

ESTADÍSTICAS E INDICADORES EN LA GESTIÓN PÚBLICA

Aproximación teórica y práctica

David Santiago Hernández Ardila
Alberto Rodríguez Rodríguez



Dirección Nacional de Planeación y Estadística
Oficina Nacional de Estadística

Estadísticas e indicadores en la gestión pública

Aproximación teórica y práctica

David Santiago Hernández Ardila

Alberto Rodríguez Rodríguez

Oficina Nacional de Estadística

Bogotá. D.C. – Colombia

2023

Universidad Nacional de Colombia
Dirección Nacional de Planeación y Estadística



**Estadísticas e indicadores
en la gestión pública**
Aproximación teórica y práctica

© Universidad Nacional de Colombia
© David Santiago Hernández Ardila
© Alberto Rodríguez Rodríguez

DIRECTOR
José Ignacio Maya Guerra

EQUIPO DE TRABAJO
Alberto Rodríguez Rodríguez
Basilio Sánchez Manrique
Cindy Catalina Vásquez Guerrero
Cindy Lorena Silva Montenegro
Jakeline Burgos Gutiérrez
Juan Carlos Iguavita Flórez
Katherin Marian Manrique Rodríguez
Liliana Andrea Gaitán Mendoza
Luis Felipe Sánchez Mesa
María Claudia Galindo González
Mónica Andrea Carrillo Rojas
Nelson Enrique Vivas Velandia

CORRECCIÓN DE ESTILO
María José Díaz Granados M.

DISEÑO CARÁTULA
Andrés Álvarez Ríos

DISEÑO GRÁFICAS
Andrés Álvarez Ríos

DISEÑO Y DIAGRAMACIÓN IMPRESA
Gloria Diazgranados M.

DISEÑO Y DIAGRAMACIÓN WEB
María Jimena López Munévar

IMPRESIÓN
Editorial Gente Nueva

ISSN: 978-958-505-424-0

RECTORA
Dolly Montoya Castaño

VICERRECTOR GENERAL
Pablo Enrique Abril

VICERRECTOR ACADÉMICO
Carlos Augusto Hernández Rodríguez

VICERRECTOR INVESTIGACIÓN
Camilo Younes Velosa

SECRETARÍA GENERAL
María Fernanda Lara Díaz

GERENTE NACIONAL FINANCIERO Y ADMINISTRATIVO
Álvaro Viña Vizcaíno

DIRECTOR NACIONAL DE PLANEACIÓN Y ESTADÍSTICA
José Ignacio Maya Guerra

DIRECTORA DE RELACIONES EXTERIORES
Melba Libia Cárdenas Beltrán

**DIRECTOR UNIDAD DE MEDIOS DE
COMUNICACIÓN - UNIMEDIOS**

Fredy Chaparro Sanabria

DIRECTOR OFICINA JURÍDICA NACIONAL
Jairo Iván Peña Ayazo

DIRECTORA NACIONAL DE VIGILANCIA DISCIPLINARIA
Diana Yamile Rodríguez Niño

DIRECTORA NACIONAL FONDO PENSIONAL
Carolina Argüello Ospina

JEFES OFICINA NACIONAL DE CONTROL INTERNO
Ángel de Jesús Múnera

VICERRECTOR SEDE BOGOTÁ
José Ismael Peña Reyes

VICERRECTOR SEDE MEDELLÍN
Juan Camilo Restrepo Gutiérrez

VICERRECTOR SEDE MANIZALES
Neil Guerrero González

VICERRECTOR SEDE PALMIRA
Jaime Eduardo Muñoz Flórez

VICERRECTORA SEDE DE LA PAZ
María Marcela Camacho Navarro

DIRECTORES DE SEDE

ORINOQUÍA: Óscar Eduardo Suárez Moreno

AMAZONIA: Eliana María Jiménez Rojas

CARIBÉ: Adriana Santos Martínez

TUMACO: Amanda Lucía Mora Martínez

JEFES OFICINAS DE PLANEACIÓN Y ESTADÍSTICA

SEDE BOGOTÁ: Geraldo Millán Cuervo

SEDE MEDELLÍN: Iván Alonso Montoya Restrepo

SEDE MANIZALES: Paula Andrea Escandón Suárez

SEDE PALMIRA: Nelson Lozano Duque

Presentación	9
Introducción	13
Capítulo I	
Información y gestión pública	15
Información	15
Definición de información	17
Campos de la información	22
Gestión pública	23
Gestión	23
Público	24
Gestión pública	25
Usos de la información en la gestión pública	28
Uso comunicativo	28
Uso de registro	29
Uso valorativo	31
Uso de procesamiento	33
CAPÍTULO II	
Las estadísticas	35
Definición	36
Características intrínsecas	40
Población o muestras	40
Datos	43
Cifras agregadas	45
Desagregaciones	46
Representación gráfica y tabular	48

Mecanismos de divulgación y de comunicación	49
Características extrínsecas	51
Incluyentes	51
Comparables	53
Legales y éticas	54
Públicas, transparentes y abiertas	56
Uso intensivo de las TIC	58
Metadatos	60
Proceso estadístico	60
De la teoría a la práctica	61
Características intrínsecas en la práctica	62
Características extrínsecas en la práctica	65

CAPÍTULO III

Los indicadores	71
Definición	72
Funciones de los indicadores	74
Estadísticas e indicadores	75
Tipologías de indicadores	76
Indicadores cualitativos y cuantitativos	77
Indicadores temáticos	79
Indicadores de gestión	81
Indicadores según jerarquía	90
Características de los indicadores	90
Naturaleza matemática de los indicadores	91
Medidas estadísticas	92
Frecuencias	92
Razones	95
Tasas	97
Índices	99

CAPÍTULO IV

Modelo P2C-T	101
Definición	102
Planeación estadística (P)	104

Conceptualización y metodología (C)	107
Producción estadística (P)	110
Tecnología (T)	112
Infraestructura tecnológica general para la gestión de datos	113
Softwares especializados en gestión de datos	114
Herramientas tecnológicas utilizadas por la UNAL para la gestión de los datos	115
Aporte de la UNAL a la producción de estadísticas institucionales. El paquete de R UnalR	117
Cultura estadística (C)	118
Soportes del modelo P2C-T	120
Apoyo institucional	120
Talento humano	121
Recursos económicos	123
CONCLUSIONES	125
REFERENCIAS	131

Cuando se trata de las instituciones, lo humano subyace en el fondo del concepto de organización. La organización está compuesta por la estrategia, los sistemas, los procesos, la estructura, las personas, los valores, las habilidades o capacidades y los estilos de trabajo¹. Todos y cada uno de estos elementos concurren para el éxito de las instituciones en el logro de sus fines.

En el diseño estratégico se plasma el ADN de la organización, la misión, la visión, los grandes propósitos y las maneras como se ha de concretar el quehacer de la institución. La eficiencia, la eficacia, la efectividad se juzgan fundamentalmente a la luz de la estrategia.

Con frecuencia se piensa que solo comprendiendo el organigrama se logran identificar las coordenadas del lugar que cada individuo ocupa en la institución, y que solo basta con leer el manual de funciones para saber qué hacer todos los días ante contextos a veces inerciales y en ocasiones inesperados. En realidad, en los procesos de inducción y capacitación de los trabajadores se pasa por alto el hecho de que la estructura está determinada por la estrategia, los sistemas² que componen la organización y los procesos que se requieren para materializar el objeto social de la institución; de hecho, aunque no parezca, en la gestión pública en países como el nuestro muchos de los procesos, las actividades y las funciones se incorporan basados en los niveles de desconfianza prevalecientes en la organización.

Las coordenadas de ubicación del individuo o de un colegiado son multidimensionales, se puede pertenecer a varios sistemas a la vez actuando desde las intersecciones o en el resultado de una unión de dichos sistemas, estar adscrito a diversos procesos sincrónicos o asincrónicos y contribuir al logro de solo una parte de los objetivos misionales que

¹ Es el modelo de las 7S de McKinsey. https://www.academia.edu/42963084/_MODELO_DE_LAS_7S_de_MCKINSEYLECTURA_COMPLEMENTARIA_No_2

² Se usa la palabra sistemas en un sentido amplio, no se refiere a la tecnología.

constituyen el todo. En realidad, la red de creación de valor de una institución poco tiene de lineal, es multidimensional. Entender esa múltiple afiliación de las personas es clave para hacer consciente el rol esencial que todo individuo tiene en relación con las responsabilidades que le corresponden como parte de la red de valor de la información; red que abarca desde el surgimiento de un dato clave hasta el uso que la organización haga de la información, las estadísticas e indicadores para comprender a la entidad en su conjunto y tomar las mejores decisiones.

Como se verá a lo largo de este libro, los datos, las estadísticas, los indicadores, la información tienen que emerger de los procesos cotidianos de la organización como algo natural, no pueden ser considerados una carga cuando se comprende la importancia de la calidad y oportunidad de la información.

Los individuos y colegiados son, a la vez, proveedores y usuarios, atienden varios procesos, se articulan en diversos sistemas y, para poder sobrevivir con éxito en medio de tanta complejidad, se requieren liderazgos efectivos que garanticen sincronía y armonía. Para lograrlo es necesario, además de liderazgo distribuido, es decir, colectivo, conciencia de la ubicación, nitidez en los roles y las funciones, y normalización.

La normalización parte de unificar el lenguaje y los significados, precisar los metadatos, decidir el día y la hora de las emisiones y recepciones, unas reglas claras de lo que significa el carácter oficial de la información y, pensando en la estructura, ahora sí en el organigrama, comprender y valorar lo que se entenderá por los canales oficiales por los que ha de fluir la información y el adecuado uso que debe hacerse de las jerarquías que están detrás del organigrama. En la normalización se debe dejar establecido qué hacer cuando las contingencias e imprevistos emergen, por ejemplo, ante una catástrofe o una pandemia, momentos en los cuales todas las variables y tendencias “se desordenan”.

Aquí es donde aparece la importancia de las otras dimensiones de la organización: las personas, los valores, las habilidades y los estilos de trabajo. Las jerarquías se ejercen con la autoridad que surge de los conocimientos, las experiencias y los estilos de liderazgo de los jefes para que logren sincronizar y armonizar al todo y a sus partes. La integridad es constitutiva de ese liderazgo.

Las personas que ejercen sus roles y sus capacidades contribuyen al grupo también con sus conocimientos, experiencias, estilos de aprendizaje,

disposición al trabajo en equipo, habilidades de liderazgo y comunicación, y disposición a acatar los ritos y las normas de la organización. Los estilos de trabajo, las prácticas, se van estructurando y, a veces, fosilizando con las vivencias de todos los días, entran al subconsciente de las personas y de la organización y difícilmente son percibidas mediante procesos de metacognición a los cuales estamos poco acostumbrados. Se requiere que con frecuencia alguien externo, los denominados “pares”, las evalúe y las critique desde el balcón.

Por tratarse del mundo de los humanos es necesario tener en cuenta que en las organizaciones operan los organigramas, pero que con frecuencia —y a veces por fortuna— conviven otras formas de organización: las informales, las relaciones verdaderas, la de las amistades, las redes, los encuentros en los corredores o las citas virtuales y por allí, desde la clandestinidad, fluye también una valiosa información. El rumor lleva estadísticas, indicadores, información, allí se despliegan a plenitud los modelos mentales, los imaginarios, el juicio despiadado y cruel que se hace sobre los líderes o sobre las partes centrales de los organigramas, que concentran el poder y los recursos. Es la “picota pública” de los pasillos. Sería maravilloso poder documentar y volver información valiosa, ojalá con indicadores cualitativos, lo que sucede en ese mundo real pero incomprensible de los rumores.

Quería con estas pocas notas hacer evidente la complejidad de este mundo de las estadísticas y los indicadores para que realmente funcionen con efectividad. Los obstáculos saltan a la vista y adoptan diversos matices y lenguajes: “que esta organización es tan compleja que no puede simplificarse como lo pretenden los estadísticos”; “que esta institución es única en el planeta tierra y que por tanto no tiene referentes para hacer comparaciones y determinar umbrales y rangos”; “que es un esfuerzo vano establecer retos o metas basándose en los *rankings* nacionales o internacionales por que hacemos cosas que ninguna otra entidad hace”; “somos muy pocos y tan mal pagados que no es razonable que nos evalúen, juzguen y nos exijan más”; “es absurdo que los estadísticos importen modelos desde lo privado pretendiendo que nuestra entidad se comporte como empresa”.

Somos conscientes de la dificultad de mirarse al espejo con detalle para lograr percibir nuestras fortalezas y debilidades. Las estadísticas y los indicadores, con las verdades crueles que ellos transportan, con los

mensajes profundos que contienen para describir lo que somos, lo que fuimos y lo que seremos (si se mantienen las cosas como van), nos van dando pistas acerca de cómo proceder para salir de la zona de confort y entrar al mundo de la innovación y el cambio.

Ahora sí, bienvenidos a los contenidos profundos de los capítulos de este libro, es un gustoso e interesante tratado de la estadística aplicable a lo público; pero, nunca pierden de vista la dimensión humana que está detrás de cada párrafo, de cada algoritmo, de cada variable, y el esfuerzo adicional que se requiere para que toda la organización opere con sincronía y armonía sin perder el horizonte estratégico de los fines misionales.

Carlos Alberto Garzón Gaitán

Profesor Emérito, Universidad Nacional de Colombia

Octubre 11 de 2023

Las preguntas sobre los modos en que se gestiona la información dentro de las entidades no son un tema menor. Para quienes están involucrados en este ámbito, pueden resultar interesantes —o, sobre todo, frecuentes— cuestionamientos acerca de cómo recopilar, analizar y emplear datos para respaldar el funcionamiento interno de las instituciones. Sin embargo, al abordar estos interrogantes, es frecuente encontrarse con el desafío de definir y comprender plenamente ciertos componentes que conforman el ámbito de la gestión. ¿Hasta qué punto es precisa la comprensión al aludir a conceptos como información, estadísticas o indicadores? ¿Hasta qué punto se comparte una visión clara de cómo estas piezas se integran de manera efectiva en la práctica?

En este escenario, este libro se presenta con el propósito de proporcionar apoyo conceptual a las entidades públicas en el estudio y la construcción de estadísticas e indicadores institucionales, claves para los procesos de gestión. Para lograr este objetivo, el libro se estructura en cuatro capítulos descritos a continuación.

En el primer capítulo se subraya la importancia de la información en entornos altamente tecnológicos y digitalizados, mientras se da cuenta del término “información” como un elemento polisémico que puede ser abordado desde una perspectiva etimológica y pragmática, donde se resalta la noción de información como cosa, que es susceptible de ser manipulada, analizada y procesada. Asimismo, se resalta la relación entre la gestión pública —que implica la formulación y ejecución de estrategias destinadas a asuntos de interés público— y la información, y se destacan tres usos fundamentales de la información en tal contexto: comunicativo, de registro y valorativo.

El segundo capítulo se centra en el análisis de las estadísticas, las cuales se definen como un conjunto de datos organizados en cifras descriptivas que desempeñan un papel fundamental en la representación y la comprensión de la realidad social, especialmente en el contexto de la gestión pública. La finalidad de este capítulo es explorar las características

de las estadísticas como un objeto de estudio, a partir de dos categorías: características intrínsecas y características extrínsecas. Las primeras se enfocan en identificar las condiciones que deben cumplirse para hablar propiamente de estadísticas. Las segundas, por su parte, se relacionan con las recomendaciones pertinentes para el uso de las estadísticas en el ámbito de la gestión pública. Ahora bien, es importante destacar que este capítulo no pretende formular directrices específicas ni establecer un modelo concreto para la recopilación y consolidación de cifras estadísticas, sino describirlas como objeto de análisis.

El tercer capítulo aborda el concepto de “indicadores” como herramientas fundamentales en la medición y evaluación de fenómenos de interés en el ámbito público. En este sentido, se desglosan las funciones primordiales que los indicadores desempeñan en dicho contexto. El capítulo avanza explorando la relación entre las cifras estadísticas y los indicadores, resaltando sus aplicaciones prácticas. Además, se introduce una tipología de indicadores que abarca diversas categorías, con énfasis en la importancia de comprenderlas y utilizarlas de manera conjunta para lograr una gestión efectiva en el ámbito público. Finalmente, se pone de manifiesto la base matemática que subyace en la formulación de indicadores, junto con la exposición de diversas estrategias, como medidas estadísticas, frecuencias, razones, tasas e índices, que se emplean en su desarrollo.

En el cuarto capítulo se presentan algunos elementos y recomendaciones prácticas que buscan dar cuenta del proceso de creación de estadísticas e indicadores, así como la implementación de áreas de estadística o análisis de datos dentro de la Universidad Nacional de Colombia. Con este propósito en mente, se introduce un modelo propio denominado Modelo P2C-T, que se compone de cinco dimensiones clave: planeación estadística (P), conceptualización y metodologías (C), producción estadística (P), gestión de herramientas tecnológicas (T) y promoción de una cultura estadística (C), y tres pilares fundamentales que se deben tener en cuenta para que el modelo funcione de manera efectiva: 1) el respaldo institucional, 2) la disponibilidad de talento humano capacitado y 3) el acceso a recursos financieros. El propósito de este capítulo no es proponer una metodología específica que deba ser adoptada por todas las instituciones, sino, más bien, dar a conocer una experiencia particular que pueda servir de apoyo.

CAPÍTULO I

Información y gestión pública

El presente capítulo explora el concepto de información y su relación con la gestión pública. Para ello, en primer lugar, se abordarán los problemas conceptuales de la información, discutiendo la importancia del concepto en el contexto histórico actual, sus problemas teóricos, definiciones y características. En segundo lugar, se analiza la gestión pública como una unidad temática aislada con el fin de brindar al lector un panorama general de sus implicaciones. Finalmente, el objetivo del capítulo es ayudar al lector a entender la relación entre información y gestión pública; comprender el papel especial que tiene la información cuantitativa en dicha relación y, por último, alertar acerca de los retos y las exigencias éticas involucradas en el manejo y uso de la información dentro del marco de la gestión pública.

INFORMACIÓN

La palabra “información” resulta ser un término bastante habitual en las interacciones sociales. Es una de esas palabras que se emplea sin mayor complicación en diferentes contextos. Cuando se dice que “hay que estar *informados*” o que “debemos cuidar nuestra *información personal*”, no se sabe lo suficiente al respecto de lo que se está entendiendo por información; simplemente, se tiene la sensación de comprender lo que se dice.

No es casualidad que el término en cuestión resulte tan habitual, parece ser que el momento actual es el momento de la información. Tanto los avances en materia de telecomunicaciones y computación como la aparición del universo digital han condicionado en gran medida los modos en que se interactúa con el mundo y con los demás. Lo anterior al punto en el que la mayoría de las experiencias humanas se ven permeadas, de

una u otra forma, por el consumo y la transmisión de información. De hecho, autores como Yoneji Masuda (1984), Manuel Castells (1995) o Nick Moore (1997) se refieren a una “sociedad de la información”, esto es, una sociedad donde los modos en que se produce, se consume y se emplea la información resultan de importancia notable para el devenir social³.

La importancia concedida al tema de la información en la época actual se puede notar en eventos como la Cumbre Mundial sobre la Sociedad de la Información, llevada a cabo en 2003 con delegados de 175 países, donde se trataron los temas de acceso a la información y de la “gobernanza de internet”.

