadística, 2017: 48). The universe was made up of "private households and the non-institutional civilian population" residing in: 1) Bogotá with 73 Zonal Planning Units (UPZ) and 17 UPZ groups; (2) Eight rural areas of Bogotá; and 3) 37 municipal capitals of Cundinamarca (Departamento Nacional de Estadística, 2017: 43). DANE indicated that the design of the survey was probabilistic, multistage, stratified and conglomerate with respect to the UPZs of the Capital District and the municipalities of the department of Cundinamarca (Departamento Nacional de Estadística, 2017: 47-50).

For this article, the questions and answers were processed considering the variables and determinants associated with student dropout: individual, institutional, socioeconomic and academic in the definitions of the Ministry of Education of Colombia (MEN) (Ministerio de Educación Nacional, 2009: 17). In addition, the variable marital status of the individual determinant was included in a special way, since as it appears in Barragán & González in the MS2014 located in Bogotá. Being single was a very influential variable in the permanence in the formal educational system (Barragán & González: p. 27).

To explore the relationship between the decision to study or not (variable of interest) it was decided to use the explanatory variables (independent): age, marital status and sex of the individual determinant, together with the socioeconomic stratum and the type of housing tenure, understood as the relationship that residents have with the ownership of the house, of the socioeconomic determinant. For the analysis between the aforementioned variables of interest and independent variables, decision trees were used to model statistically (Tan, Steinbach, & Kumar, 2006). This modeling technique was chosen in view of the fact that it locates the most influential variables when respondents must decide whether to study or not, in the MS2017 in Bogotá over 18 heterosexuals and people from the LGBTI sectors.

This technique selects the variables in order of importance to classify a person with respect to the decision to study or not; simultaneously variables that do not contribute to tree accuracy are ignored (Hernández, Ramírez, & Ferri, 2004: 281-300). The study of school dropout through this data mining technique has also been successfully addressed by Thomas & Galambos (2004) and by Sivakumar, Venkataraman & Selvaraj (2016).

The groupings of the decision trees provided the following binary categories for the independent variables: 1). Marital status: with categories of single and non-single. A non-single person is one who lives as a couple, widow, separated, divorced or married; 2). Age: whose categories were 24-year-olds or less and persons over 24 years of age; and (3). Type of housing tenure: for which the categories were people who live in leased (rented) properties and those who do not. The category of people who do not live in rented properties included those who live in their own fully paid housing, those who live in their own housing but are still paying for it, as well as those who have housing in usufruct and the other possibilities of housing tenure.

The selection of the independent variables was made considering as a base point the work of Barragán & González (2017), the novel contribution of the present work consists of focusing attention on the academic traits of people from LGBTI sectors.

Categories provide a deeper level of understanding of the influence of independent variables in heterosexual people from Bogotá and LGBTI sectors, when deciding whether to remain or leave the formal education system. In technical terms, the C&R method was chosen for the growth of the decision trees. This choice was based on the fact that the lowest Gini index is the criterion of choice for the variables influencing the decision of people from Bogotá to study or not and on the disaggregation into binary categories in each non-terminal node (Tan, Steinbach, & Kumar, 2006).

Sample

The initial sample had 221,809 people from 76,994 households in Bogotá and 98,143 people from 32,086 households in Cundinamarca, for a total of 319,952 people out of 109,111 households. However, the final sample for this work was made up of 172,900 people with an expansion of 6,127,120 since the questions on gender and orientation were applied only to people over 18 years of age. The expansion was the maximum base for calculating proportions or averages when the question was answered by each person in the sample.

Considering that the answers on sex, sexual orientation and gender were given from self-recognition, there could be duplications; in the sample of 172,900, 1,535 LGBTI people were found, that is, people who identified themselves as lesbian, gay, bisexual, transgender or intersex women. The sample included only 13 people who recognize themselves as intersex, so their percentage of the total is practically 0.

With the expansion factors, it was observed that, of the total population in Bogotá, 0.80% corresponds to the population of the LGBTI sectors while 99.20% to the heterosexual population. Figure 1 shows the distribution of people according to the categories of sex, gender and sexual orientation.

Figure 1. Distribution of people by sex, gender and sexual orientation.

Sexo	Género	Orientación sexual
<ul style="list-style-type: none"> • Hombre: 48,93% • Mujer: 51,05% • Intersexual: 0,01% 	<ul style="list-style-type: none"> • Femenino: 51,74% • Masculino: 48,20% • Transgenerista: 0,06% 	<ul style="list-style-type: none"> • Heterosexual: 99,20% • Bisexual: 0,21% • Homosexual: 0,59%

Source: Own elaboration

8 groups of people were defined to take into account in the processing of the data: 1). All the LGBTI people and the comparison with heterosexuals; 2). Lesbian women; 3). Gay men; 4). Bisexual; 5). Transgender men; 6). Transgender women; 7). Intersex; and (8). Population in general. With these groups, the following topics were addressed: household composition and demography, education, citizen participation, housing, health, perceptions about living and working conditions, due to their relationship with the objective of the work, through the research methodology set out below.

Methodology

The records of the 1,535 real people from the LGBTI sectors formed the basis of analysis, so the high degree of uncertainty is noted in estimates of proportions with low occurrence. Data processing included a comparative analysis between the heterosexual population aged 18 years or older in Bogotá and the population of LGBTI sectors.

The research methodology consisted of 5 steps: 1). Explore MS2017 and its themes in relation to LGBTI sectors; 2). Analyze from the descriptive statistics the questions and answers of the subject of education; 3). Look for relationships between the variables of the education theme with others of the survey in relation to the LGBTI sectors; 4). Analyze and discuss the results; 5). Expand the discussion with secondary sources. In step 3 the decision trees were built as statistical models, in which a variable defined with those who were studying and those who were not was used.

The methodology described allowed the results of the next section to be obtained. These results include generalities of the subject of education and particularities of the academic traits for the people from the LGBTI sectors.

Results

With the results obtained with the processing of the MS2017, the traits, attributes and characteristics were found in a general way and in a particular way the academics for the people of the LGBTI sectors in Bogotá.

General features of people from Bogotá from the LGBTI sectors

Table 1 presents the relationship between sex and sexual orientation. Here it can be seen that 51.1% were women and 48.9% men. In addition, 69.3% of homosexual people were men, 29.3% women and 1.4% intersex.

Table 1. Relationship between the variables sex and sexual orientation

Orientation	Men	Women	Intersex
Homosexual	0,693	0,293	0,014
Bisexual	0,432	0,556	0,012
Heterosexual	0,488	0,512	0
Total	0,489	0,511	0

Source: Own elaboration

Table 2 shows the possible values of each group when applying the expansion factors. Here it can be estimated, with a very large variance, that there are 25,076 homosexual men of which 2,635 identify with the female gender, 21,552 with the male and 886 with the transgender. Similar information for the MS2014 can be found in the Observatorio Política Pública LGBTI (Secretaría Distrital de Planeación, 2018).

