



# Ane ku mene

## Análisis e interpretación de datos: aportes de la información cuantitativa a los estudios sociales

### Data Analysis and Interpretation: Contributions of Quantitative Information to Social Studies

### Análise e interpretação dos dados: contribuições da informação quantitativa aos estudos sociais

Dalila Molina-Molina\*

#### Resumen

En la mayoría de los estudios sociales el investigador recopila datos cualitativos y cuantitativos, pero suele pasar por alto utilizar herramientas científicas, matemáticas, geométricas o estadísticas que permitan analizar, profundizar y correlacionar información que aporte conocimiento sustentado en cifras y otorgue rigurosidad científica al resultado de la investigación de un fenómeno. Por consiguiente, el presente artículo aborda el problema del dato y su capacidad para analizar e interpretar eventos, más allá de la mera sistematización en los estudios sociales. Plantea cómo la obtención y almacenamiento de la información en una situación hace parte de un ciclo, desde identificar elementos, establecer categorías, visualizar gráficos y cartografiar, hasta definir tendencias o comportamientos para diseñar, planear o tomar decisiones basadas en los resultados. El objetivo es explicar el significado y la presencia de los números en los análisis de corte social.

**Palabras clave:** análisis de información; cifras; correlación; datos; metodología; cualitativo; cuantitativo

\* Universidad Pedagógica Nacional.



## Abstract

In most social studies, the researcher collects qualitative and quantitative data, but often overlooks the use of scientific, mathematical, geometric, or statistical tools to analyze, deepen and correlate information that provides knowledge supported by figures and gives scientific rigor to the result of the investigation of a phenomenon. Therefore, this article addresses the problem of data and its capacity to analyze and interpret events, beyond mere systematization in social studies. It shows how obtaining and storing information in a situation is part of a cycle, from identifying elements, establishing categories, visualizing graphs and mapping, to defining trends or behaviors in order to design, plan or make decisions based on the results. The objective is to explain the meaning and presence of numbers in social analysis.

**Keywords:** information analysis; figures; correlation; data; methodology; qualitative; quantitative

## Resumo

Na maioria dos estudos sociais, o pesquisador coleta dados qualitativos e quantitativos, mas muitas vezes negligencia o uso de ferramentas científicas, matemáticas, geométricas ou estatísticas que permitem a análise, o aprofundamento e a correlação de informações que fornecem conhecimentos apoiados por números e dão rigor científico ao resultado da investigação de um fenômeno. Este artigo aborda, portanto, o problema dos dados e sua capacidade de analisar e interpretar eventos, além da mera sistematização em estudos sociais. Mostra como obter e armazenar informações em uma situação é parte de um ciclo, desde identificar elementos, estabelecer categorias, visualizar gráficos e mapeamentos, até definir tendências ou comportamentos para projetar, planejar ou tomar decisões com base nos resultados. O objetivo é explicar o significado e a presença dos números na análise social.

**Palavras-chave:** análise de informação; números; correlação; dados; metodologia; qualitativo; quantitativo



## Enfoques cualitativos y la correlación de datos para los estudios sociales

La descripción de fenómenos a partir de la investigación requiere del análisis de la información y de cifras que permitan explorar nuevos terrenos y obtener resultados para demostrar el comportamiento de una situación o problema. En la investigación, entendida como un cúmulo de saberes que pueden provenir de diferente naturaleza pero que guardan un mínimo de similitud, es el conocimiento del fenómeno el que proporciona la metodología adecuada, el que define las fuentes de datos y la forma en que se recopilan y procesan para obtener resultados veraces.

La hipótesis de este artículo centra su atención en que toda la información recopilada de una investigación requiere de rigurosidad científica para dar validez, confiabilidad y utilidad a los resultados obtenidos. En los estudios sociales podría considerarse que por tratarse de enfoques cualitativos estos están exentos del análisis matemático, geométrico o estadístico que pondere, represente, grafique, mida, correlacione y evalúe el comportamiento de los datos. Sin embargo, implementar análisis de este tipo fundamenta el resultado sin necesidad de desviar el enfoque y la naturaleza de la investigación o las convicciones del investigador.