La Cumbre Mundial sobre la Sociedad de la Información tuvo una segunda fase en 2005. Allí, además del tema del acceso a la información, se trató la cuestión del “dilema digital”. Tal cuestión tiene que ver con la opinión bastante extendida de que fenómenos como la sociedad de la información, la era digital, el internet y otros similares pueden fomentar algunas problemáticas que vale la pena considerar.

Por ejemplo, uno de los elementos centrales del “dilema” está asociado a la *cantidad* de contenidos que se producen y que se consumen. Todos los días se transmiten y recepcionan más datos de los que cualquier ser humano por sí mismo podría procesar en toda su vida. De hecho, según un reporte del gigante del contenido audiovisual YouTube, en dicha plataforma se veían para 2017 más de mil millones de horas de video por día, es decir, más de cien mil años de reproducción continua (Goodrow, s. f.). Con tanta información resulta bastante difícil organizar, dirigir y controlar el contenido que se presenta. A causa de esta “sobrecarga” de información y de las dificultades para gestionarla, se presentan problemáticas como las *fake news*, la posverdad, la pérdida de credibilidad; en una palabra, la *desinformación*.

La cantidad no es el único problema asociado a la era digital, en general, y a la información, en particular. También es importante considerar la cuestión de la *distribución*. Pongamos en cuestión el caso de las redes sociales. Los algoritmos de estas redes muestran contenidos

³ Por supuesto, esta forma de clasificar o describir las sociedades no es del todo universal, pues existen comunidades excluidas o ajenas al “Estado moderno” y a las prácticas sociales globales. Sin embargo, las sociedades de la información parecen ser el camino al que tienden la mayoría de las comunidades.

“cercanos” a las creencias de los consumidores. Así, el contenido que consume una persona con unos intereses particulares puede ser diametralmente opuesto al de otra persona con otros intereses. En principio esto no parece ser problemático. Sin embargo, en la actualidad, las redes sociales constituyen parte importante de la interpretación que se hace del mundo y de los modos en que las personas se relacionan con él. De manera que el sentido de realidad ya no es compartido, cada persona accede a una “verdad” distinta programada en su celular, o lo que Pariser (2011) denomina “burbuja filtrada”. Esto lleva, entre otras, a la polarización y la falta de empatía.

Otro elemento que vale la pena considerar tiene que ver con *los modos y los propósitos* para obtener información. Un buen ejemplo podría ser lo sucedido en 2018 con Facebook y Cambridge Analytics. De acuerdo con Greenfield (2018), una asociación privada, Cambridge Analytics, pudo acceder a información personal de miles de usuarios en Facebook y, con base en dicha información, manipular las elecciones presidenciales de los Estados Unidos. Esta situación pone de manifiesto cuestiones como a quién le pertenece la información, qué se puede hacer y qué no hacer con ella, cuál es el límite entre lo privado y lo público de la información, entre otros asuntos cuestionables.

Como se ha visto, la información emerge como un tema de relevancia contemporánea que convoca diversas problemáticas a su alrededor. Ahora bien, desde la perspectiva del presente libro, la información se presenta como un elemento necesario para la construcción de cifras estadísticas sobre las cuales es posible la gestión de los Estados y sus instituciones. De modo que en lo que sigue del presente capítulo se discute la noción de información, explorando su comprensión y evolución histórica, mientras se tejen conexiones entre información y gestión pública.

Definición de información

El concepto de información es problemático. A pesar de su cotidianidad y de su presencia en nuestras actividades más básicas, no es tan sencillo precisar sus particularidades y distinguirlo, por ejemplo, de otros conceptos como conocimiento, saber, dato, noticia o, incluso, aprendizaje. De hecho, Buckland (1991) resalta que es bastante irónico pensar en la información como algo asociado a la reducción de la ignorancia

y de la incertidumbre y, sin embargo, se use de manera tan amplia y en contextos tan diversos.

Etimología de la información

Como es costumbre, cuando se quiere precisar el significado de alguna palabra, se suele recurrir a su etimología. “Información” viene del latín *informatio*, que es la forma nominal del verbo *informare*, que quiere decir dar forma. Resulta importante resaltar que tal término implica una cierta acción o evento, no es estático; no es solo una forma, sino *dar* forma. Precisamente, el sufijo *-tio*, de *informa-tio*, denota acción o efecto. Por su parte, “dar forma” tiene que ver con definir, limitar, organizar. La forma tiene siempre una estructura. De hecho, el radical *forma* se encuentra también en *formalis*, formal. Con respecto al prefijo *in-* se relaciona con la raíz indoeuropea *en*, que designa un cierto tránsito hacia dentro, justamente este prefijo se encuentra en *internus*, interno. Así, se puede decir que el término mencionado refiere a una cierta estructura —o forma— interna (*in-forma*) que implica la acción, el movimiento, el suceso (*-tio*).

Enfoque pragmático

Con todo, esta aproximación al concepto es, aún, bastante general. Otra buena forma de continuar abordando el sentido que se le asigna al término “información” es atender a sus contextos de uso. Como menciona Buckland (1991), “frente a la variedad de significados [...] al menos podemos adoptar un enfoque pragmático” (p. 351).

De este modo, Buckland (1991) distingue tres grandes usos del término en las conversaciones habituales, a saber: información como proceso, información como conocimiento e información como cosa. En el primero, *información como proceso*, se resalta el carácter activo del concepto: la información como el acto de informar. Esta concepción involucra un *proceso* de carácter comunicativo donde el elemento central no es lo comunicado sino, justamente, el acto de comunicar, de informar. En el segundo, sobre la *información como conocimiento*, se destaca un evento posterior al acto de informar. Es decir, el foco ya no está en el proceso comunicativo, sino en su consecuencia: cuando alguien es informado sobre algún aspecto, hay algo que percibe y que, de alguna manera, cambia lo que sabe al respecto de un tema en concreto. Para Buckland estos dos usos de la palabra (como proceso y como conocimiento) se caracterizan por su inteligibilidad: “si se

puede tocar o medir de manera directa, no es conocimiento [ni proceso], sino algo físico” (p. 352). Así, el tercer uso del término, el de *información como cosa*, hace referencia a la información como una entidad física que puede transmitir o comunicar algo.

Hasta ahora se ha mostrado que Buckland (1991) expone tres usos posibles que se hacen del término *información*: como proceso, como conocimiento y como cosa. Resulta importante destacar que es posible establecer una distinción entre los términos que subyacen a los usos mencionados. Por ejemplo, el informático danés Ingwersen⁴ (1992) establece una precisión clara que distingue datos, información y conocimiento. Para este autor, los datos se definen como información en potencia, es decir, información como cosa, que solo es realmente información cuando es percibida por un receptor. Así, para Ingwersen, la diferencia entre el dato (información potencial, que no ha sido percibida) y la información propiamente dicha tiene que ver con la transformación por parte del receptor de los datos a través de sus estructuras de conocimiento. Con respecto al conocimiento, Ingwersen menciona que ocurre cuando la información, es decir, los datos percibidos, modifican la estructura de conocimiento del receptor.

Ahora bien, las definiciones de Buckland (1991) e Ingwersen (1992) no son opuestas ni contradictorias, sino que resultan complementarias. Buckland denomina *información como cosa* a lo que Ingwersen llama *dato*; *información como proceso* a lo que en Ingwersen es *información* en sentido estricto; y, finalmente, *información como conocimiento* a lo que, en términos de Ingwersen, es simplemente, *conocimiento*.

Con todo, abordar el concepto de información desde una perspectiva pragmática involucra, por lo menos, tres elementos: un objeto que se puede administrar (p. ej. almacenar, medir, distribuir, copiar) y que vendría siendo *información como cosa* para Buckland (1991) o *dato* para Ingwersen (1992); un proceso interactivo donde se comunica algo, *información como proceso* para Buckland o *información*, en sentido estricto, para Ingwersen y, por último, un aspecto conceptual, donde se sabe, se cree, se conoce

⁴ Hace parte de un enfoque subjetivo y considera que la información (como dato percibido) solamente se puede dar entre humanos.

o se opina algo, es decir, *información como conocimiento*⁵ para Buckland o *conocimiento para Ingwersen*⁶.

Buckland (1991) argumenta que el ámbito académico se ha concentrado, sobre todo, en la información como proceso y como conocimiento, no así en la información como cosa. Sin embargo, para el autor, la información como cosa posee una importancia notable en la medida en que los sistemas de información, encargados de procesar, preservar, extraer o predecir información, solo pueden operar con la información como cosa.

En este punto, cabe destacar que el presente libro hace parte de una indagación estadística y la estadística involucra, justamente, una serie de técnicas de descripción, recopilación, registro y cálculo que se aplican al dato, a las cifras, a los documentos, en una palabra, a la *información como cosa*. De modo que tal noción de la información, en tanto información que es susceptible de ser procesada, resulta crucial para la estadística y, por lo tanto, para este libro. Desde este punto de vista, en adelante el texto se concentra en la concepción de *información como cosa*.

La *información como cosa* hace referencia, justamente, a las cosas que pueden ser informativas, como, por ejemplo, un libro, una película o una base de datos. Pero ¿qué cosas pueden ser informativas? o, mejor, ¿qué cosas no lo son? Pues bien, para Buckland (1991) la información como cosa se caracteriza por ser prueba o evidencia de algo. Así, una cosa puede ser informativa en la medida en que sirva como evidencia de algo en alguna situación. De manera que, respondiendo a la pregunta de qué cosas no pueden ser informativas, el autor afirma que prácticamente cualquier cosa puede ser considerada información: “llegamos a la conclusión de que somos incapaces de decir con confianza que algo no podría ser información” (p. 356).

⁵ Como se sabe, desde la tradición filosófica de la antigua Grecia se establece una distinción categórica entre conocimiento (*episteme*), por un lado, y opinión o creencia (*doxa*), por el otro. En esta ocasión, sin embargo, con *información como conocimiento* se hace referencia a un aspecto conceptual en sentido amplio.

⁶ Sin embargo, acá no se agotan los problemas. Es posible preguntarse si la información como proceso, en tanto situacional e intencional, puede darse también en animales u objetos. De manera similar, el conocimiento, en tanto representación conceptual, remite a la pregunta por lo cognitivo y pone en cuestión asuntos actuales como el *machine learning*. No obstante, estos elementos no hacen parte propiamente de un enfoque pragmático de la información.

Aun así, Buckland (1991) establece cuatro tipos de información como cosa: los datos, los documentos, los objetos y los eventos. A continuación, se aborda cada uno de ellos.

Dato (o data)

La palabra “dato” viene del latín *datum* que deriva, a su vez, del verbo *dare* y se refiere a “lo que nos es dado”. Buckland (1991) toma esta concepción etimológica para argumentar que los datos tienen que ver con la “información como cosa que se ha procesado de alguna manera para su uso” (p. 353). Dicho de otro modo, en la medida en que los datos tienen que ver con algo que “es dado” implican un determinado proceso y ahí radica la asociación entre algo que es *dado* y algo *procesado* para ser usado. De otro modo, el autor resalta que, con frecuencia, los datos se asocian a registros que se almacenan en una computadora o, simplemente, a información de carácter cuantitativo.

Es importante mencionar que en el ámbito de las ciencias económicas se suele hacer referencia a los datos en términos de “bienes no rivales”; lo anterior en la medida en que un determinado uso de ellos no excluye otro. Así, por ejemplo, los datos de los usuarios de una red social pueden ser usados para contactar a personas con intereses comunes, para enviar publicidad personalizada o, quizás, para clasificar demográficamente a los usuarios y, sin embargo, tales datos no se agotan.

Documento

Según Buckland (1991), el término “documento” se emplea normalmente para referirse a objetos que contienen texto, como cartas, libros, periódicos o escritos. Sin embargo, el autor advierte que el término “texto” bien podría relacionarse con imágenes o sonidos. Todavía más, si se piensa en una tabla de números no resulta para nada forzado pensar en ella como un documento, texto o, incluso, como dato. Así, se resalta el carácter polisémico de las palabras en cuestión, donde una podría estar en lugar de las otras. De modo que, desde el uso de la lengua, es recomendable evitar establecer una demarcación definitiva entre datos, documentos y texto. Aun así, es importante señalar que una buena manera de definir el concepto de documento, y que aplique tanto para fotos o sonidos como para cartas o periódicos, podría ser la siguiente: un conjunto ordenado de datos alineados de acuerdo con un propósito explícito. Cabe agregar

que los datos que componen un documento pueden ser, como se mencionó, empleados para un propósito completamente distinto del original.

Objeto

Además de datos y documentos, hay otras *cosas* que son informativas, aunque no se suelan ver de esa forma: los objetos. Pongamos por caso un cono de tránsito, que, aunque no se considere *stricto sensu* dato o documento, es información como cosa en la medida en que se presenta como evidencia de un accidente o un vehículo detenido, por ejemplo. Buckland (1991) ilustra la noción de objeto informativo al hablar de fósiles y artículos de museos. Ahora bien, la representación que se hace de tales objetos pone de manifiesto su carácter informativo. Así, los artículos de un museo son contados y fotografiados y tal representación constituye, *per se*, nueva información.

Eventos

Por su parte, los eventos también se consolidan, desde la perspectiva de Buckland (1991), como un tipo de información como cosa. Sin ir más lejos, un experimento puede ser entendido como un evento que sirve como evidencia de la veracidad de una hipótesis determinada. Sin embargo, este tipo de información como cosa está en estrecha relación con la representación. De hecho, Buckland subraya que, en ocasiones, un objeto puede ser considerado información en la medida en que aparece como evidencia de un evento. Esto es claro si se piensa en las huellas de un lobo en la nieve como un objeto informativo, que lo es en la medida en que evidencia el evento que describe el paso de un lobo. Todavía más, muchos de los objetos que se conservan en los museos están allí justamente porque representan algún *evento* de alguna época, como el florero de Llorente, que se conserva en el Museo de la Independencia ubicado en la carrera séptima de Bogotá, Colombia. Aquí es importante señalar que, en ocasiones, la única manera de interactuar con un evento es, justamente, a través de su representación en fotos, videos, narraciones, entre otros.

Campos de la información

Además de ser polisémico, el término “información” es abordado desde diversas disciplinas, como son la informática, la filosofía, la

cognición, la comunicación y las ciencias del lenguaje, la biología, la física, entre otras. Cada disciplina asume diferentes concepciones del término y le da distintos usos. De hecho, García (1998) resalta que el concepto posee una extraordinaria generalidad y que lo más conveniente es especificar los sentidos con los que se emplea en ámbitos concretos.

Los distintos modos en los que las diferentes disciplinas asumen el concepto de información se pueden resumir en lo que Rendón (2017) denomina “campos de la realidad en donde se manifiesta un tipo de información específico” (p. 65), estos son: el *físico*, el *biológico*, el *psicológico* y el *tecnológico*. Vale la pena mencionar que los campos son conceptualizaciones del mundo y por ello son susceptibles de reinterpretaciones o críticas. Sin embargo, como conceptos, son útiles para exponer la historia y los enfoques en torno a la información.

El *campo físico* se enfoca en la dimensión material del mundo y las disciplinas que lo investigan. Aquí, la información existe como un elemento “estrictamente sintáctico” y “en potencia”, sin un significado intrínseco, debido a la ausencia de un intérprete. En el *campo biológico*, la información se presenta en términos de biomoléculas y procesos comunicativos cruciales para la existencia y evolución de los seres vivos. Las teorías genéticas informacionales son ejemplos de enfoques aplicados en este ámbito. En el *campo psicológico*, por su parte, la información adquiere un papel cognitivo y social, donde la inteligencia opera, razona y crea a través de la información. Aquí, tanto las teorías lingüísticas como la antropología abordan la realidad social como un proceso comunicativo en el que la información es un elemento central. Finalmente, el *campo tecnológico* comprende la manipulación de información a través de artefactos u herramientas, destacando la creciente influencia de la tecnología de la información y la comunicación (TIC) en todas las disciplinas y en la sociedad en general.

GESTIÓN PÚBLICA

Gestión

La palabra “gestionar” se define en el diccionario de la Real Academia Española (2001) a partir de tres acepciones: primero, “llevar adelante

una iniciativa o un proyecto”; segundo, “ocuparse de la administración, organización y funcionamiento de una empresa, actividad económica u organismo”; y, tercero, “manejar o conducir una situación problemática”. Estas acepciones mantienen una estrecha relación y podrían, incluso, presentarse de forma conjunta. Así, por gestión se podría entender una cierta iniciativa o proyecto, que tiene que ver con la administración, organización y funcionamiento de un elemento o situación que presenta un carácter problemático. Complementando, Muñoz (2020) expone que tal término “resalta las acciones precisas, oportunas y prácticas en pro de conseguir en el menor tiempo posible, y con la calidad del caso, los propósitos trazados” (p. 96), asimismo, menciona la autora, la gestión hace parte de las diferentes esferas de vida cotidiana, por lo que cabe hablar de gestión institucional, comunitaria, local, privada, pública, entre otras.

Cada tipo de gestión conlleva sus propios objetivos, que condicionan las acciones que se llevan a cabo para alcanzarlos. En otras palabras, las perspectivas, metas y acciones que fundamentan la gestión en una empresa privada difieren de las de una junta comunitaria. Por un lado, la gestión en una empresa privada tiende a centrarse en la obtención de beneficios en términos generales; mientras que, por otro lado, la gestión en una junta comunitaria se enfoca en identificar los problemas bajo su jurisdicción y en buscar formas de resolverlos. El tipo de gestión que se aborda en lo que sigue es la *gestión pública*.

Público

La palabra “público” viene del latín *populicus* y está compuesta por *populus* (gente, pueblo) y por el sufijo *-ico* (relativo a), es decir, “lo que pertenece al pueblo”. La noción de lo público involucra un elemento social, de comunidad, de relación, de convivencia y, por tanto, de norma. De hecho, “público” aparece también en “república” (*respublica*), la cosa o los asuntos (*res-*) del pueblo (*publica*). Pues bien, la cuestión de lo público desempeña un papel fundamental en el ámbito de la gestión. Cuando la pregunta por la administración, la organización y el funcionamiento se enfoca en lo que es del —o para el— *pueblo*, la gente, todos⁷, se espera

⁷ Este asunto es cuestionable. Uno de los sentidos que se atribuye al pueblo tiene que ver con la masa, el común de la gente, lo *popular*. Esto quiere decir que hay

una manera especial de hacer las cosas, o, al menos, se espera una cierta pregunta por los modos de ser, de actuar.

El tema de los modos de ser, o de hacer, se relaciona con la palabra griega *ethos*, ética. La pregunta ética por excelencia, como se plantea en la filosofía griega, adopta la forma de “¿cómo deberíamos vivir?” o “¿qué deberíamos hacer?”. La ética, desde esta perspectiva, involucra la reflexión sobre las “mejores opciones” para el desarrollo de la vida humana, entendida esta como una vida social, en comunidad. Así pues, el ámbito de lo público involucra siempre un aspecto ético.

Gestión pública

Los significados de las construcciones discursivas que empleamos no se limitan a los significados de las partes que las componen. Así, la noción de “gestión pública” se puede entender, también, como una disciplina que involucra algunos elementos adicionales.

Bozeman (1998) presenta una distinción entre dos orígenes de la gestión pública como disciplina. Por un lado, el enfoque B (de *business*), que está más asociado a la administración tradicional y a las disciplinas comerciales. En este enfoque la distinción entre público/privado no resulta tan destacable. Por el otro lado, hay un enfoque P (de *politics*) que considera de vital importancia el análisis de la política pública en el proceso de gestión.

