Microdata management of ENSU in the study of urban public safety perceptions

Manejo de microdatos de la ENSU en el estudio de percepciones de seguridad pública urbana

MERY-RUIZ, Miriam Elizabeth*†

Universidad Autonoma de Coahuila. Unidad Saltillo. Saltillo, Coahuila, 25280, México.

ID 1ˢᵗ Author: *Miriam Elizabeth, Mery-Ruiz* / **ORC ID:** 0000-0003-2416-0351, **CVU CONAHCYT ID:** 508959

Abstract

This paper presents a comprehensive approach to managing and analyzing microdata from the National Survey of Urban Public Safety (ENSU) by the National Institute of Statistics, Geography, and Informatics (INEGI). The core objective is to elucidate the intricacies of urban public safety perceptions through robust statistical tools and techniques, specifically utilizing Stata software. We focus on the gendered nuances of safety perception in public parks during daylight hours, a critical aspect often overlooked in public safety discourse. By merging various ENSU databases, we dissect the interplay between sociodemographic factors and safety perceptions, offering a granular view that challenges traditional narratives of urban security. The study's findings aim to bridge the gap between subjective perceptions of safety and objective data, providing actionable insights for policymakers and urban planners. Through methodological rigor and detailed data analysis, this research contributes to a deeper understanding of urban safety dynamics and fosters the development of gender-responsive strategies to enhance safety in urban public spaces.

Resumen

Este documento presenta un enfoque integral para el manejo y análisis de microdatos de la Encuesta Nacional de Seguridad Pública Urbana (ENSU) realizada por el Instituto Nacional de Estadística, Geografía e Informática (INEGI). El objetivo principal es dilucidar las complejidades de las percepciones de seguridad pública urbana a través de herramientas y técnicas estadísticas robustas, utilizando específicamente el software Stata. Nos centramos en las sutilezas de la percepción de seguridad basadas en género en parques públicos durante las horas diurnas, un aspecto crítico que a menudo se pasa por alto en el discurso de seguridad pública. Al fusionar varias bases de datos de la ENSU, diseccionamos la interacción entre los factores sociodemográficos y las percepciones de seguridad, ofreciendo una visión detallada que desafía las narrativas tradicionales de la seguridad urbana. Los hallazgos del estudio buscan cerrar la brecha entre las percepciones subjetivas de seguridad y los datos objetivos, proporcionando conocimientos prácticos para los responsables de la formulación de políticas y los urbanistas. A través de la rigurosidad metodológica y el análisis detallado de los datos, esta investigación contribuye a una comprensión más profunda de las dinámicas de seguridad urbana y fomenta el desarrollo de estrategias sensibles al género para mejorar la seguridad en los espacios públicos urbanos.

Gender, Public space, Urban studies   **Espacio público, Estudios urbanos, Género**

* Author's correspondence (E-mail: miriam.mery@uadec.edu.mx)
† Researcher contributing as first author.

Introduction

Urban public safety represents a critical challenge, encompassing not just the incidence of crime, but also the perceived security among residents in urban areas. The Encuesta Nacional de Seguridad Pública Urbana (ENSU) 2021, a significant statistical project by INEGI, is designed to provide a comprehensive snapshot of national concerns regarding public safety, including the prevalence of antisocial behavior and its impact on the population's daily routines and sense of security (INEGI, 2021). The survey's scope extends to capturing the impact of crime on the household, personal victimization, harassment, and sexual violence, providing a detailed picture of the security landscape across urban Mexico.

The perception of insecurity, however, extends beyond the fear of crime. It negatively affects trust in law enforcement and erodes the quality of life for city dwellers. This is particularly pronounced in marginalized urban areas, where concerns about street safety can lead to a societal retreat among middle-class populations and the fortification of security among the upper classes, thus exacerbating social segregation (INEGI, 2021). The ENSU identifies local "incivilities" - such as neighborhood disputes, public consumption of alcohol, and visible signs of disorder like trash and graffiti - as key factors that amplify feelings of insecurity. These elements, coupled with a lack of community cohesion and confidence in local police, can heighten the sense of vulnerability among residents (INEGI, 2021).

Demographic factors, including age and gender, also influence these perceptions. Although subjective, the quantification of insecurity is crucial as ignoring it may lead to increased distrust in authorities and reduce civic engagement in justice processes, thereby inflating the 'dark figure' of crime - crimes that go unreported and unrecorded in official statistics (INEGI, 2021). Such feelings of insecurity have tangible consequences, altering the daily habits of citizens and limiting their freedom, further diminishing their quality of life (INEGI, 2021).

To analyze the data harvested by surveys like the ENSU, it is essential to understand the data formats available for extraction and analysis. The microdata files provided by ENSU are rich with individualized, anonymized data points that need specialized software for proper analysis, as these files contain raw data without aggregation (INEGI, 2021).

The National Survey of Urban Public Safety (ENSU), conducted by the National Institute of Statistics, Geography, and Informatics (INEGI), is a critical initiative aimed at understanding the multifaceted nature of urban public safety in Mexico. The ENSU, as a part of the National Subsystem of Government Information, Public Safety, and Justice Administration, is designed to offer comprehensive information on public safety perceptions at both the national and city-specific levels (INEGI, 2021). It is an annual project that collects data to estimate the public's perception of safety in their city, focusing on a wide range of topics from witnessing criminal or antisocial behaviors to the impact of crime on personal experiences, including victimization and harassment.

The survey's methodology involves a detailed collection of data on the public's experience with and exposure to crime and delinquency. This includes documenting the changes in activities and routines prompted by fear of crime, incidents within the home, and personal experiences of victimization and harassment. The ENSU's comprehensive approach provides a nuanced picture of how public safety, or the lack thereof, shapes the lives of urban dwellers across Mexico (INEGI, 2021).

One of the ENSU's significant contributions is its investigation into the perception of insecurity. It is well acknowledged that these perceptions can have a profound negative impact on trust in institutions and affect the quality of life of the populace (INEGI, 2021) . By measuring incivilities and other local factors contributing to this sentiment, the ENSU offers invaluable insights into the variables that influence public feelings of safety or danger in their everyday environments (INEGI, 2021).

The ENSU data, rigorously anonymized to protect respondents' confidentiality, is made available in ASCII file formats that require specialized statistical software for analysis. These microdata files are essential for researchers and policymakers to understand and address the factors contributing to public safety perceptions (INEGI, 2021).

Microdata, which consist of individual records containing information on variables of interest, are pivotal in analyzing and interpreting the complex dynamics of public safety perceptions. The granularity of microdata allows researchers to dissect and understand the nuances of how individuals experience and perceive safety within their urban environments. The National Survey of Urban Public Safety (ENSU) by INEGI employs microdata to capture a wide array of variables that reflect the perceptions and experiences of the public regarding safety, incivilities, and crime in their cities (INEGI, 2021).

The detailed nature of microdata is particularly valuable when examining the influences on the public's sense of security. It enables the identification of patterns and correlations that might be obscured in more aggregated data. For instance, ENSU microdata can reveal how demographic variables like age and gender intersect with experiences of crime or observations of disorder, such as litter or vandalism, providing deeper insights into the societal and environmental factors that impact feelings of safety (INEGI, 2021).