Por consiguiente, el análisis de datos de orden cuantitativo debe ser una herramienta tenida en cuenta en los ejercicios de investigación social, la cual puede complementarse con instrumentos como las entrevistas, las encuestas, los grupos focales o los ejercicios de observación. De ese modo, los resultados obtenidos en las investigaciones tendrán un soporte técnico que robustecerá el estudio y permitirá comprender todas las dimensiones del fenómeno abordado.

## La interpretación de los datos y el comportamiento del fenómeno

El análisis de cualquier tipo de información requiere de datos, ya sea de naturaleza cualitativa o cuantitativa. Definir el objeto de estudio de una investigación permite establecer el tipo de datos requeridos, el método de recolección, la fuente de la cual obtenerlos y las variables a analizar. El enfoque

derivará del análisis y la observación del fenómeno y en el caso de los hechos de tipo histórico, sin llegar a controlar su comportamiento, pretenderá comprenderlos. Así, el enfoque de la investigación podría depender del criterio del investigador.

Las estrategias por las que se opta en una investigación no se definen de forma aislada, sino, en términos de su capacidad para resolver problemas científicos y con relación a otras estrategias o métodos que se formulan para resolver ese problema, en busca de la mejor aproximación posible para su comprensión [...]. De tal suerte que la decisión por una estrategia o técnica de investigación cuantitativa o cualitativa, o una combinación de las dos, es una decisión pragmática que se toma según la conveniencia de un método particular, para una circunstancia también particular”. (Páramo, 2011, p. 29)

Las circunstancias particulares del fenómeno permiten caracterizarlo, identificar sus componentes y discernir la dependencia mutua que guardan entre sí. Históricamente el análisis de datos para demostrar o tomar una decisión se basa en la interpretación del resultado matemático y en la lectura de parámetros previamente evaluados. Así, los datos recopilados se convierten en conocimiento previo del fenómeno que se estudia, los cuales pueden ser inmediatos o históricos. Los datos inmediatos se encuentran procesados, cuentan con instrumentos de control y se toman de fenómenos que son recreables y repetibles. Los datos históricos, por su parte, se refieren a registros de fenómenos cosmológicos, globales, sociológicos o históricos y deben ser revisados basándose en hechos.

Los datos pueden caracterizarse por ser variables o constantes, por compararse temporal o espacialmente, por correlacionarse para indicar la dirección, proporción o nivel en que una variable se relaciona con otra. Permiten poner en relación mutua dos o más cosas y calcular e interpolar cifras para una lectura concreta del fenómeno.

La clasificación de los datos se establece en constantes, variables dependientes y variables independientes. Hernández (2012) argumenta que “la constante es un dato de referencia y la variable es un dato característico de estudio” (p. 16). Los datos, cuyas variables guardan relación con los resultados del fenómeno, se tabulan para visualizar las características significativas, para ponderarlas, evaluarlas o estudiarlas. Con el fin de verificar si los datos guardan relación entre ellos, el método de medición usado y los datos obtenidos (variables que



explican el fenómeno) deben ser probados mediante la herramienta estadística de la correlación.

Es importante comprender que la observación directa de datos es insuficiente para sacar conclusiones respecto a los mismos. Hernández (2012) afirma que “utilizando las técnicas de estadística descriptiva, incluso las más elementales, podemos describir el comportamiento” p. 13). La utilidad dependerá de tabular, es decir, de ordenar y agrupar los datos en categorías para construir una gráfica que facilite el entendimiento del comportamiento estudiado.