Luego de esta distinción, Muñoz (2020), siguiendo a Bozeman (1998), resalta que la gestión y las políticas públicas son dos caras de la misma moneda. De un lado, la política pública, es decir, una suerte de plan que se ha convenido para atender a una problemática pública. Del otro lado, la puesta en práctica, es decir, la gestión de aquellos planes que se han convenido. Así, la política pública, en cierto sentido, traza unos objetivos que dan cuenta de lo que se debe hacer, mientras que la gestión pública se relaciona más con el método que responde al cómo se va a hacer eso que

una sección que no se considera parte de ese pueblo y que, de algún modo, está en una posición privilegiada. Tal concepción no es casualidad. De hecho, la palabra griega *populus* se refería a todos los hombres jóvenes (*púber* viene de *populus*) y dejaba por fuera a mujeres, niñas, niños y ancianos. Esto pone de manifiesto que la concepción de público y de todos obedece, usualmente, a representaciones y construcciones sociales que varían de acuerdo con la época.

se piensa realizar. Esta manera de entender el asunto pone de manifiesto la disyuntiva entre teoría y práctica o, si se prefiere, entre el decir y el hacer. Como se sabe, las iniciativas estatales pueden —o, mejor, suelen— quedar en el papel. De ahí que la gestión adquiera una importancia considerable.

Es importante mencionar que la concepción de *política pública* suele limitarse al plan, programa o proyecto que genera una administración específica y que tiene un alcance limitado en el tiempo, que usualmente se relaciona con el periodo de gobierno de un determinado personaje o partido político. Sin embargo, la política pública, así como la gestión pública, va más allá de lo coyuntural. Hay acuerdos estatales y sociales para atender a problemas públicos que no están acotados, al menos de manera tan radical, en el tiempo. En este sentido, tanto la política como la gestión pública pueden involucrar procesos transversales que se extienden más allá de los intereses del gobierno de turno.

Modelos de gestión pública

Dentro de la mencionada gestión, o puesta en marcha de la política pública, se pueden distinguir modelos. Estos modos de llevar a cabo la política pública no son ajenos a la noción de información. En cada modelo, la importancia concedida a la información va tomando un matiz diferente.

De acuerdo con Muñoz (2020), se pueden resaltar seis grandes modelos. En primer lugar, el *modelo burocrático*. Este tiene que ver con una organización mecanicista, racionalista y vertical que busca estandarizar los procedimientos y pone al Estado como actor fundamental de la gestión, que puede dar respuesta a todas las problemáticas sociales.

En segundo lugar, el *modelo de la nueva gestión pública*, que, siguiendo a Muñoz (2020), presenta una noción contractualista, que instala principios gerenciales y empresariales en el sector público donde el ciudadano es visto como cliente. Esa concepción redunda en el desplazamiento del servicio estatal y de la gestión hacia el entorno privado. Además, la noción de ciudadano-cliente se traduce en herramientas de evaluación y medición de los servicios públicos, así como en la gestión por resultados.

En tercer lugar, el *modelo de gobernanza*, que, de acuerdo con Muñoz (2020), reconoce que la solución a los problemas públicos no se limita al sector público-estatal o al sector privado, sino a un relacionamiento

horizontal entre estos dos últimos y los diferentes actores sociales. Así, este modelo aborda la gestión pública desde una óptica de trabajo colaborativo, donde el Estado debe coordinar y potenciar una red de distintos actores. No obstante, se ha argumentado que esta dinámica puede propiciar la preponderancia de intereses de grupos que cooptan las redes para su beneficio personal.

En cuarto lugar, el *modelo de gobierno abierto*. Para Muñoz (2020), este se centra en tres principios fundamentales, a saber, transparencia, participación y colaboración. *Transparencia* en cuanto al acceso a la información pública para supervisar el actuar del gobierno; *participación* para propiciar espacios de diálogo y toma de decisiones por parte de la ciudadanía y el sector público; y, finalmente, *colaboración* entre los sectores privado, público y civil para abordar los problemas como una red. Se le denomina gobierno abierto por el énfasis en la *apertura* de datos y procesos para promover la transparencia y facilitar la participación y la colaboración.

En quinto lugar, el *modelo de nuevo servicio público*, que, como señala Muñoz (2020), retoma las nociones de ciudadanía, un poco descuidadas en la *nueva gestión pública*. Así, tiene como objetivo central el interés público. Se deja de lado el enfoque empresarial para resaltar el *servicio público*. Así, los funcionarios pasan a ser servidores públicos. Se valoran las características particulares y las circunstancias de las personas, no solo su productividad.

Finalmente, en sexto lugar, el *modelo de gestión pública territorial*. Dicho modelo se caracteriza, de acuerdo con Muñoz (2020), por un enfoque integral que abarca elementos como las relaciones entre Estado, instituciones, sector privado y comunidad. Asimismo, el elemento distintivo tiene que ver con el reconocimiento de la noción de territorio y los modos en que influye en las redes que se establecen entre los diferentes actores sociales. De manera que el propósito central se refiere a la acción común en un territorio determinado.

Ahora bien, los enfoques de gestión pública mencionados se presentan como divisiones teóricas; no obstante, en la práctica, es frecuente encontrarse con modelos híbridos o mezclados. Esta combinación de enfoques hace que la gestión de la información se vuelva más compleja.

USOS DE LA INFORMACIÓN EN LA GESTIÓN PÚBLICA

Independientemente del modelo o enfoque que la gestión pública adopte, la información desempeña un papel preponderante y transversal. A continuación, se ofrece una perspectiva basada en la experiencia de gestión en la Universidad Nacional de Colombia (UNAL), que describe cuatro principales formas en las que las instituciones hacen uso de la información.

Uso comunicativo

La información se emplea de manera comunicativa cuando tiene el propósito explícito de transmitir un mensaje a algún receptor que, en muchas ocasiones, puede no ser un individuo en concreto, sino un receptor ideal. Algunos ejemplos de este uso son comunicados oficiales, citaciones, circulares, noticias, boletines informativos, manuales, descripción de funciones, publicaciones en redes sociales, memorandos, correos electrónicos, oficios, entre otros.

Uno de los elementos más sobresalientes del uso comunicativo tiene que ver con dar a conocer lo que las entidades están haciendo. Así, se designan muchos recursos para crear dependencias particulares, contratar especialistas o generar canales específicos para difundir las actividades, los procesos, los objetivos o los logros de una organización. Los motivos de tales esfuerzos pueden ser variados. Por ejemplo, en primer lugar, a través de la difusión de aquello que está realizando —o que tiene previsto realizar— una entidad es posible captar la atención de un determinado público que puede contribuir a la visibilidad de la organización y a la construcción de una imagen favorable. En segundo lugar, la difusión de aquello que se realiza en una organización se relaciona con una cierta estrategia de evaluación y vigilancia que permite hacer seguimiento de las entidades. Pero este es ya, como se verá más adelante, otro uso de la información.

Dentro del uso comunicativo, así como en los usos posteriores, hay un elemento ético sobre el que es importante reflexionar. En este caso, podemos hablar del control que suelen hacer algunas organizaciones de las comunicaciones que hacen sus funcionarios, por ejemplo, vigilando los mensajes de los correos electrónicos o restringiendo y monitoreando las

páginas a las que acceden los funcionarios desde los equipos vinculados a la organización⁸. Otro elemento destacable dentro del componente ético al respecto del uso comunicativo de la información podría ser el de los correos no deseados, o *spam*.

Finalmente, otro elemento importante dentro del componente ético tiene que ver con la no divulgación de aquella información que puede afectar la imagen de una entidad, aun si tal información hace parte de los procesos de gestión y es de vital importancia darla a conocer para cumplir con los procesos evaluativos y de rendición de cuentas. Todavía más, el encubrimiento de una determinada problemática dentro de los procesos de gestión en una entidad pública no permite alertar a las entidades correspondientes y consolidar un plan de acción que permita abordar desde una perspectiva informada tal problemática y sus efectos. En este sentido, en Colombia se ha consolidado la Ley 1712 de 2014 que pretende regular y fomentar la transparencia y el derecho de acceso a la información pública nacional. No en balde, dentro del discurso habitual de las entidades públicas se suele hacer referencia a la transparencia de datos.

Uso de registro

Consiste en describir, representar e inventariar hechos, objetos, actividades, individuos o características a través de técnicas de cálculo y registro. Dentro de esta función entran en juego procedimientos como el conteo, la medición, la crónica, entre otros. Algunos ejemplos puntuales de elementos informativos que desempeñen la función de registro son los directorios, los inventarios, los libros de contabilidad, las listas de personal, los planes anuales de vacantes, entre otros. Además, dentro de la función de registro se incluye no solo el objetivo de describir, representar e inventariar, sino también de conservar la información obtenida a través de archivos físicos o digitales y el manejo que se le debe dar a tales archivos.

Por otra parte, este uso no involucra el mero registro sin más. Hay una serie de normas, acuerdos y convenciones que suelen caracterizar los registros de información. Dicho de otro modo, el uso de la información en

⁸ Las amenazas a las que se enfrenta una sociedad permeada por la vigilancia son descritas en Zuboff (2020).

cuestión está determinado por una suerte de gramática que establece qué características son las que se registran del individuo (por ejemplo, el sexo en lugar del color de los ojos). La gramática no solo establece lo que se registra, sino también cómo se registra. En este sentido, hay una serie de estándares que cada organización establece para unificar sus registros. De hecho, hay una tendencia a establecer unos códigos compartidos a nivel mundial. Un buen ejemplo podría ser el de las normas ISO, que definen estándares internacionales. Pongamos por caso la norma ISO 3166 cuyo propósito es determinar una codificación específica y compartida a nivel internacional para nombrar territorios.

Finalmente, es importante mencionar que en la época actual es común que se registre la mayor cantidad de datos de los usuarios en internet. Así, los sitios web almacenan información sobre la ubicación, el dispositivo desde el cual se accede, el tiempo que el usuario permanece en la página, entre otros datos. Este fenómeno, como es obvio, supone la reflexión ética al respecto de qué debería y qué no debería ser registrado y, por supuesto, de qué se debería y qué no se debería hacer para obtener esos datos. Así, es posible hablar de fenómenos como la violación de la privacidad cuando se accede a datos particulares sin el permiso de los titulares, o cuando se ingresa de manera ilegal a bases de datos que almacenan información de carácter sensible⁹.

El uso de registro también supone un problema de carácter ético que tiene que ver con la concepción de la clasificación. En cierto sentido, toda clasificación deja algo por fuera y, al dejarlo, hace ver que eso que se queda por fuera no es. En palabras de Mèlich (2014) la clasificación

... elimina (por absorción) lo diferente, lo distinto, lo heterogéneo, lo extraño. Las clasificaciones ordenan y rigen a priori formas de comportamiento y de acción. Son performativas, fijan “lo que es” y prescriben “lo que debe ser”. Una lógica —una ordenación categorial— posee elementos ontológicos, normativos y formativos. Por eso no hay clasificación alguna que “pueda simplemente ajustarse a los hechos”. (p. 30)

¶ En Mayer-Schönberger y Cukier (2016) se aborda la idea de cómo el exceso de datos puede tener impactos significativos en la sociedad.

Con todo, la clasificación propia del uso de registro que se le da a la información es, quiérase o no, necesaria. A partir de la estandarización se pueden establecer comparaciones y, en general, procedimientos estadísticos que no tendrían cabida si no hubiese normas que permitieran clasificar y agrupar elementos. Podría decirse, pues, con Borges (1974), que “la imposibilidad de penetrar en el esquema divino del universo no puede, sin embargo, disuadirnos de planear esquemas humanos, aunque nos conste que estos son provisarios” (p. 708)

Uso valorativo

Se puede hablar de uso valorativo cuando el objetivo principal de la información es establecer, afirmar o reconocer un determinado juicio sobre algo. Esta función involucra elementos como la evaluación, las metas, los resultados o los incentivos.

El uso valorativo de la información es, sin lugar a dudas, una línea dominante en el ámbito de la información en lo público. Se espera que tanto la ciudadanía como los organismos de control vigilen la gestión de las entidades. Además, la evaluación se consolida como una estrategia que le permite a las entidades determinar si las actividades realizadas efectivamente corresponden a lo que se suponía debía realizarse. De este modo, se pueden establecer planes de acción para corregir cualquier falencia o cambiar las metas y los objetivos.

En este orden de ideas, es posible hablar de dos grandes tipos de evaluación. Por un lado, la evaluación externa, de la que se encargan, usualmente, organizaciones dedicadas a la vigilancia y el control, así como población civil. Por otro lado, la evaluación interna, que se refiere a las estrategias valorativas que aplica la misma entidad, como una auditoría interna.

En la actualidad, el elemento predominante dentro del uso valorativo de la información en el ámbito de lo público es el de los objetivos y los indicadores. Estos se establecen en estrecha relación con las metas de los procesos de gestión y, de acuerdo con el Departamento Administrativo Nacional de Estadística (DANE) (2011), tienen el propósito de describir características, comportamientos o fenómenos de la realidad desde los cuales pueda valorarse el plan, programa o proyecto. En este punto, un elemento de carácter ético que vale la pena considerarse es el

de la manipulación de la información con el propósito de mostrar que las metas y los objetivos de los planes, programas o proyectos se alcanzaron de acuerdo con lo previsto, aunque no sea así. Este elemento supone las problemáticas asociadas a la manipulación de datos, que van desde alterar una cifra hasta forzar un dato para que dé cuenta de algo ya decidido de antemano o, incluso, el recalco que se puede hacer de determinadas cifras a través de elementos gráficos, como tablas o barras, incluso colores o figuras, para priorizar información favorable a los procesos evaluativos que se estén llevando a cabo¹⁰.

Todavía más, el mero establecimiento de metas y objetivos conlleva un elemento ético. La gestión pública involucra una cierta intervención social, de modo que las metas y los objetivos que siga un plan, programa o proyecto definen y, si se quiere, priorizan las necesidades de una comunidad. Por ejemplo, un proyecto puede tener como objetivo entregar recursos tecnológicos en un territorio puntual, mientras que la población de ese mismo territorio no tiene condiciones suficientes para una alimentación adecuada. Por esto es que, además de la valoración de los procesos de gestión, vale la pena tener en cuenta la valoración a través de estrategias aritméticas básicas como el conteo, que permiten consolidar cifras (p. ej. el número de vinculados a una entidad o el número de personas en condición de vulnerabilidad) que tienen un valor fundamental y pertinente dentro de las diferentes valoraciones de interés para una entidad.

Finalmente, los diferentes usos de la información en la gestión pública pueden estar relacionados. Para este caso, la relación se presenta entre el uso de registro, el uso valorativo y el comunicativo. Es necesario buscar y recopilar información, ya sea a través de procedimientos aritméticos básicos (como el conteo) o a través de indicadores de gestión, para que la valoración se pueda llevar a cabo. De manera similar, junto con los procesos de evaluación puede haber un interés comunicativo, que se evidencia en el uso de informes para dar a conocer los *resultados* de cualquier etapa del proceso de consolidación y gestión de la información.

¹⁰ Huff (2023) da cuenta de cómo se pueden presentar datos de manera engañosa a través de gráficos, lo que puede llevar a interpretaciones cuestionables o manipuladas de la información.

Uso de procesamiento

A partir de un conjunto de operaciones, o procesos, se busca transformar la información existente en nueva información. Con frecuencia, tal transformación persigue el propósito de incidir en la toma de decisiones. De hecho, el procesamiento de información se constituye como una característica interpretativa e interactiva del ser humano con su entorno y con los demás. Así, frente a un elemento informativo, como podría ser el humo, una persona podría llevar a cabo un procesamiento de tal información para advertir un incendio.

Ahora bien, con la incidencia de las TIC, así como con el desarrollo del internet y la importancia concedida a la información en la época actual se han desarrollado técnicas sofisticadas de procesamiento de información como la minería de datos, el procesamiento de lenguaje natural o el *big data*; sin embargo, las actividades aritméticas básicas ya son, de por sí, herramientas de procesamiento de información.

El propósito de este libro no es caracterizar las estrategias de procesamiento de la información, ni argumentar en favor de unas o de otras. No obstante, la tendencia bastante extendida a supervalar sofisticados algoritmos en detrimento de procesamientos aritméticos básicos puede tener algunas implicaciones que vale la pena considerar: en ocasiones, la tecnología desdibuja la realidad social. Por ejemplo, podría darse una situación en la que los esfuerzos de procesamiento de información estén concentrados en, por decir algo, determinar los campos semánticos de los discursos de los gobernantes de un territorio a través de un sofisticado sistema de procesamiento de lenguaje natural, mientras que, por otra parte, se desconoce el número de personas en situación de vulnerabilidad de ese mismo territorio. Tal número resulta, justamente, del procesamiento a través de técnicas aritméticas básicas y puede ser más oportuno para un programa de intervención que beneficie a la comunidad.

El uso de procesamiento es transversal a los demás usos de la información, pues, constantemente, la información se procesa para generar algún tipo de *comunicación*, para ser *almacenada* y, por supuesto, para llevar a cabo procesos *valorativos*. Un ejemplo de procesamiento de la información puede ser la sistematización y el análisis de datos que hace el DANE (2018b) del Censo Nacional de Población y Vivienda, donde, *grosso modo*,

a partir de la información suministrada de manera individual por cada vivienda censada, se llega a conclusiones de la población nacional¹¹.

-
- || Conviene resaltar que el proceso de censo de población que adelanta el DANE involucra otros elementos más allá del procesamiento, a saber: especificación de necesidades, diseño, construcción, recolección y acopio, análisis, difusión y evaluación.

CAPÍTULO II

Las estadísticas

Hasta ahora se exploró el rol de la información en el contexto globalizado del siglo XXI, resaltando la importancia que esta tiene en ámbitos altamente tecnificados y digitalizados. Se mencionó, también, que “información” es un término polisémico que puede ser abordado desde una perspectiva etimológica y pragmática, donde se resalta la noción de *información como cosa*, que es susceptible de ser manipulada, analizada, procesada. Luego, se hizo un recorrido por los distintos enfoques desde los cuales se puede entender la información, dentro de ellos, el físico, biológico, psicológico y tecnológico.

Además, se destacó que la gestión pública, entendida como la definición y la puesta en marcha de estrategias que se han convenido para atender un asunto de carácter público, tiene una relación estrecha con la información. De modo que los diferentes modelos de gestión pública le dan una relevancia particular a la información para, por ejemplo, la rendición de cuentas o la toma de decisiones. También se precisaron tres usos que se hacen de la información en la gestión pública sin importar el modelo, a saber: comunicativo, de registro y valorativo.

Pues bien, en el presente capítulo se abordan las estadísticas, entendidas como un conjunto de *información* ordenada en cifras descriptivas que contribuyen, en gran medida, a la representación y articulación de la realidad social y, por tanto, son relevantes en el ámbito de la *gestión pública*. Se presentan, además, dos grandes grupos que caracterizan a las estadísticas: características intrínsecas y características extrínsecas. Las primeras buscan referenciar condiciones necesarias, sin las cuales no se podría hablar propiamente de estadísticas. Las segundas, por su parte, tienen que ver con recomendaciones pertinentes para las estadísticas en el campo de la gestión pública.

Es importante mencionar que no se presenta ni se propone un modelo o mecanismo de producción de estadísticas. Todavía más, el proceso estadístico, entendido como un conjunto sistemático de actividades encaminadas a la producción de estadísticas, hace parte de las características que se definen en este capítulo para las cifras estadísticas como objeto de estudio. Cada Estado o entidad puede definir una serie de fases, procesos, subprocessos y actividades para producir estadísticas que, si se consultan, es esperable —e incluso deseable— que mantengan relación con las características extrínsecas e intrínsecas de las cifras estadísticas como objeto de estudio. Sin embargo, como se hizo notar, lo que sigue a continuación no es una propuesta de lineamientos, ni un modelo para consolidación de cifras.