Moreover, microdata are instrumental in understanding the behavior changes that individuals may adopt in response to their perceptions of insecurity. They offer a detailed view of the modifications in daily routines or the avoidance of certain areas due to fear of crime, which can significantly alter the social dynamics within urban spaces (INEGI, 2021).

The ENSU microdata files, made available in a structured ASCII format, contain individual responses that, while anonymized to maintain confidentiality, provide the raw, unaggregated data essential for robust statistical analysis. This level of detail is crucial for policymakers and practitioners who seek to design interventions and strategies that are responsive to the specific needs and conditions of different urban populations (INEGI, 2021)

The primary aim of this research is to harness the potential of microdata from the National Survey of Urban Public Safety (ENSU) to dissect and understand the subjective perceptions of safety among urban populations in Mexico. This study intends to leverage the detailed, individual-level data provided by ENSU to explore the factors that influence perceptions of safety and insecurity, and how these perceptions correlate with various demographic and socio-environmental factors. By doing so, this research endeavors to contribute to a more nuanced understanding of public safety and inform the development of targeted, evidence-based public policies and urban planning strategies.

This study seeks to navigate the depths of ENSU microdata to decipher the varied tapestry of public safety perceptions within urban locales, examining how these perceptions differ among diverse demographic groups. By dissecting the socio-environmental factors documented in the ENSU, the research aims to determine their relationship with the feelings of safety or insecurity experienced by urban residents. An exploration into how crime and antisocial behavior, as reported in the ENSU, shape both individual and communal perceptions of safety is central to this analysis. Moreover, the investigation will consider the extent to which fear-induced alterations in daily life activities mirror the wider social and economic currents in urban settings. A pivotal aspect of the study is to ascertain if a thorough analysis of the ENSU microdata can shed light on how gender dynamics influence the perception of safety within the public spaces of the city. These inquiries are integral to the study's broader objective: to intricately understand and address the complex web of factors that contribute to the mosaic of public safety perceptions in Mexico's urban centers, thereby informing more nuanced policy responses.

Studies on urban public safety perceptions

Understanding urban public safety perceptions has been a central theme in a broad array of academic inquiries, reflecting the complexity of how individuals perceive safety and security within their urban environments. Research in this area is diverse, employing a range of methodologies to capture the nuanced experiences and perceptions of urban residents.

Survey-based research, such as that conducted through the ENSU, provides a wealth of microdata for researchers to gauge public sentiments regarding safety. These surveys typically include questions about personal experiences with crime, observations of neighborhood conditions, and general feelings of safety or insecurity. The quantitative data derived from these surveys reflect the subjective feelings and objective realities of urban living, offering a broad overview of public safety perceptions across different demographics and locales.

Complementing these are in-depth qualitative studies that delve into the personal narratives and complex emotions surrounding public safety. Through interviews and focus groups, researchers uncover the psychological and emotional impacts of crime, and the perceived effectiveness of law enforcement and community safety measures. These studies provide depth and context to the statistical patterns observed in survey data, revealing the human stories behind the numbers.

Comparative analyses are also prevalent, with researchers conducting studies across different neighborhoods, cities, or countries to understand the diverse factors influencing public safety perceptions. These studies identify patterns and trends, highlighting how urban design, policing strategies, and social cohesion might impact feelings of safety in various contexts.

Longitudinal research tracks changes in public safety perceptions over time, helping to understand how events, policy changes, or shifts in social attitudes impact feelings of safety. This approach is particularly valuable in assessing the effectiveness of policy interventions and societal changes on public sentiment and behavior.

Lastly, the field of urban public safety perceptions benefits greatly from cross-disciplinary approaches. It's an area where criminology, sociology, urban planning, and psychology intersect, revealing how interconnected factors collectively influence perceptions of safety. From street lighting and urban design to social policies and community programs, a wide range of elements are considered in understanding and improving the public's sense of safety.

Previous use of ENSU data in research

The Encuesta Nacional de Seguridad Pública Urbana (ENSU) has been utilized in various statistical studies aimed at enriching national information related to government, public safety, and justice administration. The 2021 ENSU, in particular, has provided researchers with data to estimate public perceptions of urban safety and to measure the witnessing of criminal or antisocial behaviors that affect the population. This includes examining changes in activities and routines due to fear of crime, incidents of crime in the home, experiences of victimization, harassment, and sexual violence, all of which are essential aspects of understanding urban safety from a societal viewpoint.

Researchers have employed statistical software such as Stata to analyze ENSU data, which comes in various formats including CSV, DBF, DTA, Rdata, and SAV, to perform descriptive analysis by merging different databases provided by the ENSU microdata. These efforts have been instrumental in addressing key questions such as the influence of gender on the perception of safety or insecurity in parks, showcasing the ENSU's role in informing policy and urban planning by providing a data-driven foundation for understanding and addressing public safety concerns.

Methodology

The data management procedure was meticulous, involving the integration of three primary ENSU databases: the household information table, the sociodemographic table of household members, and the main questionnaire table. These databases were merged using Stata commands, ensuring that each table was properly aligned based on their relational keys. This merging was a critical step in creating a cohesive dataset that would allow for a comprehensive analysis of the variables of interest.

Following the integration, we performed a summary check to verify the contents of the new merged database. This step was the base to ensure that the data fusion had been successful and that the dataset was correctly structured for further analysis. To enhance the readability and interpretability of the data, a thorough labeling process was conducted.

This involved assigning descriptive labels to each variable and category based on the data structure documentation provided by INEGI. These labels were essential for making the data more accessible and for facilitating an accurate analysis of the responses.

The core of the analysis hinged on the use of contingency tables, which are instrumental in examining the relationships between two variables. In this case, the focus was set on exploring the influence of gender on the perception of safety or insecurity in urban parks. By cross-tabulating gender with the perceptions of safety, it is possible to scrutinize the association between these variables, thus addressing one of the primary research questions of our study.

The methodology adopted for this study reflected a rigorous approach to data management and analysis, with a focus on ensuring the integrity and clarity of the ENSU microdata. Through careful preparation and detailed examination of the dataset, the aim was to extract meaningful insights that would contribute to a deeper understanding of public safety perceptions within urban environments.

Results

ENSU Microdata with STATA

As a first step to obtain the data, the ENSU Microdata section was accessed at INEGI's website:
https://www.inegi.org.mx/programas/ensu/#Microdatos.

The page shows the formats in which it is possible to download microdata, from the most recent database which is for the month of March 2022, and which are: CSV, DBF, DTA, Rdata and SAV, as shown in **Figure 1**.

In this case, the DTA format was downloaded for Stata.

Figure 1 Latest Microdatabases available for downloading
Source: INEGI, 2022

The file was then unzipped because it was downloaded as a .zip file, as shown in **Figure 2**.