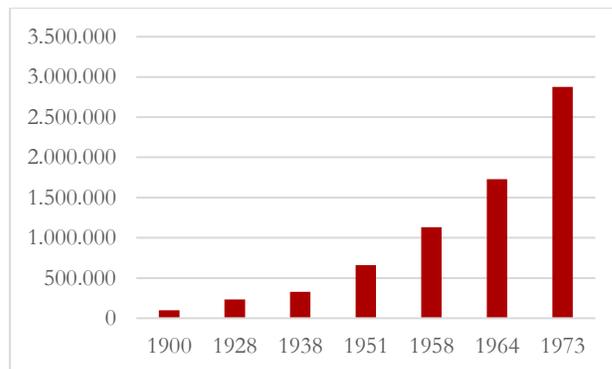
**Tabla 1.** Población de la ciudad de Bogotá entre 1900 y 1973

Años	Población	Área (ha)	Densidad
1900	100.000	909	110
1928	235.000	1.958	120
1938	330.000	2.514	131
1951	660.000	4.818	137
1958	1.130.000	8.084	140
1964	1.730.000	14.615	118
1973	2.877.000	30.423	95

Fuente: elaboración propia a partir de datos del DANE.

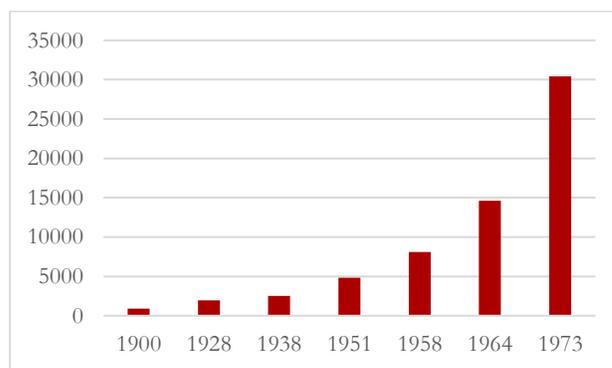
Por ejemplo, en el marco de una investigación histórica se tabularon datos obtenidos del Departamento Administrativo Nacional de Estadística (DANE) respecto a población y área construida en Bogotá, entre los años 1900 a 1973. En la tabla se observa un vínculo proporcional entre las variables *Población* y *Área*, las cuales definen el crecimiento de la ciudad, y el indicador *Densidad* que expresa la relación entre las dos variables (ver tabla 1).

Dado que hay intervalos de tiempo diferentes entre los datos tabulados, estos deben homogenizarse para visualizar correctamente la tendencia de la curva obtenida. Para ayudar a la construcción de la gráfica en el ejemplo presentado, se calculó el dato *Área construida* del año 1951, mediante interpolación entre los datos de los años adyacentes, con el propósito de empalmar la curva esperada. La interpolación es una técnica que permite calcular un valor intermedio entre dos o más datos conocidos y se utiliza cuando se requiere estimar una variable en una zona en la que se carece de la información del dato requerido.



**Figura 1.** Población de la ciudad de Bogotá entre 1900 y 1973

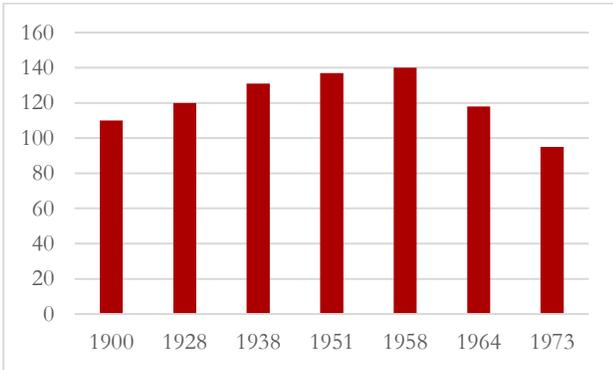
Fuente: elaboración propia.



**Figura 2.** Área de la ciudad de Bogotá entre 1900 y 1973

Fuente: elaboración propia.