DEFINICIÓN

Cuando se habla de Estadística se suele hacer referencia a una disciplina que involucra una serie de técnicas de descripción, recopilación, registro y cálculo que se aplican a un conjunto de datos¹². De ahí que se tienda a ver como una *herramienta* imprescindible en el ámbito investigativo, donde desempeña, entre otras, las funciones de registrar un fenómeno, predecir resultados, analizar los registros y falsear o probar hipótesis. Este ha sido el campo predominante de la estadística en la actualidad; sin embargo, hay otros aspectos dentro de la disciplina que vale la pena tener en cuenta.

El término “estadística” viene del latín *statisticus*, que está compuesto por la raíz *status* (estado) y por el sufijo *-icus* (relativo a). La estadística se puede entender, entonces, como lo que es relativo a un Estado y, bajo esta perspectiva, se podría argumentar que es el punto en el que converge la información y la gestión pública. De hecho, como menciona Rodríguez (2012), la aparición del término a mediados del siglo XVIII hacia referencia a la descripción —no siempre numérica— de los Estados, de sus habitantes, de su geografía, de su economía, entre otros aspectos relevantes para el desarrollo y la construcción social; dicho de otro modo, tal término hacía

¹² Como se mencionó en el enfoque pragmático de la información, para Buckland (1991) el concepto de *dato* hace referencia a un tipo de información (información como cosa) que ha sido procesada para un uso determinado.

referencia a una suerte de información relevante para la organización, o gestión, social, o pública. Pues bien, para distinguir esta concepción del término, la de “hechos notables de un Estado” (p. 18), se va a emplear el término en plural de la siguiente manera: *las estadísticas*.

Así las cosas, la concepción de las estadísticas no se limita a la noción de una herramienta empleada en el desarrollo de proyectos de investigación, sino que hace referencia a una estrategia interpretativa que permite “conocer [...] una regularidad imposible de ver contando a cada individuo por separado” (Rodríguez, 2012, p. 18). En otras palabras, las estadísticas tienen que ver con una forma de discriminación y análisis descriptivo que contribuye, en gran medida, a articular los modos en que se le da forma al mundo. Definir cantidades, comparar magnitudes o establecer porcentajes, así como su correspondiente registro, permite articular una suerte de realidad susceptible de ser representada, caracterizada, comunicada o analizada, a partir de la cual se pueden llevar a cabo tareas importantes para el desarrollo social, tales como la toma de decisiones, el fomento de la transparencia y de la rendición de cuentas en la gestión, el seguimiento y la evaluación de políticas públicas o la realización de estudios e investigaciones.

Desde esta perspectiva, como se ha venido insinuando, la información que se representa a través de las estadísticas desempeña un papel protagónico en el ámbito de lo público y, en últimas, se consolida como un pilar de las sociedades democráticas. Ya desde Platón¹³ se hacía referencia al conocimiento como reivindicación de la libertad y no es casualidad que González (2021), consejera Regional de Comunicación e Información de la Unesco para América Latina y el Caribe, afirme que una sociedad más informada es una sociedad más libre. Disponer de determinadas cifras empodera a los ciudadanos y a las instituciones, por lo que resulta fundamental discutir sobre los modos en que se recopilan, ordenan, analizan y comunican esas cifras.

Ahora bien, con el propósito de clarificar —hasta donde sea posible— y, sobre todo, de fomentar la discusión al respecto de los modos

¹³ La idea del conocimiento como fuente de libertad puede rastrearse en varios diálogos platónicos, especialmente en *La República*, donde se aborda la formación de una sociedad justa y el papel de la educación en esa sociedad. Todavía más, la conocida alegoría de la caverna, presente en el libro VII (514a) de *La República*, ilustra metafóricamente el asunto del conocimiento y la libertad.

en que se abordan los “asuntos notables de un Estado”¹⁴, se presentan a continuación algunas características intrínsecas y extrínsecas de las estadísticas¹⁵.

Se presentan seis características intrínsecas. La primera tiene que ver con que las estadísticas se extraen y se soportan a través de (1) *poblaciones o muestras*, es decir, siempre se trabaja con un conjunto de individuos u objetos *susceptibles* de ser observados y analizados. La segunda se refiere a los (2) *datos*, que resultan a partir del trabajo con poblaciones o muestras y que se organizan en distintos tipos de variables cualitativas y cuantitativas susceptibles de ser almacenadas, transferidas, visualizadas, controladas y entendidas. Los datos tienen sentido en la medida en que representan una serie de regularidades basadas en conteos o medidas, por lo que la tercera característica hace referencia a las (3) *cifras agregadas* que resultan de presentar de manera resumida las regularidades encontradas. La cuarta característica precisa que se debe poder acceder a esas cifras desde diferentes perspectivas, es decir, a través de (4) *desagregaciones*, por ejemplo, temporales, geográficas/institucionales o temáticas. La quinta característica alude a la manera en la que se representan las regularidades encontradas a través de los conteos y las medidas, es decir, por medio de (5) *tablas y gráficos*. Finalmente, la sexta característica habla de los (6) *medios de divulgación*, la estadística es intrínsecamente comunicativa, los datos no son información sino hasta que son interpretados por un receptor.

Por otra parte, las características extrínsecas están relacionadas con unas particularidades propias de las cifras estadísticas en el ámbito de la gestión pública. Para empezar, las estadísticas deben ser (1) *inclusivas*. La gestión pública tiene una connotación social, por lo que las cifras deben dar respuesta a las necesidades de los diferentes individuos que conforman los diversos grupos sociales. En segundo lugar, las estadísticas deben ser (2) *comparables*¹⁶, hay estándares, como codificaciones

¹⁴ Vale la pena resaltar que la perspectiva desde la que se construyen y se abordan “los asuntos notables de un Estado” es, para el caso de las estadísticas, cuantitativa.

¹⁵ Una primera aproximación a esta labor se llevó a cabo en el libro de Rodríguez y Bernal (2019), donde se describen once características asociadas a las estadísticas que se clasifican y complementan a lo largo de este apartado.

¹⁶ Vale la pena resaltar que el ser comparables bien podría hacer parte de las características intrínsecas de las estadísticas, pues siempre buscan establecer

y nomenclaturas, que, por un lado, acotan y orientan las estadísticas y, por el otro, facilitan las comparaciones y los intercambios a diferentes niveles (p. ej. internacional, nacional, regional o interinstitucional). De la misma manera en que las estadísticas están acotadas con el propósito de ser comparables, existen normas (3) *legales* que delimitan lo que está permitido y lo que no, por ejemplo, en materia de captura, análisis y usos de la información. Además, las estadísticas deben tener en cuenta la reflexión (4) *ética* al respecto de las acciones individuales sobre lo que se debe hacer, y que puede, o no, coincidir con las leyes. La quinta característica resalta el carácter (5) *público, transparente y abierto* por el que deben velar las estadísticas, es decir, describir, informar y explicar de la mejor manera lo que se hace. La sexta característica, sobre los (6) *metadatos*, contribuye a esa pretensión de transparencia, pues versa sobre los datos de los datos, que precisan los modos en que las estadísticas deben ser leídas y entendidas. El siguiente rasgo de las estadísticas alude al (7) *uso intensivo de las TIC*. Las herramientas tecnológicas actuales facilitan la captura, el almacenamiento, la transformación, la construcción, la visualización y la comunicación de las cifras contenidas en las estadísticas. Los avances en las TIC propician nuevas mediciones y fuentes de datos, así como la construcción de mediciones cada vez más complejas. La octava característica extrínseca es, pues, la (8) *producción de nuevas estadísticas* u otros tipos de mediciones de interés en el ámbito de la gestión pública¹⁷. Finalmente, la última característica reúne, en parte, a las demás; tiene que ver con un (9) *proceso estadístico*, es decir, con una ejecución sistemática y ordenada de un conjunto de actividades que retoman las características ya mencionadas y que buscan garantizar tanto la producción como la calidad de las estadísticas institucionales.

relaciones, semejanzas y diferencias. Sin embargo, en esta ocasión se hace referencia a convenciones más de carácter social e institucional.

¹⁷ Un ejemplo de nuevas mediciones son *los indicadores*, que son objeto de discusión en el siguiente capítulo del presente libro.

CARACTERÍSTICAS INTRÍNSECAS

Población o muestras

Las estadísticas se extraen y se soportan a través de poblaciones, es decir, a través de colecciones de unidades finitas e infinitas que, con frecuencia, corresponden a personas, pero que pueden ser también instituciones, eventos, objetos, entre otros elementos que sean de interés cuantitativo. En ocasiones, se elige un subconjunto de la población que recibe el nombre de *muestra*. Se pretende que la muestra, o subconjunto, sea *representativa* de la población. Es decir, que a través de técnicas probabilísticas se garantice una cierta aleatoriedad en la conformación de muestras que permita formular generalizaciones que se cumplan en la población con cierto nivel de certeza. Sin embargo, en el ámbito de la gestión pública, dadas las condiciones que afectan los modos en que se obtienen los datos, suele haber muestras no probabilísticas¹⁸ que, no obstante, son de utilidad para la consolidación de las estadísticas.

Las poblaciones pueden ser de tipo *transversal, anidado y longitudinal*. Las *transversales* se obtienen en un punto particular en el tiempo, sin tener en cuenta el recorrido histórico de los elementos que hacen parte del conjunto. Las *anidadas* tampoco tienen en cuenta el recorrido temporal de los elementos del conjunto, sin embargo, una población anidada hace parte de otra población más grande. Las *longitudinales*, a diferencia de las anteriores, tienen un carácter diacrónico, es decir, tienen en cuenta la historicidad de los elementos que conforman el conjunto.

Cuando se habla de poblaciones o muestras, es necesario hablar de *variables*. Las variables hacen referencia a las características que comparten los miembros del conjunto o del subconjunto y pueden ser clasificadas en dos grandes grupos: las cualitativas y las cuantitativas.

Las variables cualitativas, o categóricas, se pueden dividir, a su vez, en nominales y ordinales. Las variables cualitativas nominales hacen referencia a las *cualidades* de los individuos que conforman¹⁹ una población

¹⁸ En Wolf *et al.* (2016) se presenta de manera breve lo que concierne al tema de las muestras no probabilísticas.

¹⁹ Resulta importante anotar que los individuos que conforman una población o muestra pueden ser no humanos.

o muestra que no admiten un criterio de orden y que no representan un valor numérico, como el color de ojos, el nombre o la carrera universitaria que un estudiante ha escogido. Las variables cualitativas ordinales, por su parte, tienen que ver con cualidades que tampoco representan un valor numérico absoluto, pero que sí tienen un criterio de orden definido, como el estrato socioeconómico en Colombia²⁰ o el nivel de satisfacción de un cliente. Dicho de otro modo, las variables ordinales clasifican las características de una población o muestra en un orden determinado, pero dicho orden no representa una distribución igual en los intervalos que componen la clasificación. Así, una variable cualitativa ordinal permite determinar que un cliente está más satisfecho que otro, pero no la *cantidad* de satisfacción que tiene de más.

Por su parte, las variables cuantitativas sí representan un valor numérico absoluto y se pueden clasificar en continuas y discretas. Las cuantitativas discretas hacen referencia a números enteros (usualmente positivos), como el número de hijos de una familia o el número de admitidos a un programa curricular en una universidad. Las variables cuantitativas continuas pueden ser de intervalo y de razón. Para el caso de las de razón, el cero significa la ausencia del atributo analizado y, para el caso de las de intervalo, el cero no es la ausencia del atributo, sino un valor que toma dentro del rango. Un ejemplo de variable de razón podría ser el tiempo de vinculación de un estudiante a una institución educativa. Si el tiempo es igual a cero significa que el estudiante no está vinculado a la institución, es decir, representa una ausencia del atributo. En cambio, la nota que obtienen los estudiantes en un examen podría considerarse una variable de intervalo. Si un estudiante tiene una calificación de cero no quiere decir necesariamente que no haya presentado el examen o, mejor, no implica necesariamente una ausencia del atributo, sino que obtuvo la peor calificación posible.

Dentro de las características que comparten los miembros de las poblaciones o muestras es posible disponer de variables cuantitativas con características particulares: las espaciales y las temporales. Las espaciales se representan a través de coordenadas que designan un punto geográfico particular de acuerdo con un sistema de referencia, como la latitud y la

²⁰ Según la página del DANE, la estratificación socioeconómica en Colombia se refiere a la clasificación de inmuebles residenciales con el propósito de cobrar de manera diferencial los servicios públicos domiciliarios (DANE, s. f.a).

longitud. Por su parte, las temporales designan, a través de convenciones, como meses o años, un punto particular en una serie de tiempo.

Existe la posibilidad de transformar variables cuantitativas en cualitativas —en especial, variables cualitativas de tipo ordinal—, pero no al revés. Por ejemplo, una variable cuantitativa, como la estatura de una persona, puede ser entendida como una variable cualitativa de naturaleza ordinal a partir de la definición de rangos o clases. Así, la medida de 177 centímetros de estatura de una persona puede ser entendida como una variable cualitativa cuando se hace referencia a que dicha persona hace parte del rango de individuos que miden entre 170 y 180 centímetros.

Finalmente, en el contexto de lo público, las poblaciones y las muestras se conforman principalmente a partir de tres estrategias de recolección de datos: los censos, las encuestas y los registros administrativos.

Los censos se aplican a la totalidad de individuos que conforman una población, por lo que representan un alto costo en términos de tiempo y recursos; están reservados para ser adelantados a nivel de países o sectores económicos dentro de dichos países. Las encuestas, por el contrario, se aplican principalmente a una muestra, es decir, a un grupo parcial de la población. Y, dada su facilidad de ejecución con respecto a los censos, es la principal estrategia de los Estados, aplicada por medio de los institutos nacionales de estadística o sus sectores, para la captura y consolidación de estadísticas sobre diversos temas de interés social.

En la actualidad, con la variedad de datos disponibles a bajo costo, dada la evolución de las TIC, en los Estados y sus instituciones ha cobrado relevancia la conformación de muestras o poblaciones a través de los registros administrativos, es decir, mediante los datos que ya recopilan las entidades públicas como parte de sus obligaciones legales²¹. Así, los registros administrativos son, por excelencia, el principal mecanismo de conformación de muestras o poblaciones para la definición y consolidación de estadísticas a nivel de las entidades públicas y, dentro de estas, en las universidades, por ejemplo²².

²¹ En Colombia, el Departamento Administrativo Nacional de Estadística (DANE) cuenta con una metodología para el aprovechamiento estadístico de registros administrativos (DANE, 2018a).

²² Dado que estos registros se crean con fines administrativos y no estadísticos, es necesario llevar a cabo un proceso de fortalecimiento para que puedan ser utilizados con propósitos estadísticos.

Datos

Para que las variables cualitativas y cuantitativas que contienen los rasgos de los individuos que conforman las poblaciones o muestras puedan ser almacenadas, transferidas, visualizadas y analizadas, es necesario entender y reconocer que estas se encuentran conformadas por datos²³, es decir, por “información como cosa que se ha procesado de alguna manera para su uso” (Buckland, 1991, p. 353). Pues bien, los datos se suelen dividir en tres tipos: estructurados, semiestructurados y no estructurados.

Los datos estructurados son aquellos almacenados en tablas formadas por filas y columnas que, comúnmente, buscan facilitar su lectura y procesamiento. Sin embargo, un conjunto de datos se puede estructurar de maneras diferentes. Por esto es que, dentro del mundo de los datos estructurados, también están los datos ordenados²⁴. Para Wickham (2014), un conjunto de datos ordenado es un conjunto que sigue una forma estándar de estructuración para optimizar el análisis, que consiste en que cada variable, o atributo, forma una columna; cada observación, es decir, cada vez que se identifican las variables juntas, es una fila y cada tipo de observación, es decir, cada conjunto de variables que se identifican de la misma manera conforma una tabla.

A continuación, se muestra una tabla con la información correspondiente a la cantidad de estudiantes de una universidad que hacen parte de diferentes programas curriculares en distintas sedes. La tabla es estructurada, pero no ordenada.

²³ Para un lector que valore la precisión, podría parecer necesario establecer una distinción explícita entre datos y variables. Sin embargo, en un contexto más práctico, y para un usuario común, no especializado, resulta más sencillo referirse a esto como datos. En última instancia, es importante comprender que las variables de una población o muestra se basan en datos.

²⁴ La producción de estadísticas requiere la disposición de datos ordenados (*tidy data*) en las poblaciones o muestras. No obstante, como se mencionó, no todo conjunto de datos estructurado es ordenado por lo que estos, en ciertas situaciones, deben ser manipulados y transformados con el fin de alcanzar el ordenamiento que requiere la construcción y disposición de las estadísticas.

Tabla 1. Datos estructurados no ordenados

Programa curricular	Lingüística Sede A	Medicina Sede B	Ingeniería de Sistemas Sede A	Derecho Sede B
Inscritos	150	250	325	200

La tabla 2 presenta la misma información de la primera, pero a través de las reglas de datos ordenados. Así, se conforman 3 *variables* y 4 *observaciones* que constituyen una *tabla* ordenada y estructurada.

Tabla 2. Datos estructurados ordenados

Observaciones	Variables		
	Programa curricular	Sede	Inscritos
Lingüística	A	150	
Medicina	B	250	
Ingeniería de Sistemas	A	325	
Derecho	B	200	

La producción de estadísticas en el ámbito de lo público se vale, sobre todo, de datos estructurados y ordenados. Lo anterior en la medida en que tradicionalmente las técnicas y las herramientas de procesamiento de datos, desde las más simples a las más complejas, están pensadas para datos estructurados ordenados.

Por su parte, los datos semiestructurados están ligados al desarrollo de la web. No tienen una estructura estándar, pero siguen un flujo lógico, con etiquetas y marcadores que permiten identificarlos. Algunos ejemplos podrían ser los correos electrónicos o los *tweets*.

Finalmente, los datos no estructurados no tienen una estructura estándar, no siguen una secuencia lógica fácilmente definible o etiquetable y no se pueden almacenar dentro de una tabla convencional, por

ejemplo, los documentos multimedia o los textos. Ahora bien, como se sabe, la mayoría de los datos en la actualidad son, de hecho, no estructurados, cuyo análisis puede contribuir en gran medida a la consolidación de estadísticas y, con ello, a procesos ya mencionados como la toma de decisiones, el fomento de la transparencia, el seguimiento y la evaluación de políticas públicas, entre otras. De manera que resulta relevante preguntarse si vale la pena leer los datos no estructurados desde el mundo de lo estructurado, o si al hacerlo una buena parte de su contenido se queda en el camino y, de ser así, es importante reflexionar al respecto de qué otro tipo de estrategias de procesamiento podrían aplicarse para la consolidación y disposición de información contenida en este tipo de datos.

Cifras agregadas

Los datos —y las variables en las que dichos datos se organizan— contienen regularidades que no pueden ser percibidas a nivel individual²⁵. Tales regularidades, en el contexto de las estadísticas, se consolidan a través de conteos o mediciones. Los conteos hacen referencia a las estadísticas derivadas de variables cualitativas y, en muchas ocasiones, a variables cuantitativas de naturaleza discreta, que no pueden tomar ningún valor entre dos consecutivos²⁶. Mientras que la medición deriva de variables cuantitativas o continuas, que pueden tomar cualquier valor dentro de un intervalo. Así, contar tiene que ver con enumerar un conjunto de elementos comunes, y medir hace referencia a establecer una cierta escala para determinar la magnitud o la dimensión de un rasgo de interés social y estadístico²⁷.