Figure 2 ENSU microdata download file in .zip extension
Source: Self elaboration, 2022

Then a .pdf file was downloaded with the survey descriptors, as shown in **Figure 3**, unzipped as they were downloaded compressed in .zip as shown in **Figure 4**.
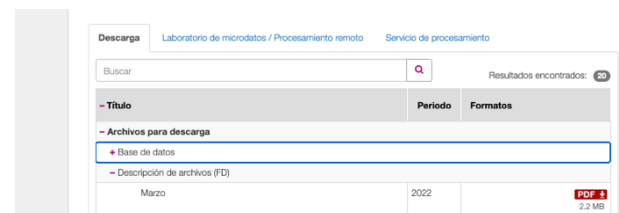
Figure 3 Survey descriptors download
Source: INEGI, 2022

Figure 4 Descriptor file, download as .zip extension
Source: Self elaboration, 2022

The next step was to review the file with the descriptors, called "Database Structure" whose cover page is shown in **Figure 5.**

Figure 5 Cover page of the file "Database structure extension
Source: INEGI, 2022

The ENSU database, according to the INEGI document (2022), consists of 3 tables in which the information captured by the survey is distributed. The following is a list of the tables and the information contained in each one of them. The notation [mm][aa] in the name of the table indicates the two digits of the month and the last two digits of the corresponding year.

Table of housing information (ENSU_VIV_[mm][yyy])
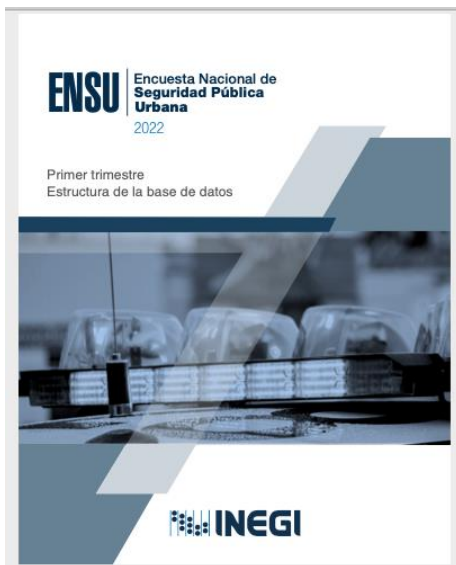
This table contains the general characteristics of the dwelling, as well as of the main household. In addition, it captures information on the data of the operating personnel and the final result of the interview.

Table of socio-demographic information of the members of the dwelling (ENSU_CS_[mm][yyy])

This table contains information on the main sociodemographic characteristics of the resident population of the main household and identifies the population aged 18 and over, since the questions of the ENSU questionnaire are applied to the latter.

Main questionnaire table (ENSU_CB_[mm]_[yyy])

This table contains information on the perception of the population 18 years of age and older about the public safety situation in the urban environment, conflicts or confrontations, government performance, as well as experiences of corruption in the performance of any payment, procedure and/or service. It should be noted that question 1.6 alternates: in the first and third quarters, the inquire is about the ways of finding out about public safety, while in the second and fourth quarters, about personal and household victimization. During the first quarter the frequency of mobility in the last three months is inquired.

Relationship between tables

The table ENSU_VIV_[mm][aa] is related to the table ENSU_CS_[mm][aa] through the primary key of ENSU_VIV_[mm][aa] formed by the fields: CVE_ENT + UPM + VIV_SEL

The description of such fields is as follows:
CVE_ENT = Federation Entity Key
PSU = Primary Sampling Unit
VIV_SEL = Selected dwelling

The table ENSU_CS_[mm][aa] is related to the table ENSU_CB_[mm][aa] through the primary key of ENSU_CB_[mm][aa] formed by the fields: CVE_ENT + UPM + VIV_SEL + R_SEL

The description of such fields is as follows:

CVE_ENT = Federation Entity Key

PSU = Primary Sampling Unit

VIV_SEL = Selected dwelling

R_SEL = Row of the selected person

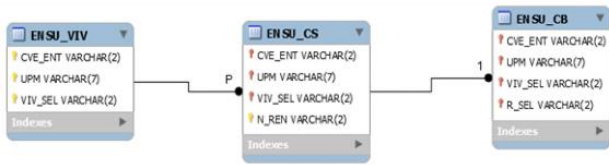**Figure 6** below shows the relationship between the ENSU tables, using the entity-relationship model.

Figure 6 Entity-Relationship model
Source: INEGI, 2022

The entity number in the file, which corresponds to Saltillo, is 17. To continue, the Stata program was opened. In the File menu, the DTA file ENSU_CS was opened, as shown in **Figure 7**.
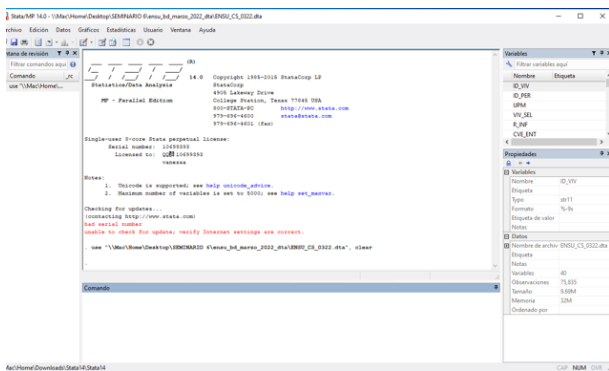
Figure 7 Stata interface to ENSU_CS database
Source: Self elaboration, 2022

Merging databases

As mentioned above, the ENSU has 3 databases: Table of housing information (ENSU_ VIV_[mm][yyy]) 2, table of socio-demographic information of household members (ENSU_CS_ [mm][yyy]), and table of the main questionnaire (ENSU_CB_ [mm]_[yyy]).

Based on the relationship between tables, the following tables were merged ENSU_VIV with ENSU_CS, ENSU_VIV being the key for the merging, and ENSU_CB with ENSU_CS, with ENSU_CB being the key for this merge.

The following commands were introduced in Stata, in order to open each of the 3 databases, save them with a different name to be able to identify them when merged with another database as shown in **Figure 8** and **Figure 9**, as well as in the syntax:
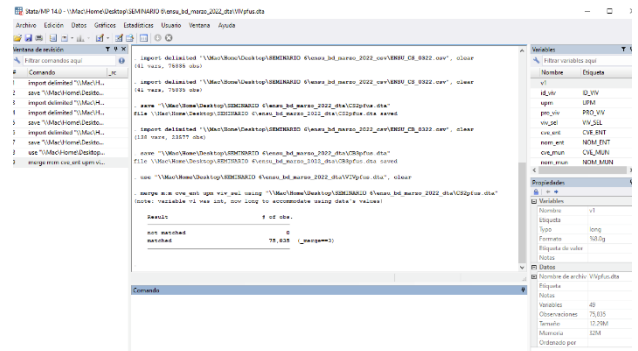
Figure 8 Stata procedure for merging databases
Source: Self elaboration, 2022
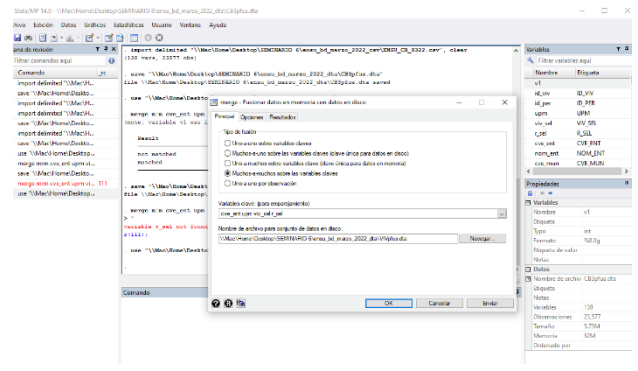
Figure 9 How to merge databases by means of a key with several variables
Source: Self elaboration, 2022