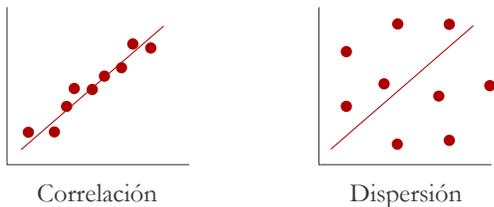
Al graficar los datos de la tabla y espacializar en diagramas o mapas se obtuvieron interpretaciones veraces, objetivas y accesibles del comportamiento o las tendencias del fenómeno. Controlar el dato o hacerlo homogéneo permite representar un punto intermedio en un gráfico de plano cartesiano facilitando la lectura de la curva. Así, en el caso presentado, las gráficas de *Población* y *Área* muestran comportamientos similares (ver figuras 1 y 2), mientras que la *Densidad poblacional* mantiene valores cercanos a lo largo de los diferentes años, comportándose como una constante para el periodo estimado (ver figura 3).



**Figura 3.** Densidad poblacional de la ciudad de Bogotá entre 1900 y 1973

Fuente: elaboración propia.

El cálculo estadístico de los datos respecto al comportamiento del desarrollo urbano de la ciudad de Bogotá muestra una correlación estadística fuerte entre las variables relacionadas. No obstante, si los resultados de medición del fenómeno hubiesen presentado una elevada dispersión, significaría que se midió la variable incorrecta o se utilizó el método inadecuado para resolver el objeto de estudio. Así, en el ejemplo, las figuras muestran una población en aumento a la par del área construida y la gráfica de densidad poblacional valida la relación entre variables.

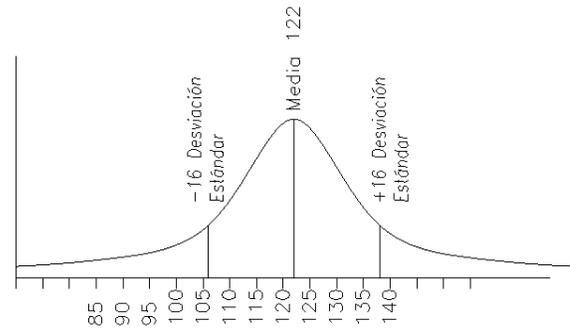


**Figura 4.** Correlación y dispersión de datos

Fuente: elaboración propia.

La densidad promedio en Bogotá, según los datos suministrados por la fuente consultada, indica que la media es de 122 habitantes por hectárea, con una desviación estándar de 16 habitantes en el periodo inspeccionado. Al representar el resultado de manera gráfica en la campana de Gauss (ver figura 5), se visualiza la media y la distribución del grupo de datos

que se limita entre dos fronteras denominadas desviación estándar. En este caso, los datos típicos se agrupan en un porcentaje del 68 % y los atípicos se encuentran por fuera de este rango. Así, los datos que están por fuera de la desviación estándar son atípicos y pueden ocasionarse por errores atribuibles al instrumento de medición empleado, al modelo asumido, al proceso o procedimiento o a la toma de la muestra que pudo o no ser controlada.



**Figura 5.** Datos de la población de Bogotá entre 1900 y 1973 (campana de Gauss)

Fuente: elaboración propia.

En el ejemplo abordado, el comportamiento de los datos en el gráfico de desviación estándar y media (promedio) explica la eventual transformación urbana de sectores consolidados en la ciudad, como es el caso de la transformación del sector del Chicó en Bogotá. Cuando se proyectó la urbanización en 1950, se determinó que para el tamaño poblacional se requerían casas de uno y dos pisos; sin embargo, cuando el sector inició el proceso de transformación espacial, en la década de los 80, al sustituir las casas iniciales por nuevos edificios, aumentó la población y cambió la ocupación del suelo y su densidad.