-
- 25 Uno de los elementos más relevantes de la conformación de estadísticas es que, a través de las cifras agregadas, es posible describir elementos de la realidad que no se podrían identificar viendo cada uno de manera aislada. En este sentido, las cifras agregadas se constituyen como un componente de las estadísticas que contribuye en gran medida a los modos de articular e interpretar la realidad.
 - 26 Vale la pena mencionar que existe un conteo intrínseco a las estadísticas que es el total de individuos que conforman una población, sin embargo, el resultado de dicho conteo no es necesariamente una variable.
 - 27 La distinción entre contar y medir puede parecer, en ocasiones, no tan definida. Por ejemplo, se podría decir que, de algún modo y en algún sentido, al contar se está haciendo un ejercicio de medición. Sin embargo, para la construcción de cifras estadísticas, resulta útil hacer la diferenciación.

En el ámbito de la construcción y la determinación de regularidades en poblaciones y muestras, son de amplio interés los modos en que se *distribuyen* los datos que conforman dichas variables. Las medidas²⁸ estadísticas descriptivas asociadas a dicha distribución se suelen clasificar en cuatro grupos: *medidas de tendencia central*, que sirven para determinar los valores centrales o medios de la distribución (como la media y la moda); *medidas de dispersión*, que indican la variación de la distribución (como la desviación media, la varianza y la desviación estándar); *medidas de localización*, que dividen la distribución en partes iguales (como la mediana, los cuartiles, los deciles o los centiles); y, finalmente, las *medidas de forma o simetría*, que permiten precisar el grado de apuntamiento, simetría o asimetría de una distribución (como la curtosis o la asimetría).

Desagregaciones

En la construcción y disposición de estadísticas, las desagregaciones tienen que ver con desglosar y presentar la información en grupos más pequeños que buscan identificar tendencias y patrones en “las partes” para entender mejor “el todo”.

Estas pueden ser de tres tipos: temporales, geográficas y temáticas. Las desagregaciones temporales consisten en agrupar la información de acuerdo con series de tiempo que permiten determinar el comportamiento de una cifra estadística en el pasado, en el presente y, a partir de tal conocimiento, estimar el comportamiento futuro de dicha cifra.

Una serie de tiempo puede entenderse como una secuencia de cifras históricas separadas en intervalos de tiempo regulares. Por ejemplo, en el marco de las estadísticas asociadas a las universidades es relevante representar la evolución de admitidos a lo largo de la historia a través de un intervalo de tiempo regular como lo son, para este caso, los semestres. En el contexto de la gestión pública, la longitud de los intervalos que definen una serie de tiempo, también conocida como temporalidad de la serie, puede variar de acuerdo con la necesidad social de la disposición de estadísticas (meses, bimestres, trimestres, semestres, anualidades, etc.).

²⁸ Las medidas estadísticas descriptivas para la conformación de cifras agregadas se pueden aplicar a variables cualitativas y cuantitativas o, dicho de otro modo, se pueden medir los conteos y se pueden medir las mediciones.

Las desagregaciones geográficas tienen que ver con la distribución de las estadísticas extraídas de una población o muestra a lo largo de un territorio o de una entidad. La *desagregación geográfica territorial* en el contexto de la gestión pública hace referencia principalmente a la forma en la que se distribuyen las estadísticas a lo largo de las divisiones político-administrativas de los países. Por ejemplo, para el caso de Colombia²⁹, la nación se divide política y administrativamente en departamentos, distritos especiales, municipios, corregimientos, centros poblados, entre otros. La *desagregación geográfica institucional*, por su parte, se refiere a la distribución en organizaciones, como universidades y, dentro de estas, las facultades o programas y demás subgrupos institucionales definidos por límites basados en funciones o tareas. Un ejemplo de *desagregación geográfica territorial* podría ser la segmentación de la población de una universidad a partir de los lugares de nacimiento de los estudiantes. La distribución de esos estudiantes en departamentos y carreras específicas sería una *desagregación geográfica institucional*.

La información ordenada en cifras descriptivas contenida en las estadísticas se puede desglosar también en características que no hacen parte de las desagregaciones temporales ni de las geográficas, por ejemplo, la edad, el género, el estrato, el nivel de formación, entre otras características que sean relevantes para la conformación de cifras estadísticas. Para ilustrar, en el caso de la UNAL, resulta relevante contar con la desagregación de los estudiantes admitidos de acuerdo con el tipo de admisión, que puede ser, por un lado, regular, cuando un estudiante presenta el examen y gana un cupo en el programa para el que se inscribió o, por el otro, especial, que busca favorecer minorías poblacionales o poblaciones apartadas de los centros de desarrollo del país a través de acciones afirmativas³⁰.

Un elemento que vale la pena destacar es que las desagregaciones se pueden combinar de acuerdo con el interés estadístico. Por ejemplo, puede haber desagregaciones temporales que, a su vez, tengan en cuenta desagregaciones geográficas e, incluso, desagregaciones temáticas. Así,

-
- 29 Cada país tiene su propia división política y administrativa del territorio, por lo que su composición puede variar.
 - 30 Es decir, las acciones implementadas con el propósito de promover y fortalecer a ciertos individuos o colectivos, con el objetivo de eliminar o disminuir disparidades de género, socioculturales o económicas que les afectan.

una segmentación que incluiría los tres tipos de desagregaciones podría ser aquel grupo de estudiantes que ingresaron por admisión especial (temática) a los programas curriculares del departamento de medicina (institucional) en los últimos seis semestres (temporal).

Representación gráfica y tabular

La implementación de recursos visuales en los procesos de comunicación ha sido una constante en el desarrollo de la civilización. Tal parece que, dentro de las interacciones humanas, lo gráfico se consolida como una estrategia de representación y de apropiación de la realidad ampliamente empleada y valorada que permite dar cuenta de los modos en que se habita y se percibe el mundo.

El desarrollo de las representaciones gráficas ha traído consigo la conformación de nociones y normas estéticas que condicionan las producciones visuales. De ahí que sea posible discutir al respecto de técnicas como la pintura, los grabados, el dibujo o la fotografía. Ahora bien, cualquier recurso gráfico puede ser entendido, en sentido amplio, como un recurso representativo o comunicativo. Así, mientras que una pintura impresionista de principios del siglo XX puede comunicar, o representar, las ideas sobre la libertad individual de un pintor, los recursos de visualización que se presentan a continuación comunican, o representan, cifras agregadas —es decir, las estadísticas— extraídas a partir de la disposición de los datos contenidos en las variables cuantitativas y cualitativas que hacen parte de las muestras o poblaciones asociadas.

Pues bien, existen dos grandes grupos de recursos para visualizar cifras estadísticas, a saber: los gráficos estadísticos y las tablas. Los gráficos estadísticos suelen contener menor información y un aparente menor grado de precisión en las cifras que representan, pero permiten destacar hechos notables, regularidades o tendencias que difícilmente podrían observarse bajo otra estrategia de visualización de la información contenida en las estadísticas. En contraste, las tablas permiten disponer de un mayor número de cifras con un alto nivel de precisión; no obstante, es difícil detectar tendencias, regularidades o hechos notables en las cifras dispuestas bajo esta estrategia de divulgación. Ahora bien, el ejercicio de divulgación de las estadísticas requiere representar el mayor número de información con los mayores niveles de precisión, así como

la posibilidad de detectar y representar hechos notables, tendencias o regularidades sociales perceptibles visualmente. De modo que tanto los gráficos estadísticos como las tablas son importantes.

Algunos de los gráficos estadísticos más comunes empleados en la visualización de estadísticas son los *gráficos de barras*, que hacen énfasis en la comparación entre elementos en un periodo de tiempo específico; *gráficos de línea*, que muestran las relaciones de los cambios en los datos en un periodo de tiempo; *gráfico circular o pie*, que se utilizan para mostrar cómo diferentes partes representan un total; *gráficos de dispersión o scatter plot*, son útiles para mostrar la relación entre diferentes puntos de datos. Tales gráficos se suelen emplear de acuerdo con los tipos de variables que conforman las poblaciones o muestras de origen, así, el *gráfico de líneas* para representar variables temporales; el *gráfico de barras* o el *pie chart* para la representación de variables cualitativas; los *box plots* o *histogramas* para la representación de variables cuantitativas (intervalo o razón); los mapas para la representación de la información estadística de manera territorial, entre otros.

Por otro lado, las tablas están conformadas, básicamente, por *filas* y *columnas* que se ordenan de maneras diferentes para representar la información. La intersección de una fila con una columna recibe el nombre de *celda*. Otros elementos que comúnmente se encuentran en las tablas son los *títulos*, que permiten nombrar e identificar la tabla y los *encabezados*, que describen el contenido de las filas o columnas. Las tablas, en comparación con los gráficos, pueden ser una mejor estrategia para representar varias desagregaciones a la vez.

Mecanismos de divulgación y de comunicación

Desde las primeras apariciones del término, es decir, como los “hechos notables de un Estado”, las estadísticas han estado ligadas a la divulgación y la comunicación. La discriminación y el análisis de regularidades, así como su sistematización y representación en tablas y gráficos tienen sentido en la medida en que pretenden *informar* algo a alguien. De modo que la última característica intrínseca de las estadísticas consiste en la necesidad de divulgar y comunicar las cifras, así como los análisis que de ellas se sigan.

En el ámbito de la gestión pública, por ejemplo, el proceso de divulgación y de comunicación es el que más tiene relevancia para la población no especializada, pues es la única manera que tiene de enterarse de las cifras estadísticas y de sus implicaciones. De hecho, conceptos como transparencia institucional o datos abiertos solo tienen sentido a través de los modos en que se articulan los medios de divulgación y comunicación.

Simplificando y adaptando el modelo tradicional de Shannon y Weaver (1949), de lo que se trata es de una *fuente de información* (como podría ser un conjunto de datos estructurado y ordenado) a partir de la cual se codifica un *mensaje* que se transmite por un determinado *canal*, que bien podría ser el cable del teléfono o el aire por el que viajan las ondas, pero en un sentido más amplio sería también el correo institucional de una universidad o sus perfiles en las redes sociales. El mensaje se transmite a un *receptor*, que podría ser un grupo general de estudiantes de una universidad o el grupo particular de egresados de un determinado programa curricular. Y entre la señal que se emite y la señal que se recibe, se establece una fuente de *ruido* o *interferencia* que afecta en mayor o menor medida el contenido del mensaje. Cuanto más clara la información, menor posibilidad de que se vea afectada por el ruido.

Así, es de especial importancia que dentro de los procesos de comunicación de las estadísticas se tengan bien definidas las fuentes de información y que la codificación del mensaje no se tome a la ligera. De hecho, como se mencionó, la implementación de un gráfico u otro puede contribuir a robustecer o debilitar el mensaje. Asimismo, la codificación del mensaje debería tener en cuenta al receptor para así reducir la posibilidad de ruido. No es lo mismo enviar un boletín informativo a los estudiantes matriculados por primera vez, es decir, que hasta ahora empiezan en la universidad, que a un grupo de docentes de carrera con amplia trayectoria y conocimiento de la institución. Con respecto al canal, resulta evidente que un medio puede ser más eficaz que otro para enviar determinado tipo de mensaje. Por ejemplo, un mensaje que tenga que ver con la cantidad de plazas disponibles para docentes en una universidad puede ser más eficaz si se transmite a través de una red social empresarial como LinkedIn que a través de una red social como Instagram.

Wiener (1961) añade al modelo de Shannon y Weaver (1949) la noción de *feedback*, que se refiere a la reacción que causa el proceso de comunicación en el receptor. El *feedback* sirve para distinguir los mecanismos

de divulgación de los mecanismos de comunicación. Mientras que los medios de divulgación no tienen en cuenta el *feedback*, como un folleto impreso, donde no se espera que el receptor pueda retroalimentar el proceso comunicativo, los medios de comunicación tienen en cuenta e, incluso, fomentan la retroalimentación.

En la actualidad, es común disponer de mecanismos de divulgación estadística, que pueden ser físicos (como los boletines estadísticos impresos) o digitales (como los *dashboards*, cubos, las presentaciones, las infografías, etc.). Sin embargo, es importante avanzar hacia estrategias que permitan que esos mecanismos de divulgación sean, a su vez, mecanismos de comunicación, es decir, que además de divulgar, permitan la retroalimentación de los usuarios. Una estrategia notable para tal propósito es la implementación de redes sociales, pues permiten alojar de manera eficiente diversos mecanismos de divulgación y, a la vez, fomentan la interacción a través de comentarios o reacciones.

CARACTERÍSTICAS EXTRÍNSECAS

Como se ha visto hasta ahora, las estadísticas tienen, cuando menos, seis características intrínsecas, esto es, indispensables para llamarse propiamente estadísticas. Asimismo, hay siete características extrínsecas que no son menos importantes; al contrario, contribuyen en gran medida a definir el sentido de las estadísticas en el ámbito de la gestión pública.

Incluyentes

Una exigencia para las estadísticas en el ámbito de la gestión pública es reconocer y responder a la diversidad. Tanto las poblaciones o muestras estudiadas como las comunidades de difusión y comunicación tienen unas características y necesidades particulares que deben tenerse en cuenta.

Vale la pena resaltar que, tradicionalmente, las estadísticas se han caracterizado por lo que se podría denominar como poco incluyentes. El análisis de grupos extensos de datos y números parece desdibujar la realidad de grupos pequeños, poco representativos, que se sugieren insignificantes al lado de los grupos grandes en la medida en que no

afectan la tendencia general. Sin embargo, las estadísticas enmarcadas en la gestión pública no pueden desligarse de los derechos y las identidades de las personas que analiza³¹.

Echeita (2013) presenta una estructura para entender la inclusión en el ámbito pedagógico que puede ser de bastante utilidad para abordar la característica incluyente de las estadísticas en lo público. Para el autor, la inclusión tiene cuatro elementos fundamentales: primero, *acceso*, se debe velar por que todas las personas de una población puedan buscar, consumir y entender las estadísticas, sin importar las limitaciones físicas, mentales, sociales o económicas; segundo, *participación*, es decir, que todas las poblaciones puedan ser representadas en la construcción de estadísticas sin discriminaciones por sexo, raza, origen, religión, opinión política, entre otras; tercero, la inclusión como *valor* o cualidad, esto es, como un principio que se manifiesta en comportamientos y actitudes de las personas y que se puede fomentar y medir³²; finalmente, cuarto, la inclusión desde una perspectiva de *propósito social*, que consiste en la idea de una sociedad incluyente como propósito fundamental de la construcción de estadísticas, es decir, que la consolidación de cifras contribuya, por ejemplo, a la formulación y promoción de políticas y prácticas de inclusión.

La inclusión en las estadísticas ya ha sido un tema de interés nacional en Colombia. El DANE presenta una serie de lineamientos que buscan impulsar la inclusión de un “enfoque diferencial e interseccional”³³ en los procesos de producción de información estadística en Colombia (DANE, 2020^a, p. 6). La guía se centra en cinco principios de la carta de datos inclusivos: primero, todas las poblaciones deben incluirse en los datos;

³¹ La característica extrínseca en consideración aplica para estadísticas basadas, sobre todo, en poblaciones humanas. Sin embargo, como se mencionó, aunque tales poblaciones son centrales en la producción de estadísticas no son las únicas en el escenario público.

³² Por ejemplo, la inclusión, entendida como valor o cualidad, se puede fomentar a través de políticas estatales de integración y reconocimiento. Además, se puede medir por medio de, por ejemplo, la frecuencia de participación de todas las poblaciones de una comunidad universitaria en actividades curriculares y extra-curriculares.

³³ Este enfoque pretende obtener y difundir información sobre grupos poblacionales con una o varias características particulares, como pueden ser pertenencia étnica, género, discapacidad, etapa del ciclo vital, entre otras.

segundo, todos los datos deben desagregarse³⁴, siempre que sea posible, para describir todas las poblaciones; tercero, los datos deben producirse en todas las fuentes posibles garantizando los principios de calidad de la información; cuarto, las personas encargadas de la recopilación de datos y elaboración de estadísticas deben hacerlo con total transparencia y responsabilidad; quinto, debe mejorarse constantemente la capacidad humana y técnica para recopilar, analizar y usar datos, especialmente mediante una financiación adecuada y sostenible.

Comparables

La segunda característica extrínseca se refiere a la tendencia de las estadísticas en el ámbito de lo público a buscar estrategias para garantizar la comparabilidad a nivel institucional, nacional e internacional, esto es, para garantizar la capacidad de ser comparadas, contrastadas y evaluadas en función de ciertas características o atributos similares. De hecho, la comparabilidad en términos generales es una condición necesaria para el desarrollo de las estadísticas³⁵. Sin definiciones precisas sobre las escalas, los valores o los elementos que van a ser sujetos de análisis estadístico es imposible realizar procesos de comparación que redunden en alguna cuantificación o medición fiable.

La comparabilidad de las estadísticas en el campo de la gestión pública se pone de manifiesto en cuatro elementos relevantes, a saber, las normas, la unificación, la especificación y la simplificación. En primer lugar, las estadísticas se construyen en un marco normativo que las condiciona. Cada Estado, institución u organismo tiene una serie de *normas* sobre las cuales es posible el desarrollo de las estadísticas en lo público.

-
- ³⁴ Como se mencionó en la primera parte del capítulo, las desagregaciones hacen parte de las características intrínsecas de las estadísticas y pueden ser temporales, temáticas o geográficas. Pero, además, en la medida en que segmentan a las poblaciones o muestras, permiten representar, describir y visibilizar subgrupos, fomentando así la inclusión.
 - ³⁵ Hay estandarizaciones intrínsecas a la noción de estadísticas, sin las cuales no podría darse el ejercicio estadístico como lo conocemos. Por ejemplo, las mediciones o los conteos a lo largo del tiempo son, en sí mismos, comparaciones de datos obtenidos en diferentes momentos. Asimismo, las unidades de medida, como los metros o las millas, son, en el fondo, convenciones estandarizadas que permiten un lenguaje común.

En segundo lugar, se busca la *unificación* a través de estándares, como codificaciones y nomenclaturas, que permiten las comparaciones a diferentes niveles. Por ejemplo, la clasificación internacional universitaria de las asignaturas de los programas curriculares en créditos hace que las estadísticas correspondientes a las notas de un estudiante en España puedan llegar a ser comparables y, por lo tanto, representativas y convencionales en Colombia. Otro ejemplo de *unificación* es el estándar nacional Divipola (División Política y Administrativa) (Datos abiertos Colombia, s. f.) en Colombia, que define y codifica a través de un código numérico la manera en la que se divide política y administrativamente el territorio, esto es, en entidades territoriales (departamentos, distritos y municipios), áreas no municipalizadas y centros poblados.

En tercer lugar, la comparabilidad se expresa a través de la *especificación*, que consiste en la implementación de un lenguaje claro y conciso para evitar ambigüedades y errores de identificación. Asimismo, la especificación tiene que ver con la definición precisa de cualquier concepto institucional o interinstitucional involucrado en los procesos de cuantificación o medida.

Por último, tanto para las normas como para las estrategias de unificación y de especificación, se recomienda velar por la *simplificación*. Es decir, que las normas, las codificaciones, las nomenclaturas y las definiciones sean tan abreviadas como sea posible.