.import   delimited "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_csv\ENSU_VIV_0322 .csv"
(29 vars, 23577 obs)

. save "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\VIVpfus.dta"
file   \\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\VIVpfus.dta saved

.   import   delimited "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_csv\ENSU_CS_0322.c sv", clear
(41 vars, 75835 obs)

.   import   delimited "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_csv\ENSU_CS_0322.c sv", clear
(41 vars, 75835 obs)

. save "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\CS2pfus.dta"
file   \\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\CS2pfus.dta saved

. import delimited
"\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_csv\ENSU_CB_0322.
csv", clear
(138 vars, 23577 obs)

. save "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CB3pfus.dta"
file \\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CB3pfus.dta saved

. use "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\VIVpfus.dta",
clear

. merge m:m cve_ent upm viv_sel using
"\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CS2pfus.dta"
(note: variable v1 was int, now long to
accommodate using data's values)

Result	# of obs.
not matched	0
matched	75,835 (_merge==3)

From the review in the data editor, it can be seen that the merging of the two databases has been satisfactory. See **Figure 10**.
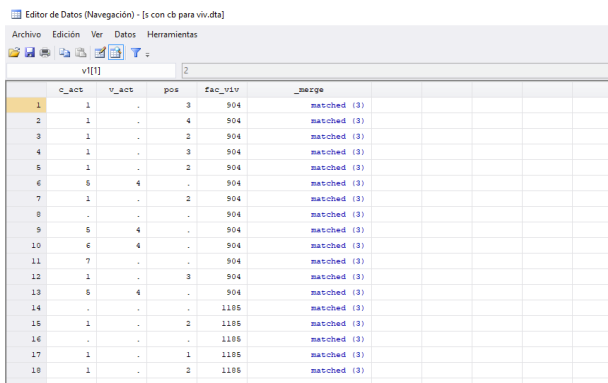
Figure 10 Variable generated by merging two tables
Source: Self elaboration, 2022

Once the databases have been merged, the file is saved as shown in **Figure 11** and in the syntax:
. save "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\VIVpfus.dta",
replace
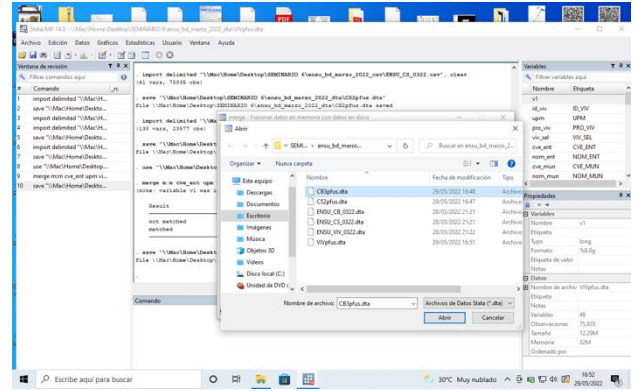
Figure 11 Variable generated by merging two table How to save a merged database in Stata
Source: Self elaboration, 2022

With the above, we have two merged databases, we proceed to merge the remaining database with the already merged file containing two previously merged databases, the following syntax was used:

file\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\VIVpfus.dta saved

. use "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CB3pfus.dta",
clear

. merge m:m cve_ent upm viv_sel using
"\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CS2pfus.dta"
(note: variable v1 was int, now long to
accommodate using data's values)

| Result	# of obs.
not matched	0
matched (_merge==3)	75,835

. save "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CS2pfus.dta",
replace
file \\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\CS2pfus.dta saved

. save "\\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\s con cb para
viv.dta"
file \\Mac\Home\Desktop\SEMINARIO
6\ensu_bd_marzo_2022_dta\s con cb para
viv.dta saved

. use "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\VIVpfus.dta", clear
r(110);

. merge m:m cve_ent upm viv_sel using "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\CB3pfus.dta", gen
> erate(_merge2)

Result	# of obs.
not matched	0
matched	75,835
(_merge2==3)	

With this the 3 databases were merged into one, the one that was used first, which is the key, as shown in Figure 12 and Figure 13.
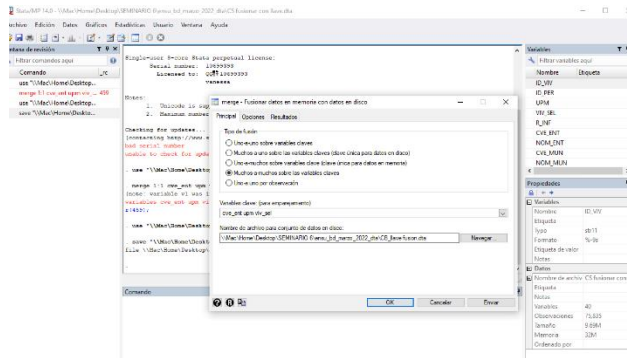
Figure 12 *Merging a previously merged database.* Source: Self elaboration, 2022.
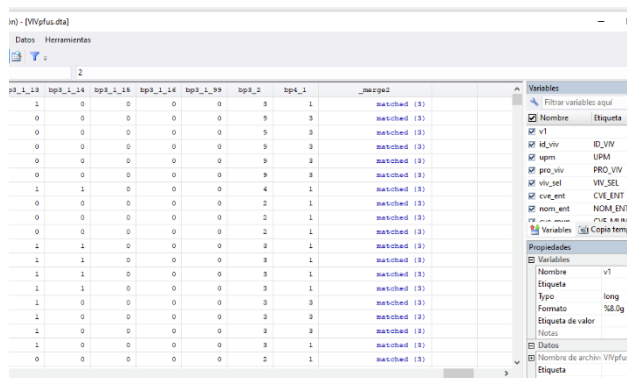
Figure 13 *Data editor with database merged with another database.* Source: Self elaboration, 2022.

Afterwards, a summary of the data was made to see the content of the new database generated and where the mergers of the two databases previously merged to form the new one can be observed, as shown in the syntax:

. sumarize
command sumarize is unrecognized
r(199);