Wiesner (1980), en su análisis espacial de Bogotá, registra la tasa de crecimiento de población y urbanización de la ciudad y del país en general y, a partir de ello, evidencia la expansión de las zonas urbanas. Constata, a partir de datos y tabulaciones, que “en el proceso de urbanización en Colombia, Bogotá ha crecido constantemente de una manera más rápida que otras ciudades de menores categorías en las que también se concentró la población, como Cali, Medellín y Barranquilla” (Wiesner, 1980, p. 17). Para el autor, la consecuencia del



aumento de trabajo en lo urbano atrajo consigo altas tasas de urbanización, no obstante, “la migración hacia las zonas urbanas continuó al ritmo acostumbrado, en parte debido a la violencia en las regiones rurales de Colombia” (Wiesner, 1980, p. 19).

Por consiguiente, la lectura de datos en los que la densidad varía es un indicador para analizar el fenómeno específico en términos históricos, sociales, demográficos, políticos o de intereses en particular. Tomar como elementos de análisis los momentos en los que varía el comportamiento de la media es relevante para obtener hallazgos. El resultado de una investigación debe permitir explicar la ocurrencia de un fenómeno y en algunos casos anticiparse a los hechos, esto cuando el fenómeno es claro, calculable y predecible. Las cifras provienen de registros y observaciones recopiladas por las diferentes fuentes, disciplinas, ciencias y oficios con intereses particulares.

## Naturaleza y tipos de los datos: categóricos y numéricos

Los datos son descripciones básicas de cosas, eventos y actividades, sensibles de ser registrados, clasificados o sistematizados. Son susceptibles de utilizar, manipular y representarse por diferentes métodos de análisis que los convierten en instrumentos o herramientas de decisiones muestrales para llegar a resultados y dar soluciones a problemas.

El diseño o planeación de la toma de una decisión o su implementación derivan del ciclo de análisis del dato. La información corresponde a datos organizados que interpretan el significado y otorga valor para quien la recibe. Las bases de datos, los datos entre sí y su sistematización ofrecen la posibilidad de relacionarlos, ordenarlos y correlacionarlos.

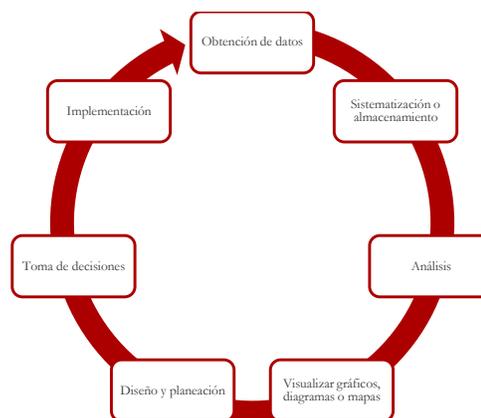


Figura 6. Ciclo de análisis del dato

Fuente: elaboración propia.

Una muestra es el número de datos que se toma para analizar. Cuanto menor es el tamaño de la muestra, mayor es el error en el cálculo de la probabilidad y, en consecuencia, mayor el valor de la falta de certeza. La probabilidad mide la mayor o menor posibilidad de que se dé un determinado resultado cuando se realiza una comprobación. Cuanto mayor es el número de datos más confiable es el estudio que se realiza y más pequeño será el margen de error.

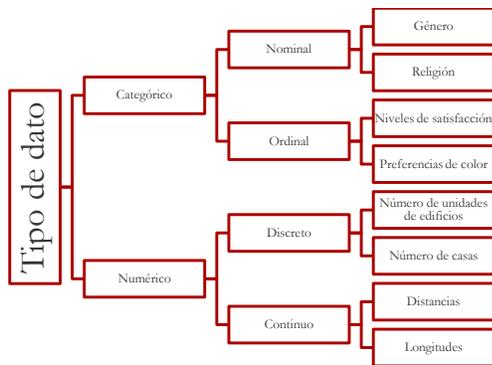
Los datos son de diversa fuente. Por su localización específica de acuerdo con un sistema cartográfico de referencia serían geográficos. En general, los datos son de naturaleza alfanumérica. Los datos cualitativos o categóricos tratan con características y descripciones difíciles de medir, pero que se pueden observar subjetivamente: olor, sabor, textura, atractivo, color. Los datos categóricos aparecen cuando una variable se mide en una escala que sólo incluye a los posibles encuestados en un número limitado de grupos: una encuesta donde se recoge información sobre variables como el género, estado civil o afiliación política. Las variables también se pueden clasificar como independientes o dependientes. (Marín, s.f., p. 1)