Table 2. Number of people in each group and category

Orientation	Gender	Sex		
		Man	Woman	Intersex
Heterosexual	Female	104545	3045896	0
	Male	2861491	64271	0
	Transgender	1585	170	20
Homosexual	Female	2635	10105	124
	Male	21552	367	85
	Transgender	886	124	308
Bisexual	Female	80	7071	0
	Male	5241	109	0
	Transgender	277	28	149
Total				6127120

Source: Own elaboration

The average age of heterosexual people is 41.87 years old (standard deviation 16.2 years) while LGBTI people are on average 37.68 years old (standard deviation 14 years). It was observed that a quarter of LGBTI people are aged 18 to 26 and that a fifth of the heterosexual population is in this age group. The predominant age groups in the people of the LGBTI sectors are: from 36 to 49 years with 27.8 years old, from 27 to 35 years old with 26.4% and from 18 to 26 years old with 25%. The proportion of people aged 65 is half in the LGBTI group (4.8%) than in the heterosexual group (9.7%).

For socioeconomic stratification, it was found that the majority of LGBTI people (36.1%) reside in stratum 3 properties, 23.5% in stratum 2, 17.4% in stratum 4, 15.1% in stratum 1 and 5.9% in stratum 5. Only 1.9% reside in stratum 6. The majority of heterosexual people (39.6%) reside in stratum 2 properties, 37.9% in stratum 3, 9.6% in stratum 4, 7.9% in stratum 1, 3.3% in stratum 5 and only 1.7% in stratum 6. The comparison showed that the proportion of LGBTI people living in stratum 1 housing is almost double that of heterosexual people. However, 25.4% (one of four) LGBTI persons reside in stratum 4 or higher housing while only 14.6% of the heterosexual sector does so.

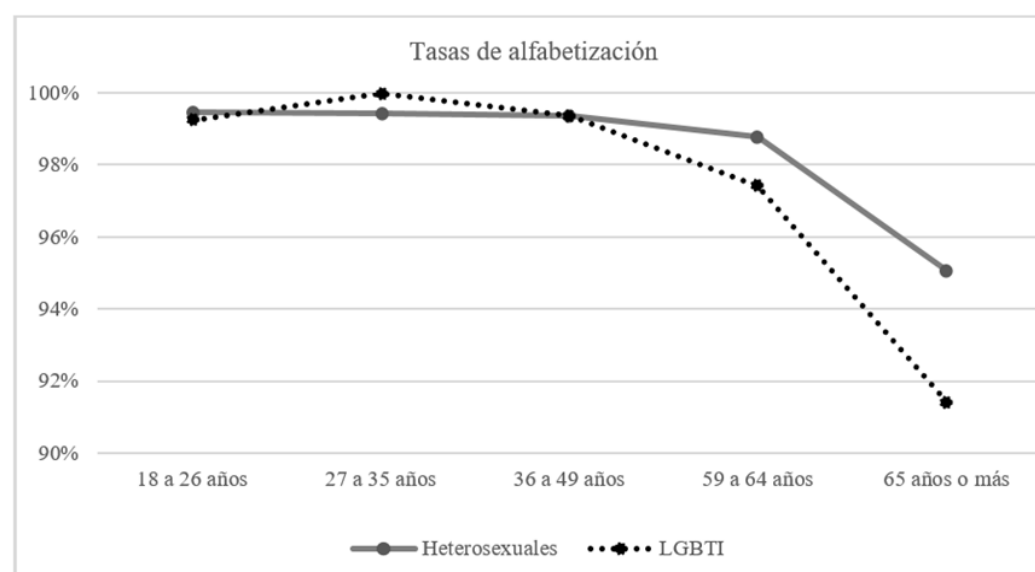
Regarding marital status, 46.6% (22,898) of LGBTI people are single, a much larger proportion than in heterosexual people where the statistic is 33.2%. It is followed by the percentage of LGBTI people who are not married but have lived as a couple for two years or more with 26.2% (12,871). The percentage of heterosexual persons who stated that they have not been married and have been cohabiting for less than two years is 2.8%, which contrasts with the fact that among "LGBTI persons this percentage is almost tripled: 7.8% stated that they have not been married and have been cohabiting for less than two years.

In MS2017, 28.5% of heterosexual people and 32.6% of LGBTI people say they have not always lived in Bogotá. The most frequent reasons why LGBTI people moved to Bogotá were labor (59.8%), then, for education (23.8%) in third place, to be near a household member (7%). In heterosexual people, the main reasons are the same with 55.2%, 12.3% and 9%, respectively.

Academic traits of LGBTI people from Bogotá

The literacy rates of LGBTI people in Bogotá and those of heterosexual people are 98.8% in both cases. In the MS2014, the global literacy rate was 98.5% (Barragán & González, 2017: 19) and that of LGBTI people was 98.8%. However, literacy rates show different behaviour for age groups. Figure 2 shows that younger people have more or less similar literacy rates; and older people, especially those aged 65 and over, have lower literacy rates, especially if they are people from the LGBTI sectors.

Figure 2. Literacy rates by age group.



It was found that the population of the LGBTI sectors studying at the time of the survey was 15% (an estimated 7,353 people), while 11.6% of heterosexual people of legal age did so. In contrast, the three main reasons for not studying in people aged 5 to 34 were: they consider that they have already finished (26.6%), due to educational costs or lack of money (26.6%) and need to work (25.6%). In heterosexual people, these same three reasons are maintained with slight variations in their distribution: 22.3%, 27.5% and 30% respectively (for details see Table 3). It is highlighted that the reason for not studying "consider that they have already finished" cannot be identified with student desertion.

Of the adults from the LGBTI sectors who are studying, approximately 81.8% do so at later than average levels. In the heterosexual population, the figure is 82.2%. The difference is that there is a higher percentage of heterosexual people aged 18 and over who choose university education over technological education 63.9 against 7.7%, while in LGTBI people 59.9% choose university education and 10.9% choose technological education.

19.3% of LGBTI people aged 35 and over who were not studying said that the highest educational level achieved was university, while only 14.5% of heterosexual people had reached it. This means that the barriers to access or permanence to education are not so strong as to be reflected in a decrease in the percentage of university graduates.

Additionally, it was found that 35.8% of the people from the LGBTI sectors, over 35 years of age and who were not studying at the time of the survey, had completed higher education (university, specialization, master's degree and doctorate), while in the heterosexual population this statistic is 23.5%.