. sum

```
Variable |      Obs       Mean    Std. Dev.      Min        Max
---------+--------------------------------------------------------
      v1 |   75,835   11882.96    6787.768        1      23577
  id_viv |   75,835    1532985    815754.5   100071    3201159
  id_per |        0
     upm |   75,835    1532985    815754.5   100071    3201159
 viv_sel |   75,835   2.996281   1.419402        1          5
---------+--------------------------------------------------------
   r_sel |   75,835   2.052759   1.344188        1         14
 cve_ent |   75,835   15.27669   8.170077        1         32
 nom_ent |        0
 cve_mun |   75,835   35.67581   43.11864        1        553
 nom_mun |        0
---------+--------------------------------------------------------
     loc |   75,835   7.191838   45.90177        0       1599
      cd |   75,835   50.08464   27.61377        1         96
  nom_cd |        0
     per |   75,835        122          0      122        122
   r_def |   75,835          1          0        1          1
---------+--------------------------------------------------------
   h_mud |   75,835   .0438979   .2275574        0          3
   s_mud |   75,835   .6404958   .5864583        0          4
tipo_cuest|       0
   n_ent |   75,835   2.990624   1.410293        1          5
    sexo |   75,835   1.559043    .496505        1          2
---------+--------------------------------------------------------
    edad |   75,835   42.99316   16.03544       18         98
   bp1_1 |   75,835   1.653761   .5825324        1          9
 bp1_2_01|   75,835    1.20476   .4266562        1          9
 bp1_2_02|   75,835   1.795226   .8836877        1          9
```

bp1_2_03 |	75,835	1.590769	.5552989
1 9			
bp1_2_04 |	75,835	2.923254	.7369165
1 9			
bp1_2_05 |	75,835	1.68206	.7522692
1 9			
bp1_2_06 |	75,835	1.458904	.6852451
1 9			
bp1_2_07 |	75,835	1.800435	.7265039
1 9			
bp1_2_08 |	75,835	1.94339	.6880422
1 9			
bp1_2_09 |	75,835	2.003112	.794936
1 9			
bp1_2_10 |	75,835	1.459234	.6781657
1 9			
bp1_2_11 |	75,835	1.658983	.7242149
1 9			
bp1_2_12 |	75,835	1.784097	.8562597
1 9			
bp1_3 |	75,835	2.698569	1.335342
1 9			
bp1_4_1 |	75,835	1.606013	.6108709
1 9			
bp1_4_2 |	75,835	1.393881	.5830183
1 9			
bp1_4_3 |	75,835	1.494758	.606736
1 9			
bp1_4_4 |	75,835	1.765451	.6058924
1 9			
bp1_4_5 |	75,835	1.704635	.9736833
1 9			
bp1_4_6 |	75,835	1.644359	.6138847
1 9			
bp1_4_7 |	75,835	2.057322	.8111155
1 9			
bp1_4_8 |	75,835	1.960084	.8967166
1 9			
bp1_5_1 |	75,835	1.518441	.5963485
1 9			
bp1_5_2 |	75,835	1.571847	.5842218
1 9			
bp1_5_3 |	75,835	1.70962	.5200383
1 9			
bp1_5_4 |	75,835	1.86319	.7926795
1 9			

bp1_5_5 |	75,835	2.30482	.8399946
1 9			
bp1_6_01 |	75,835	.5482033	.4976743
0 1			
bp1_6_02 |	75,835	.1731654	.3783927
0 1			
bp1_6_03 |	75,835	.6064218	.4885464
0 1			
bp1_6_04 |	75,835	.0498978	.2177352
0 1			
bp1_6_05 |	75,835	.2002242	.4001707
0 1			
bp1_6_06 |	75,835	.0208215	.1427874
0 1			
bp1_6_07 |	75,835	.0866223	.2812826
0 1			
bp1_6_08 |	75,835	.5995385	.4899952
0 1			
bp1_6_09 |	75,835	.209758	.4071385
0 1			
bp1_6_10 |	75,835	.097699	.2969091
0 1			
bp1_6_11 |	75,835	.2154414	.4111309
0 1			
bp1_6_12 |	75,835	.0068966	.0827592
0 1			
bp1_6_13 |	75,835	.002334	.0482555
0 1			
bp1_6_99 |	75,835	.0023472	.0483913
0 1			
bp1_7_1 |	75,835	1.519773	.8094822
1 9			
bp1_7_2 |	75,835	1.22727	.481397
1 9			
bp1_7_3 |	75,835	1.150115	.4320403
1 9			
bp1_7_4 |	75,835	1.088732	.3516962
1 9			
bp1_7_5 |	75,835	1.201648	.4951487
1 9			
bp1_8_1 |	50,014	2.722078	1.031708
1 9			
bp1_8_2 |	59,132	2.643357	1.189377
1 9			
bp1_8_3 |	65,011	2.20958	1.47597
1 9			

bp1_8_4	69,512	1.888552	1.271841	1 9
bp1_8_5	61,341	1.874619	1.521194	1 9
bp1_9_1	50,014	2.707682	.9916403	1 9
bp1_9_2	59,132	2.638639	1.108576	1 9
bp1_9_3	65,011	2.142376	1.27722	1 9

-------------+---

bp1_9_4	69,512	1.870166	1.114454	1 9
bp1_9_5	61,341	1.852334	1.393827	1 9
bp2_1	75,835	1.625199	.5299536	1 9
bp2_2_01	28,864	.3378603	.4729889	0 1
bp2_2_02	28,864	.1293307	.3355714	0 1

-------------+---

bp2_2_03	28,864	.3047395	.460305	0 1
bp2_2_04	28,864	.0374861	.1899531	0 1
bp2_2_05	28,864	.2478866	.4317931	0 1
bp2_2_06	28,864	.1238567	.3294237	0 1
bp2_2_07	28,864	.0596244	.2367938	0 1

-------------+---

bp2_2_08	28,864	.1618972	.3683628	0 1
bp2_2_09	28,864	.200492	.4003754	0 1
bp2_2_10	28,864	.1388927	.3458406	0 1
bp2_2_11	28,864	.0920524	.2891049	0 1
bp2_2_12	28,864	.0255682	.1578459	0 1

-------------+---

bp2_2_13	28,864	.0470136	.2116716	0 1
bp2_2_14	28,864	.1065341	.3085253	0 1
bp2_2_15	28,864	.1250346	.3307639	0 1
bp2_2_16	28,864	.0665535	.2492514	0 1

bp2_2_17	28,864	.01074	.1030778	0 1

-------------+---

bp2_2_18	28,864	.1335227	.3401447	0 1
bp2_3_1	25,010	.1142343	.5554046	0 9
bp2_3_2	25,010	.7568972	.625537	0 9
bp2_3_3	25,010	.0552579	.5094035	0 9
bp2_3_4	25,010	.3160336	.6507307	0 9

-------------+---

bp2_3_5	25,010	.1823271	.5969717	0 9
bp2_3_6	25,010	.0641743	.517053	0 9
bp2_3_7	25,010	.0340664	.4900944	0 9
bp2_4_01	25,010	.2602559	.4387829	0 1
bp2_4_02	25,010	.2395442	.4268138	0 1

-------------+---

bp2_4_03	25,010	.0295882	.1694516	0 1
bp2_4_04	25,010	.0397041	.1952671	0 1
bp2_4_05	25,010	.0068373	.0824062	0 1
bp2_4_06	25,010	.0017193	.0414298	0 1
bp2_4_07	25,010	.0010796	.0328397	0 1

-------------+---

bp2_4_08	25,010	.0252299	.1568258	0 1
bp2_4_09	25,010	.4683727	.4990087	0 1
bp2_4_10	25,010	.1786086	.3830319	0 1
bp2_4_11	25,010	.0187125	.1355105	0 1
bp3_1_01	75,835	.5407002	.498344	0 1