Los datos cualitativos se clasifican en categóricos y numéricos. Los datos categóricos, a su vez, se clasifican en nominales y ordinales y los numéricos en discretos y continuos. Los datos son piezas de información real que se recopilan mediante su estudio y según sus atributos definen el tipo de variable. De esta forma, las variables se clasifican en categóricas y continuas, en las primeras la naturaleza de sus datos es



cualitativa y en las segundas cuantitativa. Una variable categórica recoge información sobre el género, el estado civil o la afiliación política, mientras que una continua compila información que trata de números y cosas que se pueden medir objetivamente.

Los datos numéricos son continuos o discretos. Los recuentos son discretos y las mediciones son continuas. Los datos discretos son un conteo que no se puede hacer más preciso [...]. Implica números enteros e indivisibles: número de unidades de edificios o casas. Los datos continuos pueden dividirse y subdividirse (metros, centímetros y milímetros). La distancia o las longitudes pueden ser un dato continuo. (Hernández, 2012, p. 16)



**Figura 7.** Tipo de dato

Fuente: elaboración propia.

## Aplicación de análisis de datos

A continuación, con el propósito de clarificar el proceso de análisis de datos, se abordará una situación relacionada con la transformación urbana que ha experimentado el sector del Chicó y Santa Bárbara en Bogotá desde 1980. Para tal fin, se analizará una casa entre edificios que cumple con las tipologías estándar del sector, alrededor de la cual se encuentran edificios construidos y otros en proceso de construcción.



**Figura 8.** Fotografía del casalote 8

Fuente: elaboración propia.



**Figura 9.** Fotografía del casalote 8

Fuente: elaboración propia.

Las fotografías tomadas en marzo de 2018 corresponden al lote 8, el cual cuenta con un área de 758 metros cuadrados. La diferencia del área total ocupada por construcción del lote número 8 respecto al área total de los predios 10 y 6 revela un dato atípico por fuera de la desviación estándar. Esto indica que la casa debe igualar el porcentaje de ocupación y de densidad equivalente a las edificaciones inmediatas y asumir en proporción, por ejemplo, toda la distribución de las cargas tributarias que aportan las unidades individuales de los edificios vecinos. La casa tiene un área mayor de construcción y ocupa un porcentaje mayor del lote respecto a una unidad dentro de un edificio, por lo que el valor de la carga tributaria para la casa aumenta.

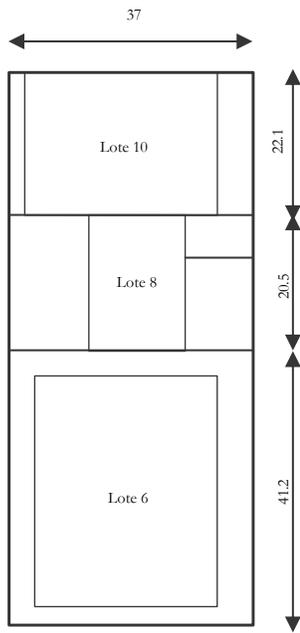


Figura 10. Distribución de los lotes 8, 6 y 10

Fuente: elaboración propia.

La tabla 2 muestra los metros cuadrados y el área total ocupada por la construcción de los lotes 10, 8 y 6. Por su parte, en el gráfico se observa la relación del área total del predio respecto al área total ocupada por construcción y la distribución porcentual comparativa por cada una de las edificaciones. En el lote 8 el área total ocupada por la construcción es inferior al área total del predio, mientras que en los lotes 10 y 6 el área total ocupada de construcción es mayor al área total del predio, lo que explica por qué la carga tributaria predial es alta obligando a sustituir la casa por un edificio similar a los vecinos inmediatos.