Legales y éticas

Como se mencionó, las estadísticas se circunscriben según una normatividad que las sustenta y, por ende, soporta la relevancia de la información generada; es decir, están sujetas a leyes institucionales, estatales o internacionales, como protección de menores, privacidad, protección de datos, *habeas data*, ley de transparencia o de datos abiertos, entre otras. El carácter legal de las estadísticas consiste, pues, en la manera en la que los procesos de conformación y comunicación de cifras descriptivas respetan las leyes vigentes. En este sentido, resulta de especial importancia velar por la formación al respecto de las diferentes normativas involucradas en los procesos estadísticos. En Colombia, por ejemplo, se cuenta con el documento del Consejo Nacional de Política Económica y Social (Conpes 3995) (DNP, 2020), que busca mejorar la confianza y la seguridad de la

información a nivel digital a través de determinadas regulaciones contenidas en una política nacional.

Dentro de los marcos normativos se suele poner en evidencia una tensión en el ámbito de las estadísticas entre el derecho particular y el derecho general, esto es, entre el derecho a estar informado y el derecho a la privacidad. Cualquier ejercicio de las estadísticas en el ámbito de la gestión pública debería reflexionar al respecto de la prioridad que se le da a un elemento o a otro en cada caso. En Colombia, la tensión se manifiesta, por ejemplo, a través de las implicaciones, en la práctica³⁶, de la Ley 1712 de 2014 —Ley de transparencia y del derecho de acceso a la información pública nacional— y el *habeas data*, que es un derecho fundamental consagrado en el artículo 15 de la Constitución Política de Colombia, la Ley 1581 de 2012 y sus normas concordantes. Por un lado, la Ley 1712 resalta que la información relacionada con cualquier entidad, organismo o persona que desempeñe una función pública esté a disposición de todos los ciudadanos e interesados de manera oportuna, veraz, completa, reutilizable, procesable y en formatos accesibles y, por el otro, el *habeas data* protege, regula y reglamenta el uso y tratamiento de la información privada de las personas.

Además de las normas que condicionan la consolidación, la difusión y el acceso al conjunto de información ordenada en cifras descriptivas —esto es, a las estadísticas— hay una serie de desafíos propios en la consolidación de estadísticas que suponen retos de carácter ético, por ejemplo, la implementación de metodologías adecuadas para capturar, procesar y representar los datos; proteger la confidencialidad de una muestra o población determinada; buscar la objetividad en el análisis; evitar resultados predeterminados, entre otros.

Para el caso particular del contexto colombiano, el DANE cuenta con un Sistema de Ética Estadística (DANE, 2021) que estructura la práctica estadística en lo público a través de cuatro ejes y principios éticos. El primer eje, la dignidad humana, la confidencialidad y el enfoque diferencial, que tiene que ver con la protección de los derechos fundamentales de las personas y el respeto de los derechos a la intimidad personal y familiar;

³⁶ Al realizar un análisis detallado de los principios y las excepciones de las leyes y normas concordantes, es posible que la tensión parezca no ser tan evidente. No obstante, en la práctica, es bastante común.

el control de los individuos sobre sus datos personales; la garantía de un enfoque diferencial en las estadísticas, y el respeto a la dignidad humana. El segundo eje, la imparcialidad, transparencia y publicidad, que consiste en la transparencia y rendición de cuentas como principios garantes de acceso a la información pública y de control social; la participación de los actores involucrados; la garantía de integridad en el tratamiento de los datos y publicación de resultados; la oportunidad y el acceso equitativo a la información y la imparcialidad en la producción estadística. Tercero, la legitimidad y generación de valor público, que destaca la contribución de los datos al bien público; la generación y promoción de la confianza en las estadísticas; y el ajuste de la producción estadística a los marcos legales. Finalmente, el cuarto eje es sobre el uso seguro de las tecnologías actuales y futuras, que alude a la ponderación de las implicaciones y los riesgos de los datos colaborativos y a la promoción de la supervisión humana, considerando los límites y riesgos del uso de nuevas tecnologías de información.

Públicas, transparentes y abiertas

La cuarta característica extrínseca tiene que ver con el carácter público, transparente y abierto que debería tener la información. Es decir, que cualquier persona esté en capacidad de acceder de manera libre y sin restricciones a los datos sobre sí misma y sobre su entorno. Además, los datos relacionados con la función pública, respetando las restricciones legales, deberían poder ser reutilizados y redistribuidos para propósitos diferentes a los originales³⁷. En este punto, es importante mencionar que para la reutilización de los datos no basta con el acceso a estos; además de ser publicados deberían cumplir características específicas para que puedan ser reutilizados fácilmente, como, por ejemplo, ser datos ordenados³⁸.

Las estadísticas deben ser abiertas y transparentes en el ámbito de la gestión pública en la medida en que contribuyen a la verificación

³⁷ En el presente libro ya se resaltó la noción de “bienes no rivales” que caracteriza a los datos. Pues bien, los datos agregados —es decir, las estadísticas— no son la excepción; pueden emplearse por diferentes usuarios y para propósitos distintos cada vez.

³⁸ Los datos ordenados son datos que siguen una forma estándar de estructuración descrita en el apartado de *Características Intrínsecas* en la sección de *Datos*.

del cumplimiento de funciones, lo que evita la corrupción y promueve el empoderamiento y la participación ciudadana en la solución de problemas públicos.

Para que los datos puedan ser públicos, transparentes y abiertos es importante tener en cuenta el proceso de construcción de estadísticas, así como los mecanismos para su difusión y comunicación. En otras palabras, la conformación adecuada de poblaciones o muestras, la estructuración ordenada de los datos, las desagregaciones representativas y detalladas, el uso de estándares y las estrategias de divulgación y comunicación favorecen el acceso libre y sin restricciones a los datos. Lamentablemente, esto no siempre se da en la práctica; en ocasiones, hay datos que son puestos a disposición del público solo por cumplir un requisito, esto es, sin tener en cuenta su calidad³⁹. Esto, al contrario de contribuir a la transparencia o la apertura, afecta el valor de los datos y su utilidad.

En el caso particular del contexto colombiano, se establece la ya mencionada Ley 1712 de 2014, que fomenta la característica pública, abierta y transparente de la información. Así, las entidades públicas y las personas naturales o jurídicas que desempeñen una función pública están en la obligación, entre otros deberes, de publicar la información de naturaleza cuantitativa en el portal de datos abiertos⁴⁰ del Estado colombiano⁴¹. Por su parte, en el contexto internacional, un buen ejemplo de datos abiertos podría ser el sitio web de datos abiertos UNdata de la Organización de las Naciones Unidas (ONU), donde están publicadas algunas estadísticas representativas de diferentes países⁴².

Ahora bien, el rasgo público de las cifras estadísticas pone de manifiesto el asunto abordado en la característica anterior sobre la tensión entre lo privado y lo público. La tensión, como hemos visto hasta ahora,

³⁹ El concepto de calidad en relación con la producción de estadísticas generalmente se interpreta de manera particular, caracterizándose por una serie de atributos precisos y bien definidos, tales como coherencia, regularidad, y otros elementos que los organismos reguladores a nivel estatal o institucional dispongan.

⁴⁰ GOV.CO. Datos abiertos. <https://www.datos.gov.co/>

⁴¹ Para el caso de Colombia, los lineamientos de publicación de datos abiertos por parte de las entidades públicas y demás sujetos obligados están definidos en la Resolución 1519 de 2020 del Ministerio Tecnologías de la Información y Comunicaciones, así como en su Anexo 4 (Requisitos Mínimos de Datos Abiertos).

⁴² UNdata. <http://data.un.org/>

radica en que, por un lado, se promueve el libre acceso y la utilización de la información y, por el otro, la confidencialidad y la privacidad, al menos, en la práctica. Un ejemplo para ilustrar la problemática podría ser la desagregación de estudiantes víctimas del conflicto armado que se hace en la UNAL. Para la institución, resulta relevante conocer los datos de los estudiantes que pertenecen a dicho grupo, así como los programas curriculares a los que se inscriben. Sin embargo, compartir dicha información podría resultar en el señalamiento o la estigmatización de los diferentes programas curriculares, los departamentos universitarios, la universidad e, incluso, llegar a contextos nacionales o regionales a través de, por decir algo, la publicación de los datos mencionados en los portales del Estado, lo que podría poner en riesgo a las personas.

Vale la pena mencionar que a pesar de las tensiones que se puedan presentar entre las dimensiones pública y privada, es importante velar por que las poblaciones minoritarias se representen a través de cifras estadísticas.

Uso intensivo de las TIC

En la actualidad, los datos empleados para la consolidación de cifras estadísticas se capturan, almacenan, transforman, construyen, visualizan y comunican a través de herramientas relacionadas con las tecnologías de la información y la comunicación. Herramientas como Excel, Python o R son cada vez más comunes en el ámbito no especializado, y términos como analítica, *big data*, procesamiento de lenguaje natural, minería de datos o ciencia de datos son cada vez más empleados en nuestras conversaciones habituales.

De hecho, la relación entre estadísticas y tecnología no es un tema nuevo. Si se entiende a la tecnología como aquel conjunto de conocimientos, herramientas y habilidades aplicados de manera sistemática para la resolución de problemas, se puede afirmar que la producción de estadísticas ha estado siempre vinculada a la vanguardia tecnológica del presente, sin importar si esa vanguardia se refiere a recuentos en papel y lápiz, escuadras de celuloide, transportadores, compases, tinta china, tablas de madera, codificaciones digitales, encuestas virtuales, algoritmos o *scripts* de Python o R.

Pues bien, el avance tecnológico actual ha significado ventajas tales como procesar una mayor cantidad de datos en una menor cantidad de tiempo; difundir y comunicar los resultados a más personas; almacenamiento y procesamiento en la nube; simplificación de los procesos de búsquedas en bases de datos, entre otras. De hecho, pareciera que la tecnología de vanguardia se presenta como un imperativo de la época y hay quien pudiera pensar que el último programa de gestión y análisis de datos podría garantizar, en sí mismo, mejores resultados.

No obstante, el uso de las tecnologías más avanzadas no es sinónimo de una mejor producción de cifras estadísticas, ni de un mayor nivel de transparencia institucional ni tampoco lo es de una mejor comunicación. De modo que la consolidación de estadísticas en el ámbito de la gestión pública debe estar permeada por los avances más representativos en materia tecnológica, pero sin olvidar la importancia de la planeación estadística, de precisar los objetivos de cuantificación o medida, así como los elementos que van a ser susceptibles de dicha medición o conteo. Lo anterior en la medida en que la planeación estadística y las definiciones pertinentes son pasos previos y necesarios para la implementación de cualquier tecnología.

Finalmente, vale la pena mencionar que, en la actualidad, la implementación de nuevas herramientas tecnológicas supone temas que no se deberían pasar por alto, como la comercialización y el mercado de datos⁴³. De esta manera, los datos, además de ser de interés para las sociedades, son de interés para un amplio número de organizaciones que ofrecen diferentes servicios tecnológicos y comunicativos para apoyar los procesos de consolidación de cifras y que, en ocasiones, cobran por hacer uso de las herramientas que permiten procesar los datos e, incluso, pueden llegar a involucrar acuerdos para hacer uso de dichos datos con fines comerciales.

⁴³ El mercado de los datos involucra dos tipos de herramientas, a saber: las libres, como Python y R, y las comerciales, como Tableau o Power BI. En ocasiones, las entidades optan por modelos mixtos que emplean herramientas libres y comerciales para el análisis de datos.

Metadatos

La sexta característica o recomendación de las estadísticas en el ámbito de la gestión pública tiene que ver con el uso de metadatos, es decir, de los datos que hablan sobre los datos. Estos desempeñan un papel fundamental en la difusión y la comunicación de las estadísticas, pues definen la manera en la que los datos deben ser leídos, tanto para personal especializado como no especializado, reduciendo así malas interpretaciones o tergiversación. De hecho, según el DANE (2020b), los metadatos hacen referencia a la información necesaria para el uso e interpretación de las estadísticas y se requieren para articular y generar de manera adecuada la producción de información estadística.

En el presente libro, *Estadísticas e indicadores en la gestión pública*, se propone clasificar los metadatos en dos categorías: los metadatos *descriptivos* y los metadatos de *instrucción*. Los primeros describen las características de los datos, es decir, explican cómo son y cómo están construidos, así como los demás elementos necesarios para entenderlos. Los segundos definen los modos en que los datos pueden ser usados, por ejemplo, cómo se almacenan, cómo se modifican, sus restricciones de propiedad intelectual, entre otras instrucciones de uso⁴⁴.

Proceso estadístico

La séptima y última característica extrínseca determinante de las estadísticas en el ámbito de lo público tiene que ver con que estas se construyen y comunican a través de la ejecución de una serie de pasos y actividades que involucran, en mayor o menor medida, las características intrínsecas y extrínsecas descritas hasta ahora. Tal serie de procesos y subprocesos recibe el nombre de *proceso estadístico*. Según los *Lineamientos para el proceso estadístico en el Sistema Estadístico Nacional* del DANE (2020b) se entiende por proceso estadístico aquel conjunto sistemático de actividades encaminadas a la producción de estadísticas que comprende la detección

⁴⁴ Una descripción sobre la gestión de los metadatos y su lugar dentro del proceso estadístico se encuentra en el apartado “Gestión de metadatos” del DANE (2020b).

de necesidades de información, el diseño, la construcción, la recolección, el procesamiento, el análisis, la difusión y la evaluación.

La primera fase consiste en la *detección y análisis de necesidades*, donde se valida la necesidad de realizar la operación estadística⁴⁵ y la viabilidad de la misma; la segunda fase, de *diseño*, tiene que ver con la definición de la metodología de cada fase; la tercera es de *construcción*, donde se desarrollan las herramientas y los instrumentos necesarios para la operación estadística; la cuarta fase es de *recolección* y hace referencia a la ejecución de las fases anteriores con el fin de obtener los datos; la quinta fase comprende el *procesamiento* que resulta en una única base de datos con ponderaciones y resultados preliminares; la sexta es la de *análisis*, es decir, donde se verifica la coherencia y la consistencia de los datos para su posterior interpretación; la fase siete abarca el proceso de *difusión* que consiste en poner a disposición de los usuarios la información estadística generada; finalmente, la fase ocho alude a la *evaluación*, donde se verifica el cumplimiento de los objetivos, se consolidan los hallazgos y se formula un plan de mejora.

DE LA TEORÍA A LA PRÁCTICA

Con el fin de ilustrar la descripción teórica que se ha expuesto hasta ahora de las características de las cifras estadísticas como objeto de estudio, se presentan a continuación las características intrínsecas y extrínsecas a través de un ejemplo práctico basado en las cifras contenidas en uno de los principales recursos de la página de estadísticas de la UNAL⁴⁶, esto es, el módulo de cifras oficiales⁴⁷.

-
- ⁴⁵ La diferencia entre operación estadística y proceso estadístico no es fácil de describir y no está documentada. Sin embargo, tales términos se emplean, en ocasiones, con significados distintos. A lo largo de este libro, el proceso estadístico se entiende como una serie de pasos que debe cumplir cualquier operación estadística. Operación estadística, por su parte, es la aplicación regular del proceso estadístico a cada estadística que un Estado o entidad construye.
- ⁴⁶ UNAL en un vistazo. <http://estadisticas.unal.edu.co/home/>
- ⁴⁷ El módulo de cifras oficiales constituye solamente uno de los elementos dentro de dicha página. A lo largo de esta sección del presente libro, cuando se mencione la página de estadísticas oficiales de la UNAL, se estará haciendo referencia específicamente al módulo de cifras oficiales.

Características intrínsecas en la práctica

Para ilustrar las características intrínsecas, se considera la estadística o población de estudiantes de primera matrícula en la UNAL. Esta población abarca el número total de estudiantes que se inscriben por primera vez en la institución, como se refleja en la cifra de 5675 para el segundo semestre del año 2022⁴⁸. Es importante destacar que esta *población* constituye un *subconjunto anidado* dentro del conjunto más amplio de estudiantes matriculados en la universidad durante más de un periodo.

En cuanto a las *variables*, o características relevantes para el ejercicio estadístico dentro de la UNAL, en esta población se pueden identificar algunas como el sexo de los estudiantes, la sede de la universidad en la que estudian, el nivel de formación (pregrado, especialización, maestría, doctorado) y la edad. Las variables en cuestión se pueden dividir en *nominales* (el sexo y la sede), *ordinales* (nivel de formación) y de *intervalo* (edad)⁴⁹.

Las variables que describen la población empleada para exemplificar las características intrínsecas se transforman para que sus *datos* presenten una estructura ordenada (*tidy data*), que se consolida como una estrategia a nivel primario que facilita el almacenamiento y, en general, el trabajo con los microdatos. Sin embargo, la estructura *tidy data* no es la mejor para la comunicación y la visualización, por lo que, en la página de estadísticas oficiales de la UNAL, los datos se *agregan* a través de estrategias de visualización tabular que no necesariamente siguen la estructura *tidy data*.

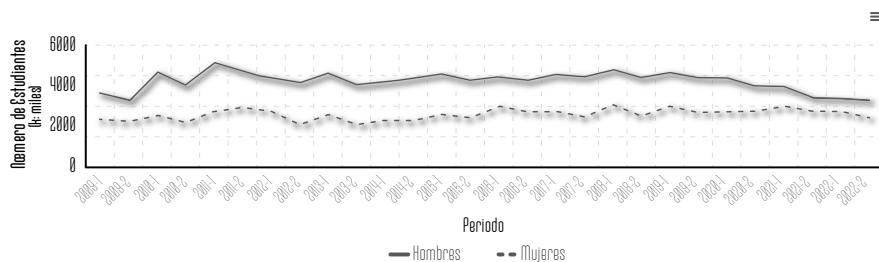
A partir de los datos consolidados y agregados, es posible *desagregar* la población en grupos más pequeños, por ejemplo, la *desagregación geográfica* de lugar de procedencia de los estudiantes; la *desagregación geográfica institucional* de los estudiantes matriculados por primera vez a nivel general, por sedes, facultades o programas académicos; y las *desagregaciones temáticas* de modalidad de formación, sexo, áreas del conocimiento, entre otras.

⁴⁸ En la página de las estadísticas se cuenta con la serie de tiempo completa desde el año 2009.

⁴⁹ Para efectos prácticos, en la página de estadísticas oficiales de la Universidad Nacional de Colombia esta variable de intervalo se vuelve ordinal por un ejercicio de transformación a través de rangos.

Las cifras, sin importar si son agregadas o desagregadas, se representan de manera tabular y gráfica. Por ejemplo, en la página de estadísticas de la UNAL se emplean algunas estrategias, como la *representación gráfica* que describe una serie de tiempo de la variable sexo y que evidencia, además, una desagregación temporal de esa variable temática (Gráfica 1); la *gráfica circular* que da cuenta de la situación actual (2022-2) de la distribución por sexo de los estudiantes matriculados por primera vez (Gráfica 2); una *representación tabular* que resume el comportamiento de la variable sexo a nivel temporal y actual (Tabla 3); y, para terminar, la representación de la *desagregación geográfica* del lugar de nacimiento (municipios de Colombia) de los matriculados por primera vez en 2022-2 a través de un mapa (Gráfica 3).