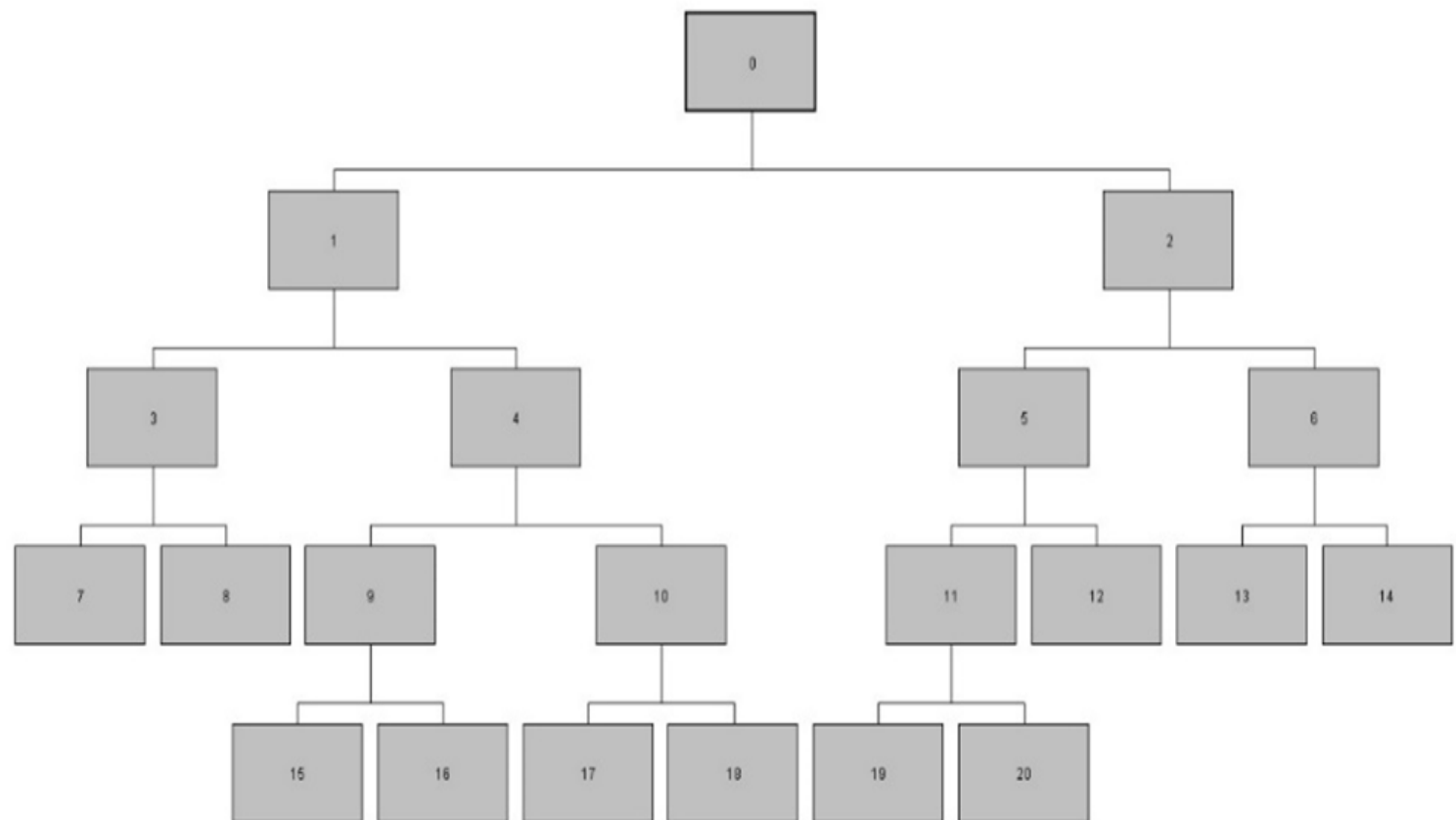
Table 3. Reasons not to study

Reason	LGBTI		Heterosexual	
	Number	Percentage	Number	Percentage
Consider that they are not of school age	1240	2,5	61147	1
Consider that they have already finished	4746	9,7	385055	6,3
Educational costs or lack of money	4733	9,6	475033	7,8
Must take care of household chores	314	0,6	95740	1,6
Need to work	4568	9,3	517143	8,5
Don't like it or Are not interested	737	1,5	40879	0,7
Married or in a relationship	298	0,6	65488	1,1
Lack of vacancies	97	0,2	6556	0,1
Pregnancy	88	0,2	13445	0,2
Disease	185	0,4	10902	0,2
Other	809	1,6	45613	0,8
Total	17814	36,2	1725807	28,4

Source: Own elaboration

Figure 3 is the panoramic map of 20 nodes of the decision tree that examines the most influential variables in the decision to study or not of the people from the LGBTI sectors. These nodes and the variables will be made explicit in figures 4 to 6. As indicated in the methodology, the independent variables included in the analysis are: marital status, age, sex and the type of housing tenure.

Figure 3. Decision tree map: Influencing variables in the decision to study of LGBTI people aged 18 and over



Node 0 of Figure 4 says that 15% of LGBTI people aged 18 and over were studying when the survey was applied. The tree states that the most influential variable in the decision to study of this group of people is age. This being a quantitative variable in a C&R tree, successive cuts are made with the purpose of converting it into a binary variable. The chosen cut is in 24 years, which minimizes the Gini index and better explains why people decide to study or not. In Node 1, almost half of the LGBTI people (46.7%) aged 24 and under were studying at the time of application of the survey, this figure is high compared to those over 24 years of age (8.1%). The 24 and 44 years are milestones (critical ages) in the determination of citizens of the LGBTI sectors in their permanence in the educational system. Less than 1% of the student population is over the age of 44.

Figure 4. First nodes for the tree for LGBTI people.