Tabla 2. Área total del predio y área total ocupada por construcción en metros cuadrados

Urbanización Santa Bárbara, Manzana 20			
	Lote 10	Lote 8	Lote 6
Área total del predio	822	758.5	1.500
Área total ocupada por construcción	3.600	324	6.000

Fuente: elaboración propia.

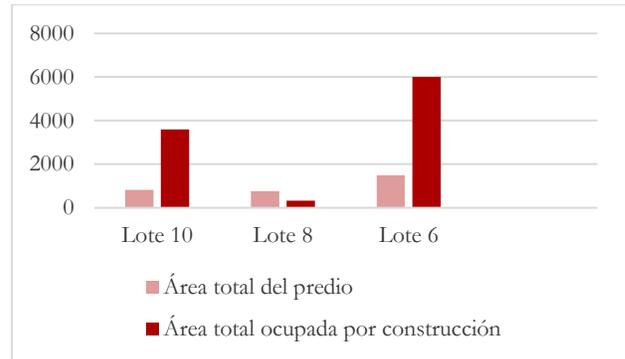


Figura 11. Área total ocupada y área construida

Fuente: elaboración propia.

En este caso, el número de lotes es el dato discreto (lote 10, lote 8 y lote 6) y las dimensiones del área total del predio y el área total ocupada por construcción es la información continua. El análisis para estimar la relación entre las variables es igual al área total del predio (datos continuos), correlacionada con el área de ocupación en el primer piso. El área total es proporcional entre sí a los tres predios, lo que significa que el potencial desarrollo de construcción para el predio número 8 es similar al área real construida en los predios 6 y 10 respectivamente. Así, se confirma que las cargas y obligaciones siempre serán equivalentes al área del predio.

En los datos presentados en la tabla 3 se observan relaciones inversamente proporcionales entre área libre y área construida; además, en el índice de ocupación, el área total del predio es mayor al área ocupada en el primer piso y el área libre respecto al porcentaje de ocupación disminuye si aumenta el área de construcción. El análisis para estimar la relación entre las variables, igual al porcentaje de área libre con el porcentaje de área ocupada (datos continuos), está correlacionado con el área total del predio y con el área de ocupación del primer piso (datos discretos). Por consiguiente, el análisis de este caso permite determinar que las casas tradicionales desaparecerán toda vez que la proporción de cargas y obligaciones caen sobre el desarrollo futuro potencial del predio, a la par del hecho de que la norma cuantifica el desarrollo de los terrenos mediante análisis similares para tomar decisiones de planeación y desarrollo futuro de una ciudad.



**Tabla 3.** Índice de ocupación, área libre por lote y área total construida

Índice de ocupación			
	Lote 10	Lote 8	Lote 6
Área total del predio	822	758.5	1.500
Área de ocupación primer piso	600	324	1.000
Área libre por lote y área total construida			
	Lote 10	Lote 8	Lote 6
Porcentaje de área libre	27%	57.3%	33.3%
Porcentaje de área ocupada	73%	42.7%	66.7%

Fuente: elaboración propia.

En este sector, la tipología dominante explica la desaparición de las casas y su sustitución por edificios. La casa es el dato disperso o atípico que tiende a convertirse en edificio para que, como dato, se agrupe con otros para normalizar la muestra. La variación en los análisis de ocupación del suelo y área construida se calcula sobre el área del lote, sin embargo, distribuir cualquier valor por el área construida es diferente a distribuir el valor por el área del lote. El área del lote es la información con datos constantes y permite identificar el tipo dominante y la tendencia esperada una vez cambie el tipo edificatorio.