-------------+---

bp3_1_02	75,835	.3169776	.4653017	0 1
bp3_1_03	75,835	.4060131	.4910903	0 1

Variable	Obs	Mean	Std. Dev.	Min	Max
bp3_1_04	75,835	.1865102	.3895204	0	1
bp3_1_05	75,835	.5876574	.4922595	0	1
bp3_1_06	75,835	.27275	.4453763	0	1
bp3_1_07	75,835	.1382343	.345148	0	1
bp3_1_08	75,835	.4207424	.4936815	0	1
bp3_1_09	75,835	.0572559	.2323324	0	1
bp3_1_10	75,835	.7613107	.4262852	0	1
bp3_1_11	75,835	.3645414	.4813045	0	1
bp3_1_12	75,835	.5345553	.4988078	0	1
bp3_1_13	75,835	.3528582	.4778622	0	1
bp3_1_14	75,835	.3633283	.4809614	0	1
bp3_1_15	75,835	.0102591	.1007671	0	1
bp3_1_16	75,835	.0157711	.1245896	0	1
bp3_1_99	75,835	.001345	.0366502	0	1
bp3_2	74,537	2.94977	.9175371	1	9
bp4_1	75,835	1.782752	1.126794	1	9
fac_sel	75,835	2212.092	2111.538	38	41469
dominio	0				
est	75,835	2.624672	.7591174	1	4
upm_dis	75,835	27092.53	15535.73	10	54000
est_dis	75,835	1375.959	803.9109	10	2740
r_inf	75,835	2.00265	1.295575	1	14
n_ren	75,835	2.735096	1.768148	1	19
nom	0				

Variable	Obs	Mean	Std. Dev.	Min	Max
c_res	59,734	1.084675	.4027202	1	3
paren	75,835	2.486082	1.325617	1	6
nac_d	14,382	17.03776	14.06581	1	99
nac_m	14,382	8.209846	12.79087	1	99
cod	56,375	1.581783	.4932706	1	2
cod_sel	0				
i_niv	73,297	4.489774	2.615157	0	9
i_anio	70,186	3.429829	1.426816	1	6
a_esc	73,297	1.732008	.4761679	1	9
c_act	56,375	2.620754	2.260659	1	8
v_act	19,825	3.73367	.79527	1	4
pos	38,058	2.317568	.5983632	1	5
fac_viv	75,835	774.0074	603.5436	29	13800
_merge	75,835	3	0	3	3

The interest, for this exercise, is to know the perception of insecurity by gender, so it is necessary to use the ENSU_CB and ENSU_CS databases, which although they are already merged for practical purposes of this exercise, it is important to know in order to be able to consult the nomenclatures of the results.
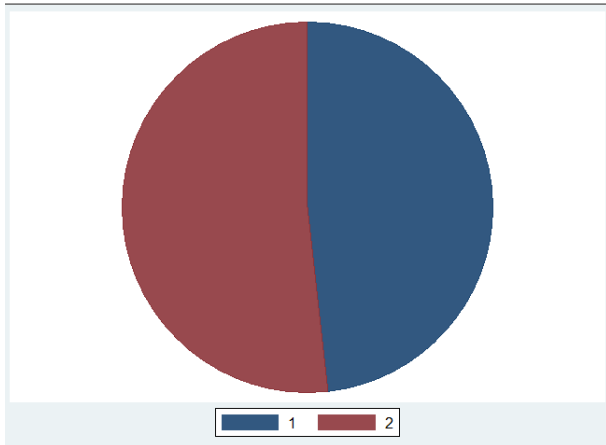
ENSU_CS database

Gender data were obtained from the ENSU_CS database (Page 15). The tabular command sex was applied to know the number of people in the survey by gender, as shown below:

```
. tab sexo
```

SEXO	Freq.	Percent	Cum.
1	36,585	48.24	48.24
2	39,250	51.76	100.00
Total	75,835	100.00	

A pie chart was generated for the sex variable:

Figure 14 Pie chart for the variable "gender"
Source: Self elaboration, 2022

To read the graph, the database structure file was consulted and the required nomenclature was found. See **Figure 15.**

Figure 15 Data for categories
Source: Self elaboration, 2022

As a result of the tabulation and the graph, the data shows that 48.24% of the respondents were male and 51.76% female.

ENSU_C database

From the ENSU_C database (Page 19), question 1.2: In terms of crime, tell me if in (PLACE) you feel safe or unsafe. The options are shown in **Table 1**. with their keys or mnemonics were. Possible answers are shown in **Table 2**.

Question	Mnemonic
Your home	BP1_2_01
Your work	BP1_2_02
The streets you usually use	BP1_2_03
Your school	BP1_2_04
The market	BP1_2_05
The shopping mall	BP1_2_06
The bank	BP1_2_07
The ATM located on the public street	BP1_2_08
Public transportation	BP1_2_09
The automobile	BP1_2_10
The road	BP1_2_11
Recreational park or recreation center	BP1_2_12

Table 1 Response options for ENSU question 1.2
Source: Self elaboration, 2022

1	Secure
2	Insecure
3	Not applicable
9	Don't know/no answer

Table 2 Response options for ENSU question 1.2
Source: Self elaboration, 2022

All the variables in the previous question were tabulated in frequencies:

. tab bp1_2_01

```
BP1_2_01 |   Freq.    Percent     Cum.
------------+-----------------------------------
     1 |  60,489    79.76     79.76
     2 |  15,320    20.20     99.97
     9 |      26     0.03    100.00
------------+-----------------------------------
Total |  75,835   100.00
```

. tab bp1_2_02

```
 BP1_2_02 |   Freq.    Percent     Cum.
------------+-----------------------------------
     1 |  37,725    49.75     49.75
     2 |  16,268    21.45     71.20
     3 |  21,783    28.72     99.92
     9 |      59     0.08    100.00
------------+-----------------------------------
 Total |  75,835   100.00
```

. tab bp1_2_03

```
 BP1_2_03 |   Freq.    Percent     Cum.
------------+-----------------------------------
     1 |  32,089    42.31     42.31
     2 |  43,111    56.85     99.16
     3 |     565     0.75     99.91
     9 |      70     0.09    100.00
------------+-----------------------------------
 Total |  75,835   100.00
```

. tab bp1_2_04

```
 BP1_2_04 |   Freq.    Percent     Cum.
------------+-----------------------------------
     1 |   4,359     5.75      5.75
     2 |     972     1.28      7.03
     3 |  69,859    92.12     99.15
     9 |     645     0.85    100.00
------------+-----------------------------------
 Total |  75,835   100.00
```

. tab bp1_2_05

BP1_2_05	Freq.	Percent	Cum.
1	34,007	44.84	44.84
2	32,886	43.37	88.21
3	8,783	11.58	99.79
9	159	0.21	100.00
Total	75,835	100.00	

. tab bp1_2_06

BP1_2_06	Freq.	Percent	Cum.
1	46,504	61.32	61.32
2	24,695	32.56	93.89
3	4,497	5.93	99.82
9	139	0.18	100.00
Total	75,835	100.00	

. tab bp1_2_07

BP1_2_07	Freq.	Percent	Cum.
1	25,961	34.23	34.23
2	39,941	52.67	86.90
3	9,784	12.90	99.80
9	149	0.20	100.00
Total	75,835	100.00	

. tab bp1_2_08

BP1_2_08	Freq.	Percent	Cum.
1	16,606	21.90	21.90
2	47,948	63.23	85.12
3	11,109	14.65	99.77
9	172	0.23	100.00
Total	75,835	100.00	