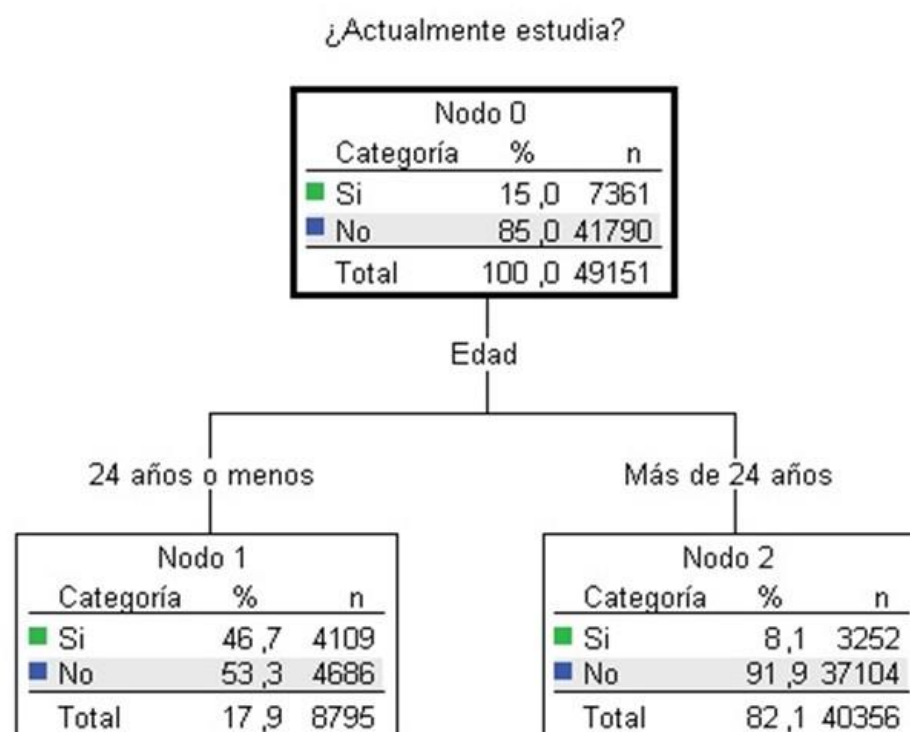
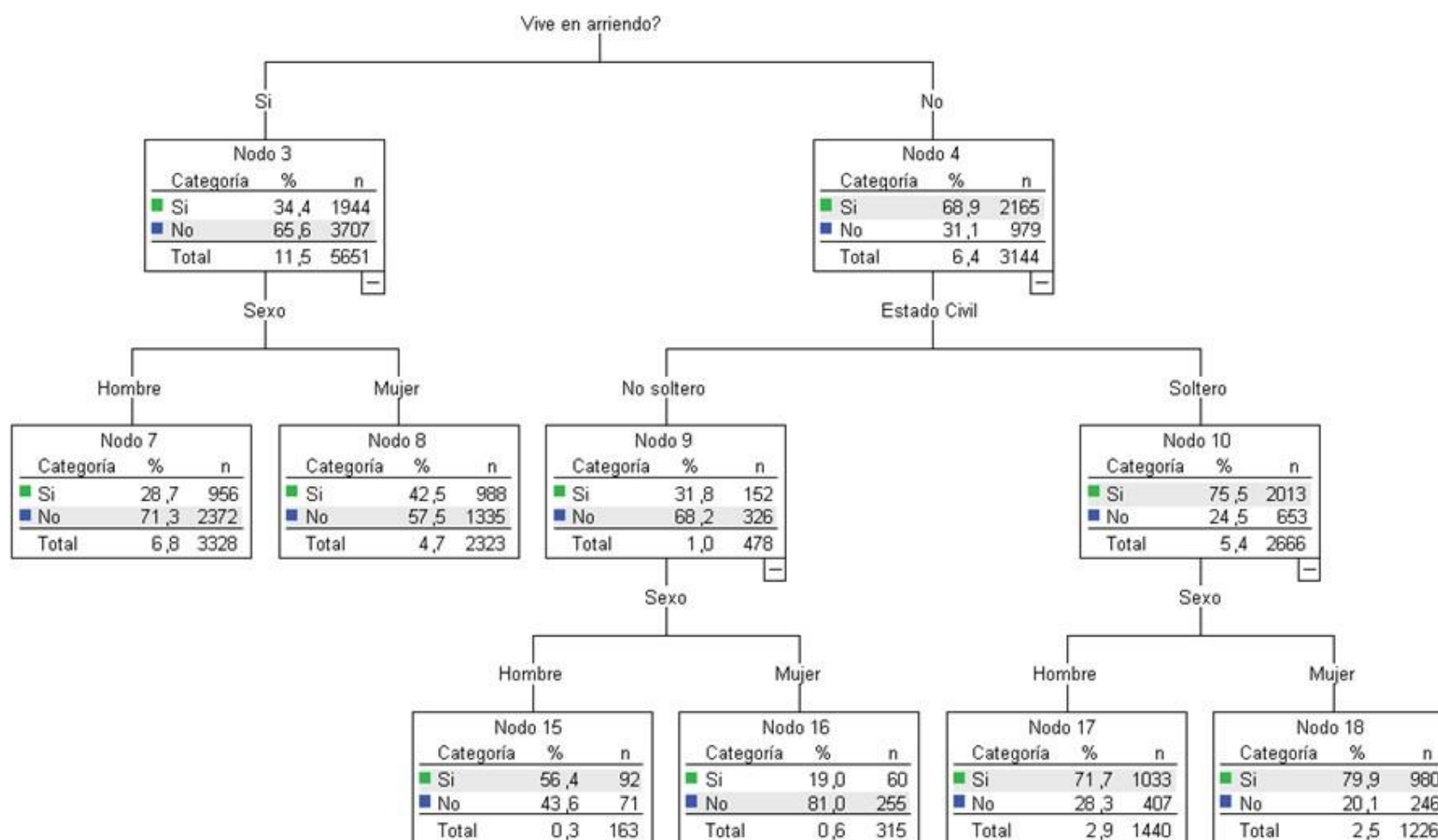


Figure 5 shows that in younger people from the LGBTI sectors, the type of housing tenure is a very important variable in the decision to study, probably because of its association with the household economy. A lower Gini index was obtained by converting this last variable into binary: people who live on rent and people who do not (own housing, own housing and have not finished paying, housing in usufruct). Thus, in young people from the LGBTI sectors, the percentage of those who were studying goes to only 34.4% in those who live on rent. On the contrary, if these people do not live on rent, the percentage is almost 67% (nodes 3 and 4). In the segment of people aged 24 and under from the LGBTI sectors living on rent, sex is an important variable, since it is women who in greater percentage remain in the educational system 44.5% compared to 28.7% of men (nodes 7 and 8).

Going through node 4, marital status is the most decisive variable for young and non-rental LGBTI people. This variable was also added in a binary way, achieving a lower value of the Gini index: single and non-single people (living in a couple, widowed, separated, divorced, married). In node 4, the percentage of people in the education system rises to more than 75% in those who are single (people 24 years old or younger who do not live on rent) and is reduced to less than 32% if the person is non-single. In nodes 15 and 16, of people aged 24 and under who do not live on rent and are not single, the probability of remaining in the education system is higher in men than in women (56.4% and 19%, respectively). The relationship is reversed in people aged 24 and under from the LGBTI sectors who do not live on rent and who are single, 80% of women study compared to 71.7% of men.