El área construida de los predios debería presentar valores continuos. Cuando un predio parcialmente ocupado aparece dentro de un sector con densidad ocupacional, este es absorbido por la tipología. En términos tributarios, la carga se calcula sobre el área del terreno y esta es absorbida en proporción por el tipo edificatorio dominante del sector, aumentando el valor de la carga para los predios ocupados con baja densidad respecto a los predios densamente ocupados por la construcción.



**Figura 12.** Tipología edificatoria de la manzana

Fuente: elaboración propia.

La fotografía, que corresponde a la tipología edificatoria de manzana, se tomó en marzo de 2018. Los dos edificios ubicados sobre la Carrera 19 con Calle 96, sentido sur norte, muestran una tipología edificatoria distinta caracterizada por la altura máxima permitida, toda vez que los predios tienen un área de superficie similar. La altura de la nueva edificación constituye el futuro desarrollo para el edificio que se observa de ladrillo y que en menos de dos décadas, paradójicamente, aumentó su capacidad edificatoria o de densidad.



## Conclusiones

Los resultados estadísticos en las ciencias sociales y en disciplinas diferentes a la financiera o matemática, brindan por su naturaleza descriptiva una herramienta adecuada para el estudio de fenómenos en sus ámbitos sociales, normalmente complejos y enmarcados en un contexto variable. Toda investigación cualitativa tiene un trasfondo cuantitativo que el investigador minimiza en la mayoría de los casos, descartando el análisis de la información sin prestar atención a su potencial.

La muestra procede de un subconjunto representativo de la población y la variable de una característica de esta. Por su naturaleza, la variable puede ser cuantitativa o cualitativa, por lo que la población, la muestra y la variable pueden provenir de distinta naturaleza, espacio o tiempo. Tanto las variables cualitativas como las cuantitativas están sujetas a ser analizadas con expresiones nominales, ordinales y continuas, siempre susceptibles a ponderarse, medirse o evaluarse.

En las variables cuantitativas el resultado de una expresión numérica puede convertirse en la representación de un atributo gráfico expresado en un plano, el cual permite lecturas objetivas del comportamiento de un evento. El empleo de modelos de organización y análisis de la información, por ejemplo en estudios sociales como los de la ciudad, permite validar comportamientos o resultados representativos, reunir datos, sistematizar, graficar y representar atributos en planos, imágenes o mapas. Los resultados analizan ciertos patrones, visualizan tendencias, probabilidad o comportamiento de eventos para obtener diagnósticos y tomar decisiones, inferir o hacer hipótesis de futuros resultados.

Los estudios sociales usualmente toman distancia del hecho cuantitativo por considerarlo propio de las ciencias exactas y

exentas de ello a las ciencias sociales. El análisis de los datos, lejos de pretender convertir los estudios sociales en enfoques cuantitativos, aporta herramientas para fundamentar y explicar con argumentos un fenómeno y para profundizar en las interpretaciones de orden social. Así, es la naturaleza de los datos la que puede ser cuantitativa o cualitativa.

De esta forma, el análisis de datos de orden cuantitativo debe considerarse en los ejercicios de investigación social ya que complementa instrumentos como las entrevistas, las encuestas, los grupos focales o los ejercicios de observación. Con ello, los resultados obtenidos en las investigaciones tendrán un soporte técnico que robustecerá el estudio y que permitirá comprender las múltiples dimensiones del fenómeno que se aborda mediante un marco metodológico definido.

## Referencias

- Hernández, Z. (2012). *Métodos de análisis de datos: apuntes*. La Rioja: Universidad de la Rioja.
- Marín, J. (s.f.). Tema 1. Introducción al análisis de datos categóricos [documento en línea]. <https://cutt.ly/snMBNWx>
- Páramo, P. (Ed.). (2011). *La investigación en Ciencias Sociales: estrategias de investigación*. Universidad Piloto de Colombia.
- Wiesner, G. (1980). *Cien años de desarrollo histórico de los precios de la tierra en Bogotá*. Cámara de comercio de Bogotá.