. tab bp1_2_09

BP1_2_09	Freq.	Percent	Cum.
1	19,139	25.24	25.24
2	38,665	50.99	76.22
3	17,807	23.48	99.70
9	224	0.30	100.00
Total	75,835	100.00	

. tab bp1_2_10

BP1_2_10	Freq.	Percent	Cum.
1	46,868	61.80	61.80
2	23,726	31.29	93.09
3	5,138	6.78	99.86
9	103	0.14	100.00
Total	75,835	100.00	

. tab bp1_2_11

BP1_2_11	Freq.	Percent	Cum.
1	33,174	43.74	43.74
2	36,506	48.14	91.88
3	5,962	7.86	99.75
9	193	0.25	100.00
Total	75,835	100.00	

. tab bp1_2_12

BP1_2_12	Freq.	Percent	Cum.
1	32,063	42.28	42.28
2	29,708	39.17	81.45
3	13,793	18.19	99.64
9	271	0.36	100.00
Total	75,835	100.00	

Labeling of variables

The above tables show that the data do not have labels to be able to read the data easily, so the document containing the INEGI data structure (2022) was consulted and the labels corresponding to the variables were placed.

. label variable bp1_2_01 "su casa"

. label variable bp1_2_02 "su trabajo"

A frequency table was then generated to test whether the command worked or not:

. tab bp1_2_01

su casa	Freq.	Percent	Cum.
1	60,489	79.76	79.76
2	15,320	20.20	99.97
9	26	0.03	100.00
Total	75,835	100.00	

The command was tested to see if it had worked and the labels continued to be applied to each variable of the chosen question. See **Figure 16** and syntax:

. label variable bp1_2_03 "las calles que habitualmente usa"

. label variable bp1_2_04 "la escuela"

. label variable bp1_2_05 "el mercado"

. label variable bp1_2_06 "el centro comercial"

. label variable bp1_2_07 "el banco"

. label variable bp1_2_08 "el cajero automático localizado en la vía"

. label variable bp1_2_08 "el cajero vía pública"

. label variable bp1_2_09 "transporte público"

. label variable bp1_2_10 "automóvil"

. label variable bp1_2_11 "la carretera"

. label variable bp1_2_09 "el transporte público"

. label variable bp1_2_10 "el automóvil"

. label variable bp1_2_12 "el parque recreativo o centro recreativo"

. label variable bp1_2_12 "el parque recreativo"
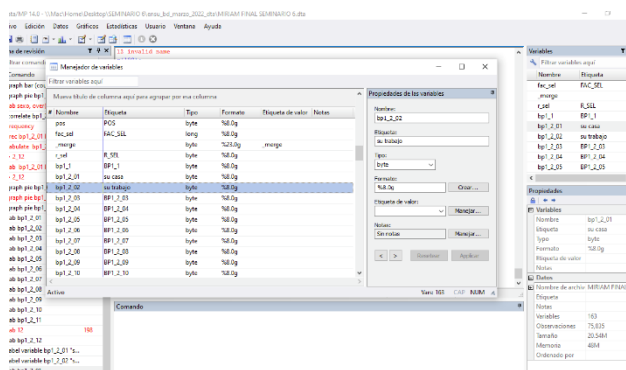
Figure 16 Variable handler for manual editing
Source: Self elaboration, 2022.

Based on the question to be solved, and knowing the names of the variables, it was decided to work only with the variable bp1_2_12, which is the one that corresponds to parks; the other variables of the question were discarded for this work.

Category labeling

The next step was to label the categories of the variables for the study in order to identify what each response refers to. In Stata, the commands label define and label values were used to obtain the corresponding labels, as shown in **Figure 17**, **Figure 18** and **Figure 19**.

. use "\\Mac\Home\Desktop\SEMINARIO 6\ensu_bd_marzo_2022_dta\MIRIAM FINAL SEMINARIO 6.dta", clear

. label define bp1_2_12 1"seguro o segura" 2"inseguro o insegura" 3"no aplica" 9"no sabe o no responde"

. label values bp1_2_12 bp1_2_12

. tab bp1_2_12

```
tab bp1_2_12
```

el parque recreativo	Freq.	Percent	Cum.
seguro o segura	32,063	42.28	42.28
inseguro o insegura	29,708	39.17	81.45
no aplica	13,793	18.19	99.64
9	271	0.36	100.00
Total	75,835	100.00	

Figure 17 Table of frequencies for question 2.1 of the ENSU
Source: Self elaboration, 2022

The table above shows that 42.28% of respondents feel safe in recreational parks, 39.17% feel unsafe.

. label define sexo 1"Hombre" 2"Mujer"

. label values sexo sexo

. tab sexo

SEXO	Freq.	Percent	Cum.
Hombre	36,585	48.24	48.24
Mujer	39,250	51.76	100.00
Total	75,835	100.00	

Figure 18 Frequency table for the ENSU variable "gender"
Source: Self elaboration, 2022.

Regarding the gender variable, 48.24% of the respondents were male and 51.76% female.

. tab sexo bp1_2_12

SEXO	seguro o	el parque recreativo inseguro	no aplica	9	Total
Hombre	15,965	13,922	6,567	131	36,585
Mujer	16,098	15,786	7,226	140	39,250
Total	32,063	29,708	13,793	271	75,835

Figure 19 Frequencies of perception of safety in parks by gender
Source: Self elaboration, 2022.

A frequency table was also generated with the two variables, resulting in 15,965 men feeling safe in parks and 13,922 feeling unsafe, while 16,098 women feel safe in these spaces and 15,786 feel unsafe.

The above information at a glance and without inferential analysis, that there is a tendency for women to feel safe, but also unsafe in recreational parks.

Contingency tables

Contingency tables are useful to see relationships between two variables, so that the question posed for the present study can be answered: Does gender have an influence on the perception of safety or insecurity in parks?

tab sexo bp1_2_12, row

```
Key

   frequency
 row percentage
column percentage
```

SEXO	seguro o	el parque recreativo inseguro	no aplica	no sabe o	Total
Hombre	15,965	13,922	6,567	131	36,585
	43.64	38.05	17.95	0.36	100.00
	49.79	46.86	47.61	48.34	48.24
Mujer	16,098	15,786	7,226	140	39,250
	41.01	40.22	18.41	0.36	100.00
	50.21	53.14	52.39	51.66	51.76
Total	32,063	29,708	13,793	271	75,835
	42.28	39.17	18.19	0.36	100.00
	100.00	100.00	100.00	100.00	100.00

Figure 20 Contingencies with frequencies in percentages for the variable perception of safety in parks and gender
Source: Self elaboration, 2022.

When reading **Figure 20**, it can be seen that if the data were arranged differently, a different reading would be obtained, so to answer the question of whether gender has an influence on the perception of security or insecurity in parks, it is necessary to invert the data in the columns and rows.

The contingency table presented in **Figure 21** was generated, showing frequencies and percentages that give another reading, since it indicates that gender influences the perception of security in parks.