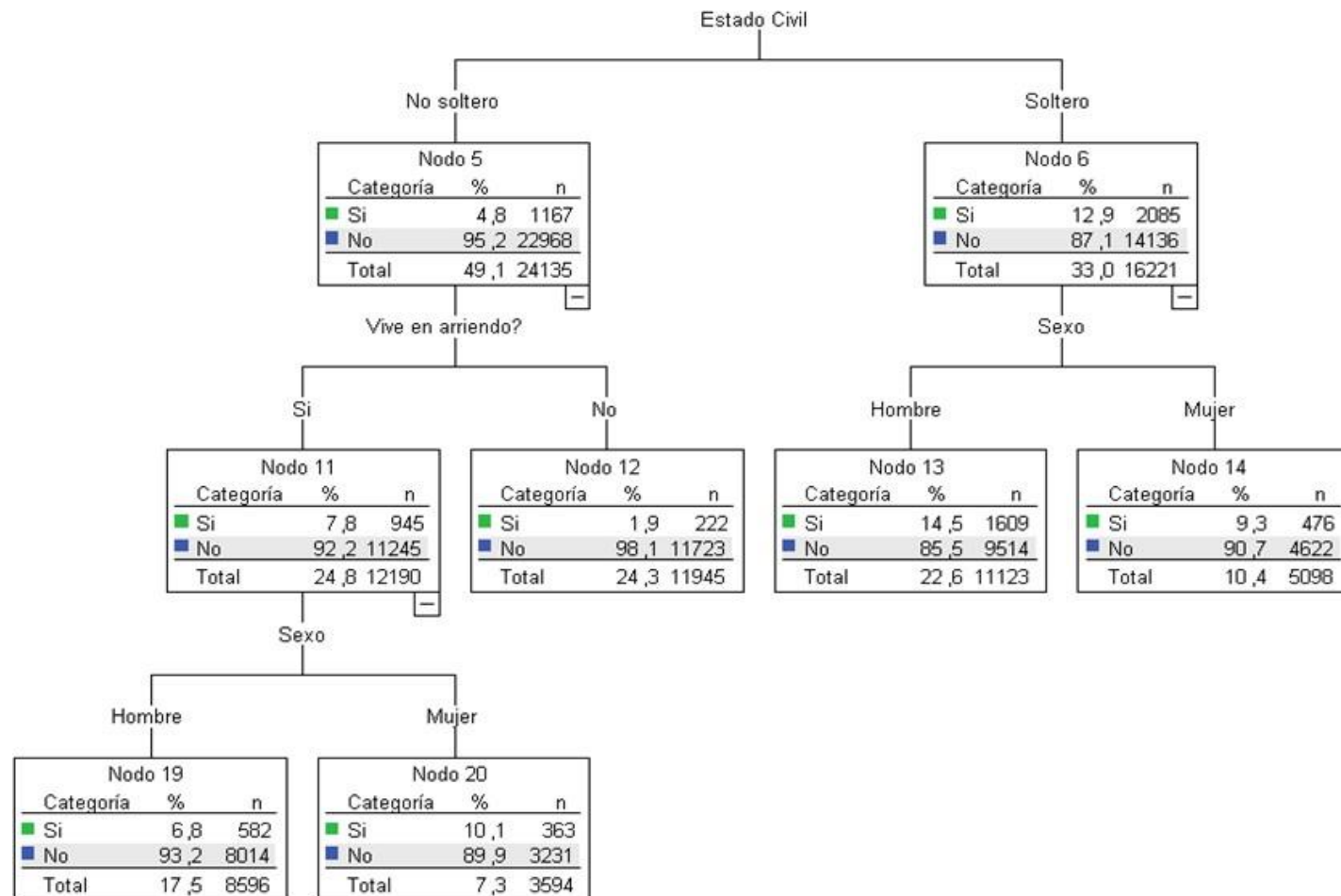
Figure 5. Nodes 3, 4, 7, 8, 9, 10, 15, 16, 17 and 18 of the LGBTI decision tree



Source: Own elaboration

Nodes 5 and 6 of Figure 6 show that LGBTI people over the age of 24 contain low percentages of people in the education system: 12.9% in the education system if they are single and only 4.8% if they are non-single. Here the best scenario happens in single men because 14.5% of them are in the education system compared to 9.3% of women (nodes 13 and 14). In nodes 11, 12, 19 and 20, the percentages of those who remain in the education system are always low: people over 24 years old, non-single, living or not on rent, men and women.