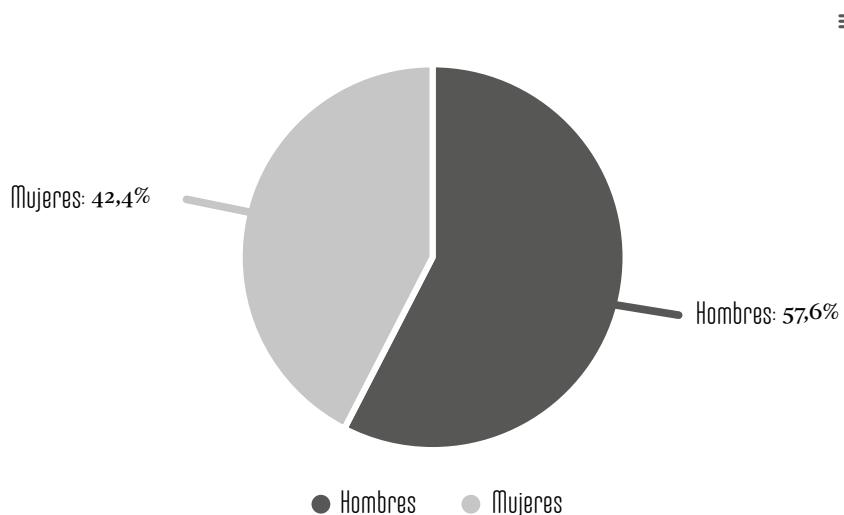
Gráfica 1. Evolución del número de estudiantes matriculados por primera vez de acuerdo con el sexo



Fuente: elaboración propia.

Finalmente, para concluir con las características intrínsecas, la página de estadísticas de la UNAL se consolida como un *medio de difusión y comunicación* que ofrece una amplia gama de cifras relevantes sobre diversas poblaciones de interés universitario. Mediante esta herramienta, es posible acceder de manera clara y concisa a las estadísticas oficiales de la universidad y, además, evidenciar cómo intervienen las nociones teóricas de las características de las estadísticas en la construcción de un ejercicio práctico.

Gráfica 2. Distribución de estudiantes matriculados por primera vez de acuerdo con el sexo



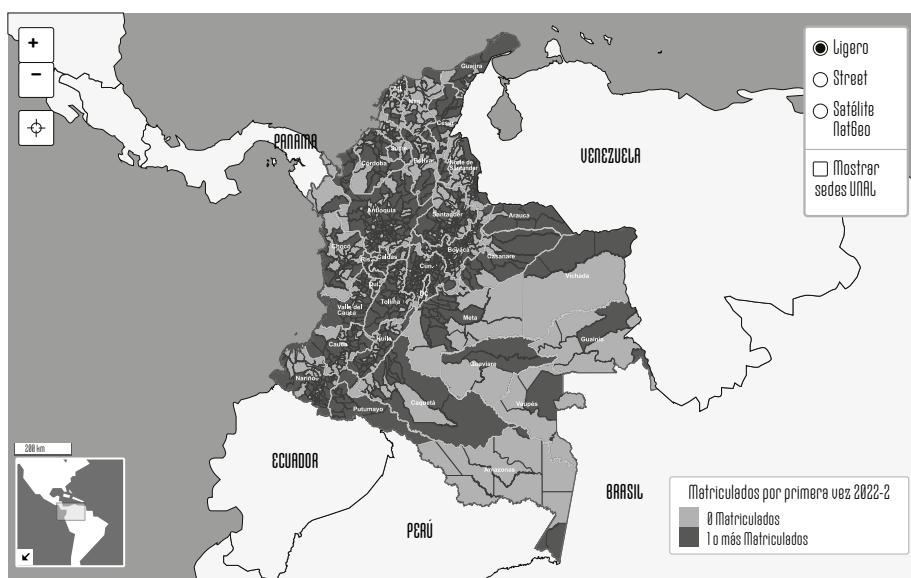
Fuente: elaboración propia.

Tabla 3. Total de estudiantes matriculados por sexo

Año	Periodo	Total estudiantes matriculados por primera vez por sexo			Total
		Hombres		Mujeres	
All	All	All		All	All
2022	2	3269		2406	5675
2022	I	3383		2741	6124
2021	2	3421		2760	6181
2021	I	3964		2998	6962
2020	2	4017		2734	6751

Fuente: elaboración propia.

Gráfica 3. Mapa del lugar de nacimiento en municipios de Colombia de los matriculados por primera vez



Fuente: elaboración propia.

Características extrínsecas en la práctica

Con respecto a las características extrínsecas, la disposición de las cifras que se hace en la UNAL pretende ser *inclusivo*, por lo que se vela por que todas las poblaciones que involucran personas sean descritas. Así, dentro de la población de estudiantes matriculados por primera vez, como en otras poblaciones universitarias, se describen también las poblaciones anidadas del Programa de Admisión Especial (PAES) y el Programa de Admisión Especial y Movilidad Académica (Peama)⁵⁰, que

⁵⁰ La modalidad de admisión Peama busca proyectar a la UNAL en el territorio nacional y, en específico, en departamentos y municipios con baja o nula cobertura de programas de educación superior de calidad, de difícil acceso o con problemas de orden público. La modalidad PAES, por su parte, está dirigida a la asignación regular de cupos en los programas de pregrado, y la admisión y disposición

representan, por ejemplo, el 9,4 y 6,1% respectivamente de todos los estudiantes admitidos por primera vez para 2022-02.

Además, dentro de la página de cifras oficiales de la UNAL la característica de la *comparabilidad* se refleja a través de la aplicación de normas, la unificación, la especificación y la simplificación. De esta manera, se tienen en cuenta las normas y las unificaciones, por ejemplo, a través de los códigos contenidos en estándares nacionales e internacionales que permiten identificar y definir lo que es un programa curricular y cada uno de ellos, así como los códigos de los municipios para los lugares de nacimiento⁵¹.

Un lugar de la página de estadísticas en el que se puede evidenciar el uso de un lenguaje conciso y de definiciones precisas es en la sección de *metadatos*, donde se describen los datos y la manera en la que estos deben ser leídos o empleados. A continuación, se presenta una tabla en la que se ven los modos en que, a través de metadatos, se precisan las definiciones de algunos atributos relevantes para la presentación de las cifras de la UNAL en la página de estadísticas oficiales (Tabla 4).

Los metadatos desempeñan un papel fundamental al proporcionar los detalles para comprender conjuntos de datos, siendo de utilidad tanto para un público especializado como para uno sin experiencia en el campo o no especializado. Esta consideración adquiere una importancia crucial, ya que garantiza el acceso sin restricciones a los datos, especialmente cuando se trata de asuntos públicos, es de gran relevancia. Desde esta perspectiva, la característica de las cifras estadísticas que resalta su naturaleza *pública, transparente y de acceso abierto* adquiere un valor aún más significativo. Pues bien, en la página de la UNAL hay una sección particular que contiene un hipervínculo a la página de datos abiertos del Estado colombiano⁵², a fin de facilitar el acceso y la disponibilidad de los microdatos oficiales de la

de apoyos especiales durante su proceso de formación para estudiantes provenientes de comunidades indígenas, de colegios ubicados en municipios en estado de extrema pobreza, de bachilleres de los mejores colegios del país, de integrantes de comunidades afrocolombianas y de víctimas del conflicto armado interno.

⁵¹ En Colombia, el estándar Divipola permite identificar los departamentos y municipios, entre otros aspectos, a través de los cuales se conforma política y administrativamente el territorio colombiano.

⁵² GOV.CO. Datos abiertos. <https://www.datos.gov.co/>

Tabla 4. Metadatos de atributos y desagregaciones

Atributos generales de la estadística	Desagregaciones de la estadística
Atributo	Definición
Nombre	Matriculados primera vez en la Universidad Nacional de Colombia.
Definición	Hace referencia a la población de estudiantes matriculados por primera vez en la Universidad Nacional de Colombia.
Tipo de medida	Estadística oficial
Orientaciones conceptuales y metodológicas	La construcción de esta estadística se orienta conceptual y metodológicamente por las orientaciones definidas en el libro “Gestión de la información cuantitativa en las universidades” y, en específico, en lo definido en el Apartado 4.3 del Capítulo 4. estadisticaun.github.io/L_Conceptual/ estadisticaun.github.io/L_Conceptual/4-3-once-caracteristicas-asociadas-a-las-estadisticas.html
Proceso estadístico	
Serie de tiempo disponible	Desde el año 2009
Frecuencia de medición	Semestral

Atributos generales de la estadística

Desagregaciones de la estadística

Desagregación	Descripción	Opciones de respuesta	Estándares, nomenclaturas o normatividad	Observaciones
Evolución matriculados primera vez	Hace referencia a la evolución histórica (serie de tiempo) del total de estudiantes matriculados por primera vez en la Universidad Nacional de Colombia			1. Se cuenta con información histórica del total de estudiantes matriculados por primera vez desde el primer periodo del año 2009. Aquellos interesados en obtener cifras históricas anteriores a este año, estas se encuentran inmersas en las distintas revistas de estadísticas e indicadores publicadas por esta institución de manera periódica desde el año 1965 y disponibles en el componente Publicaciones del sitio web http://estadisticas.unal.edu.co/
Modalidad de formación	Hace referencia a la modalidad de formación en la cual se encuentra el estudiante matriculado por primera vez en la Universidad	Pregrado postgrado		

Fuente: elaboración propia.

UNAL de manera pública, abierta, accesible y transparente para todos los ciudadanos.

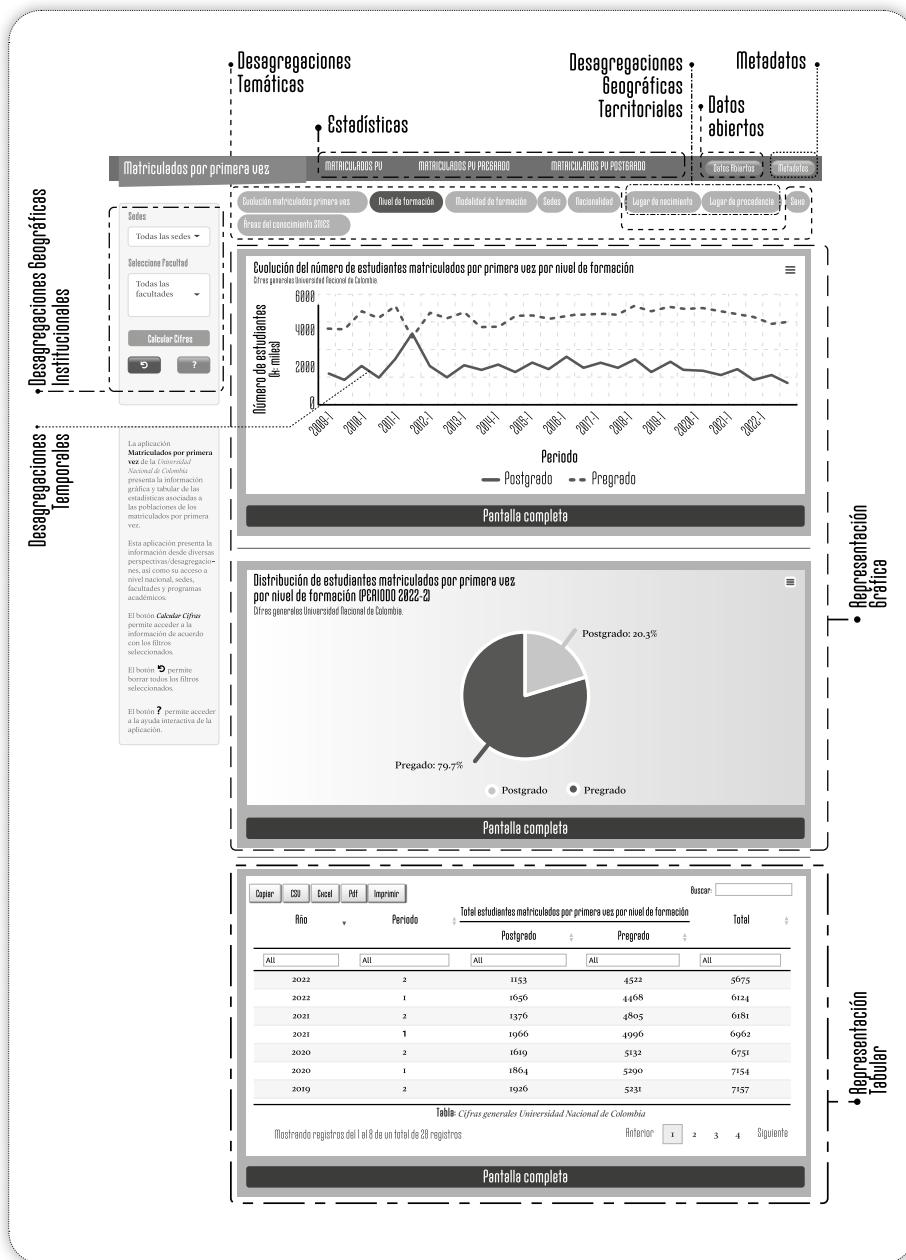
Otra característica extrínseca es la de *uso intensivo de las tecnologías de la información y comunicación (TIC)*, que se ve de manera clara en la página de estadísticas de la UNAL, en donde la influencia de las TIC condiciona la forma en que se recopila, procesa, almacena y presenta la información estadística. Lo anterior en la medida en que se emplean diversas estrategias de automatización, recopilación y comunicación de los datos para representar de manera eficiente y oportuna la información cuantitativa de la UNAL.

Al mismo tiempo, dentro de la estructuración de la página de estadísticas oficiales de la UNAL se aplican rigurosamente diversas normativas legales. Por ejemplo, se considera la Ley de Protección de Datos Personales en relación con la información recopilada de los estudiantes. Asimismo, se atienden los aspectos relativos a los derechos de autor y propiedad intelectual, garantizando la protección de los creadores de la página, las aplicaciones y otros elementos. También se aborda la normativa sobre *cookies* y tecnologías de seguimiento. Además de estas consideraciones legales, en términos éticos, el contenido de la página se esfuerza por ser imparcial y se ajusta a los lineamientos del Sistema de Ética Estadística del DANE (2021) y sus principios rectores.

Finalmente, la página de estadísticas de la UNAL representa la noción de *proceso estadístico* en la medida en que para llegar a ella se recopilan, analizan, interpretan y presentan datos numéricos de manera sistemática y organizada, siguiendo un flujo lógico de pasos que se definen para responder a unos objetivos definidos.

A continuación, se resaltan algunas de las características en elementos puntuales de la página:

Gráfica 4. Módulo de cifras en la página de estadísticas de la Universidad Nacional de Colombia



Fuente: elaboración propia.

CAPÍTULO III

Los indicadores

Con el fin de brindar un contexto adecuado para el contenido que se presentará en este capítulo, es necesario hacer una breve recapitulación de los capítulos anteriores. En primer lugar, en el capítulo I se exploró el concepto de información, el cual demostró ser un término polisémico con múltiples interpretaciones. Se hizo especial énfasis en la noción de “información como cosa”, es decir, como un conjunto de datos que puede ser manipulado, analizado y procesado. Se resaltó la estrecha relación entre la información cuantitativa y la gestión pública, evidenciando cómo los datos desempeñan un papel fundamental al respaldar la implementación de estrategias y políticas de interés. Por otro lado, en el capítulo II se examinó el concepto de estadísticas, que fueron definidas como un conjunto estructurado de información derivada de cifras descriptivas. Se subrayó, además, la existencia de condiciones intrínsecas que son necesarias y constitutivas de las estadísticas, así como de condiciones extrínsecas que indican la dirección que deben tomar en el ámbito de lo público.

En este capítulo se aborda el concepto de indicador como una herramienta de medición empleada en la evaluación y el monitoreo de fenómenos de interés en el ámbito de lo público, asimismo, se enuncian las funciones de los indicadores en dicho contexto. Además, se explora la relación entre las cifras estadísticas y los indicadores, así como sus usos en la práctica.

El capítulo continúa con la presentación de una tipología de indicadores que busca abarcar los distintos tipos existentes y resaltar la importancia de comprender cada categoría y utilizarlas de manera conjunta para lograr buenos procesos de gestión en lo público. Asimismo, se examinan las características y los requisitos que deben considerarse en la construcción de indicadores, como el enfoque SMART.

Por último, se destaca la naturaleza matemática subyacente en la conformación de indicadores, así como las diversas estrategias, como medidas estadísticas, frecuencias, razones, tasas e índices, que se emplean para su creación. Estas estrategias proporcionan un marco sólido para desarrollar indicadores significativos y representativos de los fenómenos que se pretenden medir.

DEFINICIÓN

El trabajo con indicadores ha ganado bastante popularidad en diversos ámbitos, sobre todo, en el mundo del *marketing*. Así, con términos como indicadores clave de rendimiento (KPI) o indicadores clave de riesgo (KRI), y con métricas alusivas al alcance de una marca, el número de clientes potenciales, las ventas realizadas por períodos de tiempo, entre otras, los profesionales del *marketing* pueden entender mejor al público, analizar el impacto de sus estrategias y hacer ajustes basados en la información resultante.

Así, los indicadores han adquirido una connotación privada, empresarial, con una clara orientación a la eficiencia en términos de maximizar el rendimiento financiero. Esto supone una serie de problemáticas como, por ejemplo, la tendencia a enfocarse en resultados a corto plazo en detrimento de los objetivos a largo plazo; la poca relevancia atribuida a factores externos, como la sostenibilidad; la falta de equilibrio, al reducir presupuesto y tiempo para, por ejemplo, la investigación o la innovación; y, en general, una consideración no tan exhaustiva de elementos de carácter ético.

Sin embargo, los indicadores no siempre están relacionados con el *marketing*, las empresas, el crédito y, en general, el ámbito privado. Como se muestra a continuación, los indicadores pueden ser entendidos desde otra perspectiva, fundamental para lo público.

La palabra “indicador” viene del latín *indicare* y hace referencia a algo, como una pista, una señal, un *índicio*, “que sirve para dar a entender”; para precisar, aclarar, para, justamente, *indicar*. Sus componentes léxicos son el prefijo *in-* (hacia dentro), la raíz *dicare-* (señalar con el dedo), y el sufijo *-dor* (agente, el que hace la acción). Esto se relaciona con la definición generalizada de los indicadores como una suerte de herramientas

de *medición* empleadas principalmente para la evaluación o el monitoreo, es decir, para señalar o mostrar determinadas características de algo, ya sea un proceso, producto o servicio.

Desde la perspectiva que nos ocupa, a saber, la de las cifras estadísticas en el marco de la gestión pública, los indicadores se consolidan como herramientas empleadas para la medición y el monitoreo de los diferentes aspectos del devenir social. En términos de Gutiérrez (2009), una de las funciones de los indicadores tiene que ver con “recabar y sistematizar información para el análisis estadístico sobre problemas y fenómenos sociales, donde se traducen los conceptos y nociones teóricas en unidades de análisis, como unidades de medida para cuantificar, de manera precisa, la realidad” (p. 17).

Bajo esta perspectiva, no es extraño pensar que el origen de los indicadores está relacionado con la vida en sociedad y con la idea moderna de gestión pública. De hecho, la aparición de indicadores se remonta a prácticas sociales como la implementación de balanzas o varas para medir, registros escritos o medidores artesanales y demás estrategias de las que se valían las sociedades antiguas para controlar y monitorear diferentes aspectos de su realidad social, como la producción agrícola o la distribución de alimentos a la población. Luego, con el paso del tiempo, y con la aparición de la estadística y los sistemas de información, los indicadores se convirtieron en herramientas clave para la medición y el monitoreo de fenómenos más consolidados, como la economía, la salud o el bienestar social.