. tab sexo bp1_2_12, col row

```
Key

   frequency
 row percentage
column percentage
```

| el parque recreativo	SEXO Hombre	Mujer	Total
seguro o segura	15,965	16,098	32,063
	49.79	50.21	100.00
	43.64	41.01	42.28
inseguro o insegura	13,922	15,786	29,708
	46.86	53.14	100.00
	38.05	40.22	39.17
no aplica	6,567	7,226	13,793
	47.61	52.39	100.00
	17.95	18.41	18.19
no sabe o no responde	131	140	271
	48.34	51.66	100.00
	0.36	0.36	0.36
Total	36,585	39,250	75,835
	48.24	51.76	100.00
	100.00	100.00	100.00

Figure 21 *Contingencies with frequencies in percentages for gender variable with perception of safety in parks.*Source: Self elaboration, 2022.

Graphs

For this type of variables, the most convenient is to make a bar chart. A bar chart was generated by means of the syntax:

. graph bar (count), over(sexo) over(bp1_2_12)
. graph bar (count), over(sexo) over(bp1_2_12) ytitle(Frecuencias) title("Percepción de seguridad o inseguridad en parques por género")

In both cases the resulting graph is the same as shown in **Figure 22**.
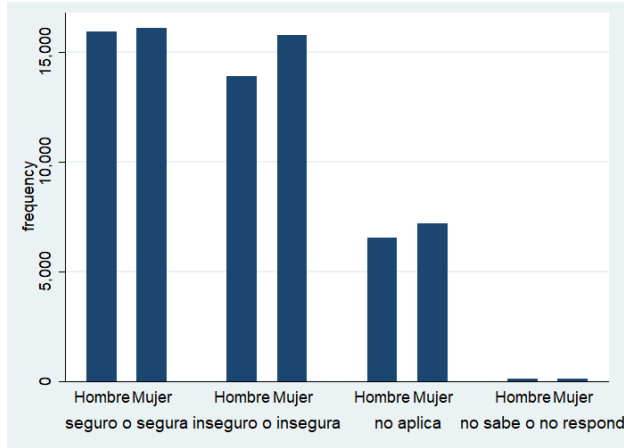
Figure 22 *Frequencies of gender and perception variables in parks*
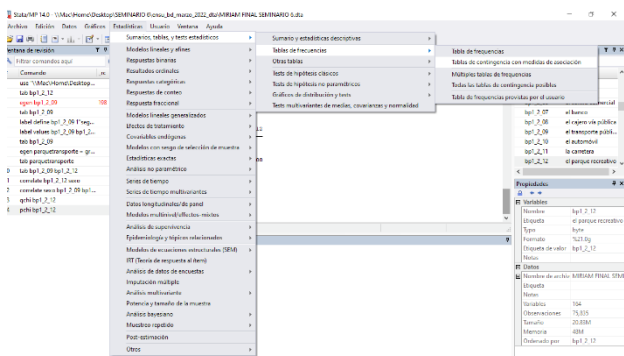*Source: Self elaboration, 2022*

Pearson's Chi-square test

This test evaluates, by means of a comparison of proportions, the association between two categorical or qualitative variables by means of observed versus expected frequencies.

In this case, two categorical variables were analyzed: sex and perception of safety in parks (bp_2_12).

The following path was followed in Stata. See **Figure 23** and **Figure 24**.

1. Statistics

2. Summaries, tables, and statistical tests

3. Frequency tables

4. Contingency tables with measures of association

Figure 23 Path for making contingency tables with measures of association
Source: Self elaboration, 2022

The null hypothesis was stated that both variables are independent. Alternate hypothesis, the variables were associated by obtaining a P value lower than P= 0.05, the null hypothesis was rejected, leaving the alternate hypothesis, indicating that the variables are dependent on each other, or that they show an association, as shown in **Figure 24** and **Figure 25**.
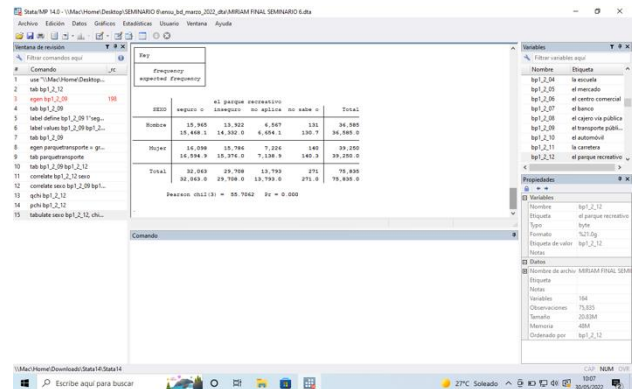
Figure 24 Pearson's Chi-square test for the variables "gender" and "perception of security or insecurity in parks"
Source: Self elaboration, 2022.

SEXO	el parque recreativo seguro o	inseguro	no aplica	no sabe o	Total
Hombre	15,965	13,922	6,567	131	36,585
	15,468.1	14,332.0	6,654.1	130.7	36,585.0
Mujer	16,098	15,786	7,226	140	39,250
	16,594.9	15,376.0	7,138.9	140.3	39,250.0
Total	32,063	29,708	13,793	271	75,835
	32,063.0	29,708.0	13,793.0	271.0	75,835.0

Pearson chi2(3) = 55.7062　Pr = 0.000

Figure 25 Pearson's Chi-square test
Source: Self elaboration, 2022

The differences between the proportions observed between the two variables are statistically significant, indicating that gender has an influence on whether a park is safe or unsafe for users.

Acknowledgements

ISSN 2524-2016
RINOE® All rights reserved

MERY-RUIZ, Miriam Elizabeth. Microdata management of ENSU in the study of urban public safety perceptions. Journal-Public Economy. 2023

Your steadfast encouragement and guidance have been pivotal, and your dedication to academic excellence remains a driving force behind this scholarly work.

Conclusions

The comprehensive analysis conducted utilizing Stata for data management proved to be highly beneficial and efficient. This advanced statistical software facilitated the merging of databases, and the creation of frequency and contingency tables, as well as the labeling of variables and categories. It also proved instrumental in generating relevant graphics and conducting statistical analyses tailored to the variable types. Through the detailed examination of three separate databases, this study focused on two specific variables from the ENSU: gender, as recorded in one database, and the availability of recreational parks, as recorded in another. The investigation sought to determine whether gender impacts individuals' perceptions of safety or insecurity in park areas.

The findings confirm that gender indeed plays a significant role in shaping perceptions of safety or insecurity in recreational parks. This outcome underscores the importance of considering gender perspectives in the development and management of urban recreational spaces to ensure they meet the safety expectations and needs of all users.

References

INE (2022) Instituto Nacional de Estadística de España. Accessed May 2022. Retrieved from: https://www.ine.es/prodyser/microdatos.htm

INEGI (2022) Instituto Nacional de Estadística Geografía e Informática. Encuesta Nacional de Seguridad Pública Urbana (ENSU). Retrieved from:
https://www.inegi.org.mx/programas/ensu/

INEGI (2022) Instituto Nacional de Estadística Geografía e Informática. Encuesta Nacional de Seguridad Pública Urbana 2022. ENSU - Primer trimestre. Estructura de la base de datos. 2022 Retrieved   from: https://www.inegi.org.mx/programas/ensu/

Serrano, R. M., y Valcarce, E. V. (2000). Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial (Vol. 44). Edicions Universitat Barcelona.

.