Figure 6. Nodes 5, 6, 11, 12, 13, 14, 19 and 20 of the LGBTI decision tree

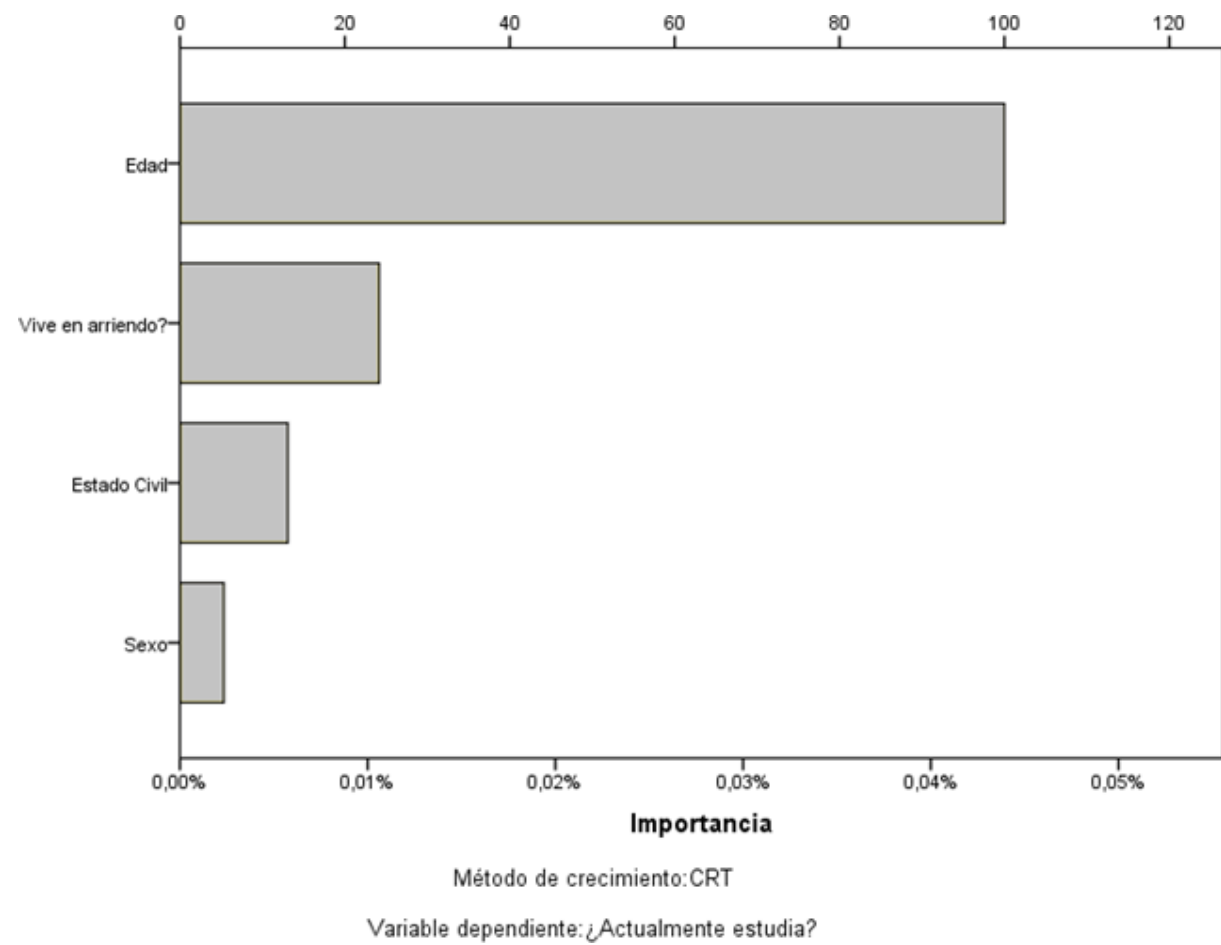


Source: Own elaboration

Figure 7 presents the evaluation of the weight of the variables included in the decision tree. It is concluded that in order of importance are age, whether or not they live on rent, marital status and sex.

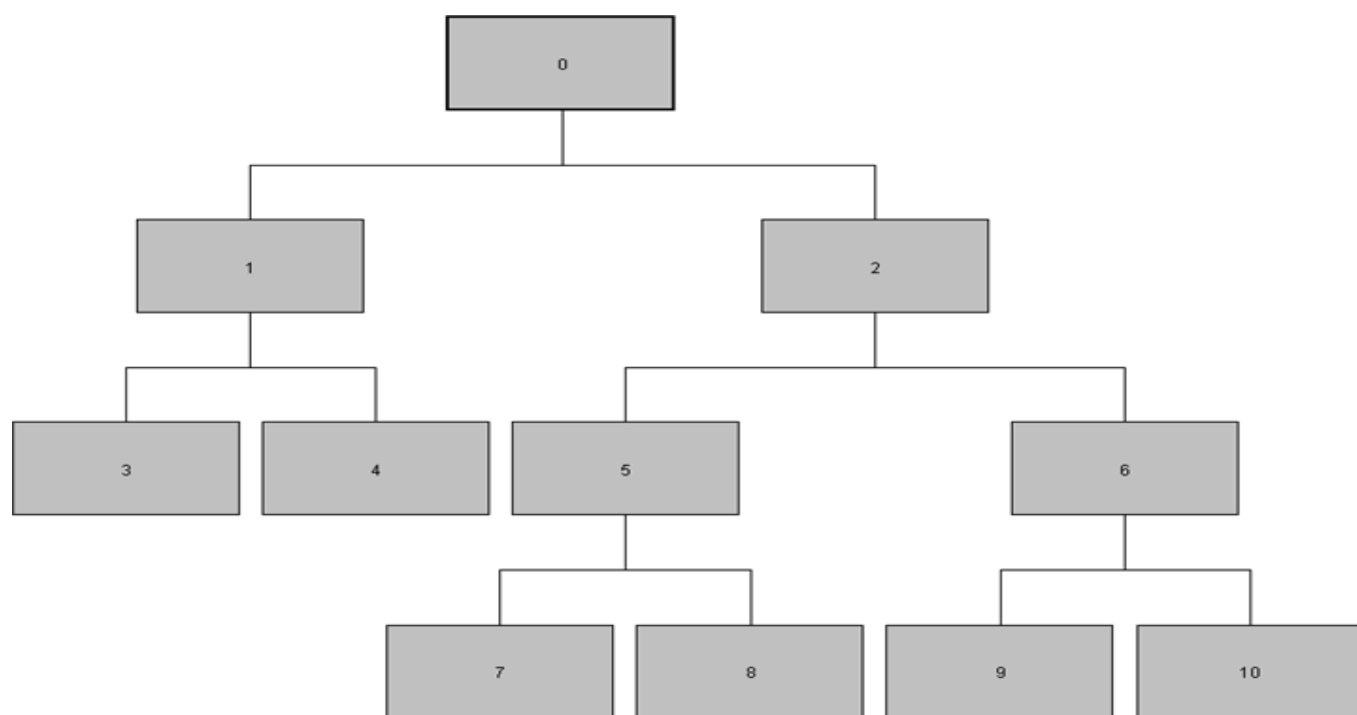
Similarly, Figure 8 is the map of 10 nodes corresponding to the decision tree that examines the most influential variables in the decision to study for heterosexual people surveyed in the city of Bogotá, aged 18 and over. Again, the variables included were: marital status, age, sex and the type of housing tenure.

Figure 7. Importance of variables for LGBTI



Source: Own elaboration

Figure 8. Decision tree map: Influential variables in the decision to study heterosexual people aged 18 or more



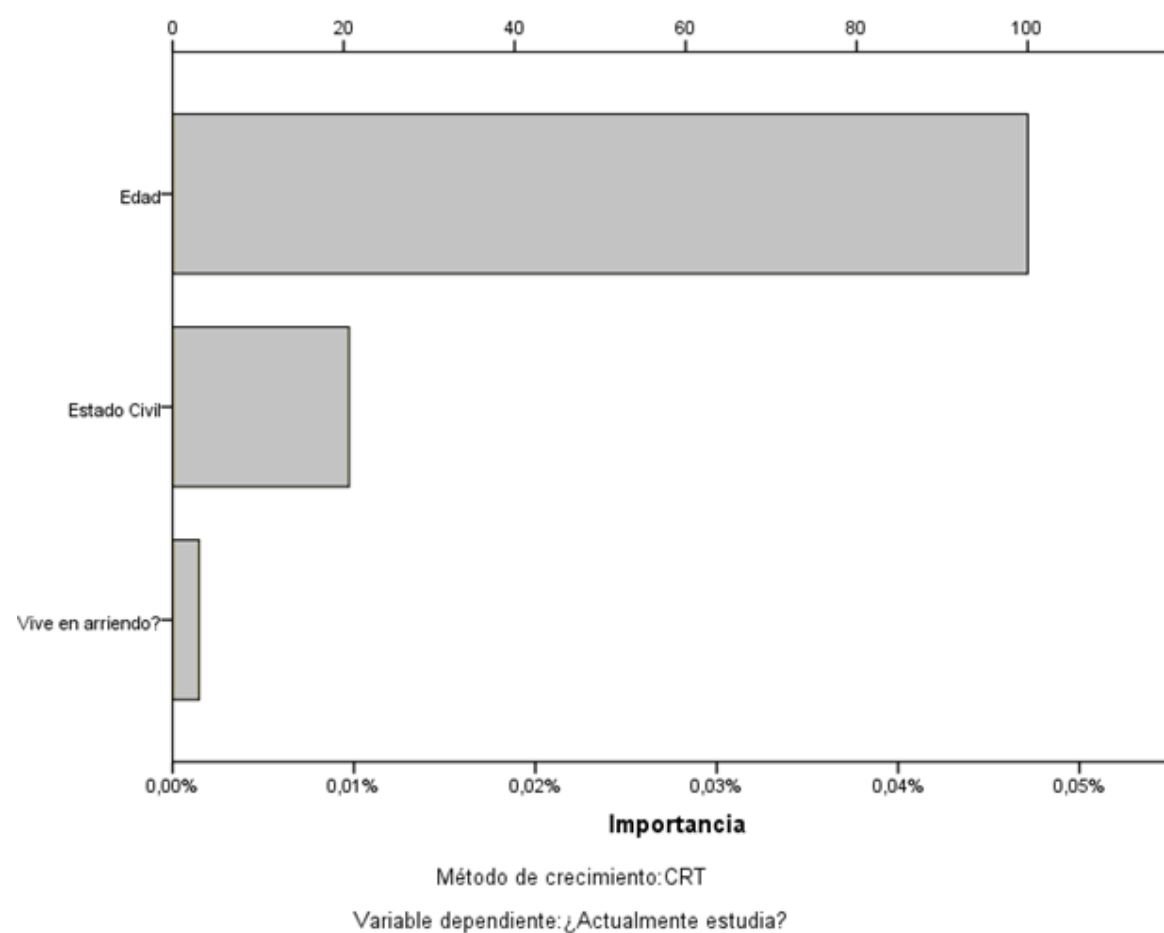
Source: Own elaboration

According to the MS2017, 11.6% of Bogotá residents aged 18 and over who recognize themselves as heterosexual were studying at the time of application of the survey. This percentage is 3.4% less than that of LGBTI people, which coincides with the SDP reports for 2014 and 2017 where they are placed with higher levels of schooling and higher percentage of graduation at the highest levels of the education system. The tree also established that age is the most influential variable in the decision to study for this group of people. In Node 1, almost half of the LGBTI people (46.7%) aged 24 and under were studying at the time of application of the survey, this figure is high compared to those over 24 years of age (8.1%). While for heterosexual people who were over 24 only 5% did.

Marital status is the most influential variable on the decision to continue in the education system for heterosexual people over the age of 24 with 10.9% of those who are single and 3% of non-single people.

The path of node 2 of Figure 8 leads to heterosexual people from Bogotá aged 24 years or younger, showing that the most important variable is also marital status. 54.5% of singles and 13.4% of non-singles continued in the education system. In importance, after marital status, is the type of housing tenure. 60.5% of heterosexual people from Bogotá aged 24 and under who do not live on rent remain in the education system, almost 15% more than those who live on rent. In the path of node 6 of Figure 8, heterosexual people aged 24 or less are not single, only 11% of those who live in rent were studying and 21.1% of those who do not live in rent were in the educational system.

Figure 9. Importance of heterosexual decision tree variables



Source: Own elaboration

When evaluating the weight of the variables included in this second decision tree, the order of importance of the variables was determined: age, marital status and whether or not they live on rent (see Figure 9). The difference in the importance of the variables between people from LGBTI sectors and heterosexual people is in the order of the last two variables and the exclusion of sex as a deciding variable. The latter

because both men and women from LGBTI sectors are similarly inclined to remain in the educational system and that among heterosexual people it is women.

As mentioned, the two previous decision trees have descriptive purposes, in particular to establish how variables such as marital status, age, sex, type of housing tenure influence the decision to study of heterosexual citizens or LGBTI sectors. However, if the trees are used as predictive models of the decision to study, the rates of correct classification on the same sample analyzed are 87.8% in the case of people from LGBTI sectors and 90% in the case of heterosexual people. This means that, if the factors influencing the decision to study or not remain constant, the probabilities of being right whether a citizen studies or not based on the same variables is 0.90 in the case of heterosexual people and 0.87 in the case of people from LGBTI sectors.

Discussion

This section particularly highlights aspects related to the most influential variables in the decision to study or not for people from Bogotá aged 18 years or older in the MS2017 because they may become explanatory variables for student dropout.

The hierarchies of the variables present in the decision trees (see Figure 7 and Figure 9) showed that the most influential variables in the decision to study or not are: age, type of housing tenure, marital status and sex for LGBTI people; and for heterosexual people: age, marital status and type of housing tenure. A discrepancy is observed between the reasons indicated by people from LGBTI sectors for having stopped studying (obtained by direct question in the MS2017), which are: they consider that they have already finished studying, educational costs and need to work and the most influential variables when deciding whether to study or not.

In 2016, the MEN updated and recorded the quantifiable variables in the determinants of dropout through the System for the Prevention and Analysis of Dropout in Higher Education Institutions: stratum, gender, parents' educational level, family income, classification of the family nucleus in the system of potential social beneficiaries, number of members of the family nucleus, employment status and age at taking the State tests as well as the results (Ministerio de Educación Nacional, 2016). Of the hierarchical variables for Bogotá, in this work, only the sex variable coincides. It would be advisable for analyses located in Bogotá to include: age, type of housing tenure and marital status.

On the other hand, at the national level it has been reiterated that men have higher annual dropout rates than women (Consejo Nacional de Educación Superior, 2014: 58; Ministerio de Educación Nacional, 2009: 93; Ministerio de Educación Nacional, 2016: 3). However, the decision trees showed that both men and women from the LGBTI sectors, with similar intensity, remain in the educational system and among heterosexual people, it is women who remain in the educational system in higher percentages.

Conclusions

Literacy rates for people from LGBTI sectors according to the MS2014 and MS2017 were 98.8% in both cases. People aged 65 and older from LGBTI sectors have lower literacy rates than heterosexual people.

LGBTI people are on average younger than heterosexuals and show a greater willingness to be linked to the educational sector.

The gender split (men and women) in the population over 18 years of age is close to 50-50. The percentage of men who recognize themselves as homosexual is 69.38% compared to 29.3% for women. This can possibly be explained by the fact that it is more difficult for women to express their sexual identity. Through the decision trees, men and women from the LGBTI sectors, with similar intensity, remain in the educational system while in heterosexual people it is women.

The proportion of LGBTI people residing in strata 1 real estate (low) is almost double that of heterosexual people. However, one out of every four LGBTI people resides in stratum 4 or higher,

while only 14.6% of heterosexual people reside in this type of property. In other words, a higher proportion of LGBTI people reside in the most economically vulnerable stratum, while a larger proportion reside in stratum 4, probably due to their climbing up the social ladder through education.

The proportion of LGBTI singles is 13 percentage points higher than the proportion of heterosexual singles.

The reasons why LGBTI people and heterosexual people have dropped out of school are the same: they consider that they have finished school, educational costs and need to work. Pregnancy was not relevant for either group.

For people from LGBTI sectors, the most influential variables when deciding whether or not to study are: age, type of housing tenure, marital status and sex. For heterosexual people, the hierarchy of variables is: age, marital status and type of housing tenure. It would be desirable to include these variables in the analysis of student desertion at the Bogotá level.

References

- Alcaldía Mayor de Bogotá D.C. (January 10, 1993). *Decree 16 of 2013: By which the internal structure of the District Planning Secretariat is adopted and other provisions are issued*. Retrieved June 8, 2016, from <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=51241>
- Alcaldía Mayor de Bogotá D.C. (1999). *Alcaldía Mayor de Bogotá D.C.* Retrieved on October 12, 2018, from Organization Chart: http://www.alcaldiabogota.gov.co/sisjur/organica/tabla_organigrama.htm
- Barragán, S., & González, L. (2017). Hallazgos sobre la educación en Bogotá con base en la encuesta Multipropósito 2014. *Sophia*, 13(2), 13-29.
- Colombia. Presidencia de la República. (July 21, 1993). *Decree 1421 of 1993: By which the special regime is dictated for the Capital District of Santa Fe de Bogotá*". Retrieved from <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=9027>
- Congreso de la República de Colombia. (February 8, 1994). *Law 115 of 1994*. Retrieved from: https://www.mineducacion.gov.co/1621/articles-85906_archivo_pdf.pdf
- Consejo Nacional de Educación Superior. (August 19, 2014). *Ministerio de Educación Nacional*. Retrieved May 27, 2015, from Acuerdo por lo superior 2034. Proposal of public policy for the excellence of higher education in Colombia in the scenario of peace: http://www.dialogoeducacionsuperior.edu.co/1750/articles-319917_recurso_1.pdf
- Constitución política de Colombia. (1991). Retrieved from http://www.secretariasenado.gov.co/senado/basedoc/constitucion_politica_1991.html
- Departamento Nacional de Estadística. (October 2017) *Documento metodológico Encuesta Multipropósito - EM 2017*. Retrieved from http://www.sdp.gov.co/sites/default/files/disenio_metodologico_em2017.pdf
- Departamento Nacional de Estadística. (July 18, 2018). *The Multipurpose Survey (Encuesta Multipropósito)* Retrieved from <https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/encuesta-multiproposito>
- Hernández, J., Ramírez, J. M., & Ferri, C. (2004). *Introducción a la minería de datos*. Madrid: Pearson Educación, S.A.
- Ministerio de Educación Nacional. (2009). *Deserción estudiantil en la educación superior colombiana. Metodología de seguimiento, diagnóstico y elementos para su prevención*. Bogotá: Imprenta Nacional de Colombia.
- Ministerio de Educación Nacional. (August, 2016). *Boletín educación Superior en Cifras*. Retrieved from Factors determining drop-out: https://www.mineducacion.gov.co/1759/articles-358136_resource.pdf

- Real Academia de la Lengua Española. (2005). *Linguistic queries*. Retrieved from Los ciudadanos y las ciudadanas, los niños y las niñas: <http://www.rae.es/consultas/los-ciudadanos-y-las-ciudadanas-los-ninos-y-las-ninas>
- Secretaría de Integración Social. (June 24, 2009). *Construyendo la ciudad de derechos*. Retrieved from http://intranetsdis.integracionsocial.gov.co/anexos/documentos/1.3_proc_direc_estrategico/3-Documento%20Construyendo%20la%20Ciudad%20de%20Derechos%2026%20de%20enero-10.doc
- Secretaría Distrital de Planeación. (June 9, 2012). *Formulario EM2011*. Retrieved October 12, 2018, from http://www.sdp.gov.co/sites/default/files/formulario_emb2011.pdf
- Secretaría Distrital de Planeación. (December 2014). *Socioeconomic characterization. Sisbén III Survey*. Retrieved from http://www.sdp.gov.co/sites/default/files/caracterizacion_sisben-04032015.pdf
- Secretaría Distrital de Planeación. (2015). *¿Qué es la encuesta multipropósito?* Retrieved June 8, 2016, from http://www.sdp.gov.co/PortalSDP/Encuesta_Multiproposito_2014/Que_es_la_encuesta
- Secretaría Distrital de Planeación. (December 17, 2015). *Documento Lineamiento Conceptual de la Política Pública LGBTI*. Retrieved from http://old.integracionsocial.gov.co/anexos/documentos/2_cdv/catalogo_2015/Lineamiento%20Conceptual_PP_LGBTI.pdf
- Secretaría Distrital de Planeación. (April 2015) *Multipurpose Survey 2014. Main results in Bogotá and the region*. Retrieved from http://www.sdp.gov.co/sites/default/files/1_boletin_resultados_encuesta_multiproposito_2014.pdf
- Secretaría Distrital de Planeación. (September 9, 2015). *Formulario EM2014*. Retrieved from http://www.sdp.gov.co/sites/default/files/formulario_em2014.pdf
- Secretaría Distrital de Planeación. (2017). *Observatorio Política Pública LGBTI*. Retrieved from Characterization of sectors: <http://www.sdp.gov.co/gestion-socioeconomica/diversidad-sexual/observatorio>
- Secretaría Distrital de Planeación. (July 18, 2018). *Formulario EM2017*. Retrieved from http://www.sdp.gov.co/sites/default/files/formulario_em2017_1.pdf
- Secretaría Distrital de Planeación. (December 2018). *Observatorio Política Pública LGBTI*. Retrieved from Preliminary Information of the results of the Biennial of Cultures and Multipurpose surveys 2017: http://www.sdp.gov.co/sites/default/files/boletin_18_informacion_distrital_sector_igbti.pdf
- Secretaría Técnica Dirección de Diversidad Sexual. (2015). *Balances and perspectives of LGBTI public policy 2012-2015*. Bogotá: Secretaría Distrital de Planeación. Retrieved from http://www.sdp.gov.co/sites/default/files/balances_y_perspectivas_2012_2015_0.pdf
- Sivakumar, S., Venkataraman, S., & Selvaraj, R. (2016). Predictive modeling of student dropout indicators in educational data mining using improved decision tree. *Indian Journal of Science and Technology*, 9(4), 1-5. doi: 10.17485/ijst/2016/v9i4/87032
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Boston: Pearson Education, Inc.
- Thomas, E., & Galambos, N. (2004). What satisfies students? Mining Student-Opinion Data with Regression and Decision Tree Analysis. *Research in Higher Education*, 45(3), 251-269. doi:doi.org/10.1023/B:RIHE.0000019589.79439.6e