Ahora bien, el asunto de los indicadores, tanto en las sociedades antiguas como en las sociedades actuales, no se limita a la mera definición y explicación de procesos, productos o servicios relevantes para la vida en sociedad. A través de las definiciones y explicaciones propias del trabajo con indicadores, antes y ahora, las sociedades y sus dirigentes fundamentan, entre otras, la toma de decisiones. Podría decirse, pues, a modo de conclusión, que, a pesar de las transformaciones a nivel conceptual y tecnológico, la importancia y la pertinencia de los indicadores para la comprensión y la mejora de la sociedad ha perdurado desde las comunidades antiguas hasta hoy.

En la actualidad, el trabajo con indicadores en el ámbito de lo público está ligado a la figura del Estado y sus instituciones, pues tienen la responsabilidad de definir indicadores para recopilar y analizar la

información sobre, primero, grandes objetivos estatales e institucionales relevantes para la idea de desarrollo social o institucional, como la salud, la economía, el ambiente, la educación, entre otros. Segundo, los procesos de gestión interna que se dan en la práctica dentro de las instituciones para perseguir tales objetivos. Y tercero, la puesta en práctica de planes, programas o proyectos que buscan intervenir en aspectos específicos y precisos de esos grandes objetivos.

FUNCIONES DE LOS INDICADORES

Como se destacó, la definición y el uso de indicadores es clave para la gestión de los Estados y sus instituciones. A continuación, desde la experiencia de la Dirección Nacional de Planeación y Estadística de la Universidad Nacional, se presentan las funciones más comunes de los indicadores en el ámbito público.

- **Medición y evaluación:** los indicadores se utilizan para medir y evaluar un determinado fenómeno, como puede ser el desempeño de una empresa o el impacto de una política pública.
- **Monitoreo:** los indicadores permiten monitorear, es decir, hacer un seguimiento del estado de un proceso a través del tiempo, usualmente, en términos de desempeño o rendimiento.
- **Información:** los indicadores sirven para consolidar y dar a conocer información relevante e imparcial.
- **Comunicación:** los indicadores pueden ser una herramienta que permita consolidar información de interés de manera clara y concisa para procesos efectivos de comunicación, por ejemplo, a través de gráficos o tablas.
- **Focalización:** a través de los indicadores, los procesos de medición y descripción de fenómenos relevantes para los Estados y sus instituciones se pueden dirigir específicamente a los elementos necesarios.
- **Transparencia y rendición de cuentas:** los indicadores sirven para que los estados y sus instituciones presenten información de acceso público a fin de que se conozcan los modos en los que se están dando sus procesos de gestión. Asimismo, a través de los

- indicadores las entidades de control pueden hacer seguimiento de los procesos.
- **Toma de decisiones:** una de las funciones más populares de la consolidación de indicadores es que, a través de las funciones ya mencionadas, se pueden establecer elementos que contribuyen a orientar la planeación de estrategias y la toma de decisiones.

ESTADÍSTICAS E INDICADORES

Llegados a este punto es importante abordar la distinción entre estadísticas e indicadores. Para empezar, las estadísticas y los indicadores se encuentran en estrecha relación. Se emplean para recopilar, presentar y analizar información que permita describir y explicar fenómenos de interés, a través de, sobre todo, datos numéricos. Sin embargo, aunque son complementarios, no son iguales.

La diferencia radica en que las estadísticas son el resultado de *conteos* u operaciones aritméticas básicas, que buscan describir de manera general poblaciones o muestras relacionadas con un fenómeno de interés⁵³. Los indicadores, por su parte, suelen emplear *medidas*, como frecuencias, proporciones, razones, tasas, índices, entre otras, para resumir y comunicar información focalizada sobre un aspecto específico del fenómeno de interés.

Con frecuencia, los indicadores se basan en las estadísticas, así, se seleccionan cifras estadísticas relevantes y se combinan para formar un indicador que pretende ser claro, preciso y específico. Por ejemplo, el índice de precios al consumidor (IPC) es un indicador económico *específico* que se basa en las cifras estadísticas *generales* sobre un conjunto representativo de los precios de los bienes y servicios finales en los que gastan los hogares. Se podría decir, pues, que de la misma manera en que los datos se consolidan, a nivel descriptivo, como cifras estadísticas en potencia, las cifras estadísticas pueden ser indicadores en potencia.

⁵³ Un elemento que caracteriza a las estadísticas en el ámbito institucional es que se producen cumpliendo estándares, metodologías y un proceso estadístico que garantiza que la información generada cumpla con determinados atributos de calidad. Sin embargo, es posible hablar de estadísticas por fuera de las reglas de la institucionalidad.

En definitiva, las estadísticas son un mecanismo de recopilación y análisis general de datos descriptivos sobre poblaciones o muestras relevantes a nivel social o institucional, que se ordenan a través de conteos y operaciones aritméticas básicas. Mientras que los indicadores son medidas complejas empleadas para la focalización y la profundización en los fenómenos de interés.

Ahora bien, en el ámbito de la gestión pública, tanto estadísticas como indicadores se emplean en conjunto para caracterizar aspectos de interés, monitorear y evaluar programas, diseñar políticas, asignar recursos, garantizar la transparencia y la rendición de cuentas, tomar decisiones informadas, entre otras.

TIPOLOGÍAS DE INDICADORES

Las diferentes funciones de los indicadores, así como la característica polisémica del término *indicador* y el uso extendido en las dinámicas económicas, administrativas e investigativas de la actualidad, redundan en la consolidación de múltiples tipologías para la clasificación de indicadores.

Cada tipología tiene sus propios criterios de clasificación, como la fuente de los datos, el alcance, la naturaleza y el propósito de la medición. Por ejemplo, una tipología puede clasificar a los indicadores según si son cuantitativos o cualitativos, y otra puede clasificarlos según su relevancia para el sector privado o el sector público.

Por ejemplo, en el escenario de los indicadores cuantitativos y cualitativos, los primeros son fáciles⁵⁴ de medir y comparar, pero pueden no proporcionar información completa sobre la complejidad de los problemas o las experiencias de los individuos que se van a medir. Los segundos pueden proporcionar información más rica y detallada sobre los contextos y las experiencias, pero son más difíciles de medir y comparar.

⁵⁴ Los indicadores cuantitativos se consideran relativamente fáciles de medir debido a su naturaleza numérica. No obstante, en la práctica, medir cuantitativamente puede presentar desafíos como la definición precisa de poblaciones y variables, la selección de metodologías adecuadas, entre otros. Al final, para medir cuantitativamente hay que seguir un enfoque riguroso que garantice la validez y confiabilidad de los resultados.

A pesar de estas limitaciones, todas las tipologías de indicadores son útiles para diferentes propósitos. En la práctica, lo más importante es seleccionar la tipología correcta con los indicadores adecuados para las necesidades y los objetivos específicos de cada contexto.

Ahora bien, hay una serie de dificultades propias de la variedad de tipologías existentes para clasificar indicadores, como la confusión y la falta de estandarización. Si no se establece una clasificación clara y comúnmente aceptada para los indicadores, puede ser difícil entender y comparar diferentes conjuntos de datos. Además, si hay demasiadas tipologías, puede resultar difícil para los usuarios elegir la que mejor se adapte a sus necesidades. Esto puede llevar a la selección de una tipología inapropiada, a la combinación de diferentes tipologías o a la construcción de indicadores adecuados, pero empleados en categorías inadecuadas, lo que puede dificultar la interpretación de los resultados o generar indicadores solapados, es decir, que digan lo mismo.

Antes de abordar en detalle cada una de las tipologías de indicadores presentadas en este libro, es necesario señalar que se ofrece solo una aproximación inicial a la clasificación y definición de las principales tipologías que se consideran relevantes en el ámbito público. Es importante tener en cuenta que este trabajo no tiene la pretensión de agotar todas las posibilidades de uso de diferentes tipologías de indicadores y que se reconoce la existencia de otras perspectivas y enfoques que podrían ampliar el panorama y enriquecer la comprensión de la medición en el contexto público.

Indicadores cualitativos y cuantitativos

Los indicadores cuantitativos se basan en mediciones numéricas y se utilizan para medir cantidades y magnitudes. Estos indicadores se expresan generalmente en términos de números, porcentajes, tasas, entre otras. Algunos ejemplos de indicadores cuantitativos incluyen el producto interno bruto (PIB), la tasa de inflación, la tasa de mortalidad, la tasa de desempleo o la cantidad de emisiones de gases de efecto invernadero.

Por otro lado, los indicadores cualitativos se basan en observaciones subjetivas y se utilizan para medir las características no numéricas de un fenómeno, como la calidad, la percepción, las actitudes, los valores, entre otros aspectos no numéricos. Tales indicadores se expresan generalmente en términos de descripciones verbales o visuales, como palabras o imágenes.

Sin embargo, en el ámbito de los indicadores —si no en todos los ámbitos de la medición en general— las características no numéricas se suelen entender, configurar y medir a partir de escalas de orden y clasificación de elementos en categorías o niveles que involucran magnitudes o números. Por ejemplo, para medir un indicador como la satisfacción, que está relacionado con las percepciones y, por tanto, con lo cualitativo, se terminan empleando escalas psicométricas⁵⁵ que ordenan las percepciones en segmentos simétricos que se analizan y se procesan numéricamente, como las *escalas Likert*.

Si las cualidades o los constructos se miden numéricamente, podría resultar más conveniente hablar de estrategias de medición directas e indirectas, en lugar de cualitativas y cuantitativas. El peso o la estatura de una persona serían medidas directas, mientras que su nivel de felicidad o de ansiedad serían indirectas⁵⁶. Ahora bien, los grupos focales, las entrevistas o la etnografía podrían ser, en sí mismas y sin un procesamiento numérico, ejemplos más propios de la medición en el mundo de lo cualitativo, sin embargo, no es fácil determinar hasta qué punto podrían ser indicadores⁵⁷.

-
- §§ Para que estas escalas puedan ser utilizadas de manera efectiva, es fundamental que estén respaldadas por medidas de fiabilidad, como la validez de constructo y la confiabilidad. La validez de constructo se refiere a la capacidad de la escala para medir el constructo o la variable de interés de manera precisa y coherente. Por otro lado, la confiabilidad se relaciona con la consistencia y estabilidad de los resultados obtenidos a través de la escala en diferentes momentos y contextos.
 - §§ En el ámbito de las estrategias de medición indirectas es importante preguntarse cuánto se pierde en el proceso del fenómeno que se mide.
 - §§ Valdría la pena explorar la noción de medición en el marco de lo cualitativo; pensar si, por ejemplo, se puede medir sin números y, así, una descripción etnográfica, un producto literario o una pintura pueden ser entendidas como herramientas puramente cualitativas de medición. La reflexión al respecto de las mediciones en el ámbito de lo cualitativo y cuantitativo puede ser de gran utilidad para el fenómeno actual del big data y de los datos no estructurados, que conforman la mayor parte de datos disponibles en la actualidad y cuyo análisis puede contribuir en gran medida a la consolidación de nuevas estadísticas y, con ello, a la toma de decisiones, la transparencia, el seguimiento, la evaluación de políticas públicas, entre otras. No obstante, hasta ahora, la mayoría de los esfuerzos de procesamiento, sistematización y análisis de datos no estructurados se centra en estrategias cuantitativas.

En la práctica, muchos indicadores combinan elementos de medición directa e indirecta para proporcionar una imagen más completa y precisa de un fenómeno determinado. Por ejemplo, volviendo sobre el indicador de satisfacción, se podrían incluir para su medición encuestas cuantitativas y preguntas cualitativas abiertas.

Indicadores temáticos

Los indicadores también se pueden clasificar a partir de la problemática o área de interés sobre la que se concentran. Así, se generan indicadores específicos que se focalizan en diferentes dimensiones, como la educativa, la social, la política, la económica, la demográfica, entre otras. Dichos indicadores permiten, además de la focalización, la comparación y el contraste con respecto a situaciones puntuales entre diferentes países, regiones o instituciones.

Los indicadores temáticos no son una tipología menor. Al contrario, estos indicadores pueden considerarse como indicadores primarios, que se construyen a partir de cifras estadísticas y que buscan dar cuenta de grandes acuerdos sociales considerados relevantes para el desarrollo humano, tales como la salud, la educación, la economía o la demografía. En este sentido, los indicadores temáticos funcionan como herramientas de carácter, sobre todo informativo, que señalan los síntomas generales del cuerpo social y deberían emplearse como línea base para establecer los tan populares indicadores de gestión, que son de carácter, especialmente, interventivo. Dada la naturaleza de los indicadores temáticos, estos se encuentran principalmente a nivel de organismos multilaterales o países, no obstante, son ampliamente utilizados en entidades públicas. Un buen ejemplo podría ser la implementación de un sistema de indicadores temáticos en una universidad que describa áreas de interés como formación, investigación, extensión universitaria, talento humano, bienestar universitario, entre otras.

Sin tener en cuenta los indicadores temáticos se pueden plantear procesos a nivel estatal o institucional —ya sean recurrentes o no recurrentes⁵⁸— que ignoren las cifras globales y la situación histórica y actual

⁵⁸ Como se resalta en el apartado de “Indicadores de gestión”, en el ámbito de la gestión, los procesos pueden ser de dos tipos: los recurrentes, que no están delimitados en el tiempo, y los no recurrentes, que están delimitados en el tiempo.

de los temas de interés. Dicho de otro modo, desconocer la noción de indicadores temáticos deriva en acciones estatales o institucionales que no tienen en cuenta los síntomas del cuerpo social o institucional, como un tratamiento médico que ignora la enfermedad.

En el contexto público colombiano es común que, aunque se reconozcan y definan los indicadores temáticos, no siempre se materialicen en procesos específicos dentro de las instituciones, ni en planes, programas o proyectos. Esto se debe, en parte, a que resulta difícil para un gobierno de turno, limitado en tiempo y recursos, solucionar de manera definitiva las problemáticas que involucran aspectos tan amplios como la salud o la economía. No obstante, es fundamental que los gobiernos establezcan estrategias que contribuyan a definir indicadores temáticos a partir de la consolidación de cifras estadísticas. Además, se deben diseñar planes, programas y proyectos que, aun cuando estén alineados con las convicciones y limitaciones de un gobierno de turno, respondan a la situación que se refleja a través de los indicadores temáticos y contribuyan, de manera articulada, a dar respuesta, al menos parcialmente, a esos grandes temas de interés social o institucional.

A continuación, algunos ejemplos de indicadores temáticos aplicables, principalmente, a nivel estatal.

1. *Indicadores económicos*: se utilizan para medir y evaluar la actividad económica de una sociedad⁵⁹. Por ejemplo, la inflación, el IPC o el PIB.
2. *Indicadores sociales*: miden y evalúan lo relacionado con el bienestar social. Por ejemplo, la calidad de vida, el acceso a servicios básicos o la tasa de alfabetización.
3. *Indicadores ambientales*: son usados para medir y evaluar el impacto de las actividades humanas en el medio ambiente. Por ejemplo, la huella de carbono o la calidad del aire.

⁵⁹ Para el uso y entendimiento de la medición de la dimensión económica de un país, por ejemplo, se puede acudir al libro de Eduardo Lora y Sergio I. Prada titulado Técnicas de medición económica.

4. *Indicadores demográficos*: se utilizan para medir y evaluar la dinámica poblacional, su composición y estructura. Por ejemplo, la tasa de natalidad, de mortalidad o de migración.

Indicadores de gestión

Como se mencionó en el primer capítulo de este libro, la gestión en el ámbito de lo público no es un tema menor. Al contrario, tiene que ver con la puesta en práctica de una estrategia que se ha convenido para atender a una problemática que les compete a todos los individuos que comparten un cierto lugar en el Estado y sus instituciones.

Desde una perspectiva práctica sobre cómo se aplica la gestión en lo público, se han definido varios modelos particulares, por ejemplo, el modelo de gobernanza, nueva gestión pública, burocrático, gobierno abierto, entre otros. Estos modelos conviven de manera no tan definida al interior de los Estados y sus instituciones. Sin embargo, sin importar el modelo —o los modelos— que se apliquen, los Estados en la actualidad, o por lo menos el colombiano, se han decantado porque la gestión se dé a través de una noción de procesos.

Pues bien, teniendo en cuenta que la ausencia de un modelo preciso y consensuado para la gestión en las entidades públicas puede causar distorsión en la construcción de indicadores de gestión, a continuación se exponen los indicadores de gestión a través de procesos.

Los procesos son secuencias de pasos o acciones definidas que se llevan a cabo de manera sistemática y coordinada para lograr un objetivo. Estos se pueden entender de dos maneras, los *recurrentes*, que no están acotados en el tiempo, y los *no recurrentes*, que están acotados en el tiempo. Los primeros hacen referencia a la medición y evaluación de los procesos internos de las organizaciones, los cuales permanecen a lo largo del tiempo sin importar las políticas o las convicciones de los gobiernos de turno y que, en últimas, mantienen el funcionamiento de las organizaciones y, más importante, su institucionalidad. Los segundos, por su parte, hacen referencia a la ejecución de planes, programas o proyectos del orden estatal o institucional; tienen un periodo de vida definido y están en estrecha relación con los períodos de los gobiernos de turno.

Un ejemplo de proceso recurrente es la dinámica interna de una universidad que se propone siempre graduar estudiantes; si ese proceso se deja de llevar a cabo, la universidad como institución deja de funcionar. Un proceso no recurrente podría ser el plan de una rectoría de una universidad para mejorar las habilidades de lectura y escritura de los estudiantes recién admitidos a carreras de humanidades. El plan en mención debería tener una fecha de inicio y final, y está supeditado a las prioridades y convicciones del gobierno de turno.

A diferencia de los indicadores temáticos, los indicadores de gestión necesitan que se establezcan *niveles de desempeño* para los procesos recurrentes y *metas* para los no recurrentes⁶⁰. Los niveles de desempeño miden, monitorean y evalúan el rendimiento o la calidad de manera continua, mientras que las metas establecen objetivos específicos que se deben alcanzar en un plazo determinado.

Llegados a este punto, es importante destacar que, aun teniendo muy buenos procesos de gestión, con sus correspondientes indicadores, esto podría no ser suficiente si esos procesos y sus indicadores no están alineados con las necesidades reales de la población a nivel social, económico, educativo, ambiental, entre otros. En este sentido, es fundamental considerar los indicadores temáticos, junto con las estadísticas oficiales, como una herramienta esencial que da cuenta de los grandes temas de la humanidad que permiten identificar las problemáticas que requieren ser atendidas y evaluadas de manera regular. De modo que es conveniente establecer una interacción entre los indicadores temáticos, de gestión y las estadísticas a fin de garantizar que los procesos y las decisiones de las instituciones estén alineados con los grandes temas de interés social.

Volviendo sobre los procesos, de acuerdo con el Departamento Nacional de Planeación (DNP) (2018), estos siguen una lógica de estructura de cadena de valor, es decir, en su interior se agrupan en varias tipologías que involucran *insumos, actividades, productos y resultados*⁶¹. A continuación,

-
- 60 En lugar de niveles de desempeño, los procesos recurrentes podrían tener metas. Sin embargo, esto no es usual.
 - 61 Entender las dinámicas no recurrentes a través de esquemas provenientes del ámbito privado, como la lógica de insumos, actividades y productos, suele ser objeto de críticas. Es importante plantearse si existen alternativas para comprender de manera más integral y profunda las dinámicas internas de las instituciones, sobre todo, cuando tienen un carácter humanístico, como la educación.

se presentan los indicadores para cada uno de los subtipos, que conforman un proceso.

Insumo

• *Indicadores de insumo o capacidad:* miden la capacidad de una organización para realizar sus actividades de manera recurrente o no recurrente. Por ejemplo, el número de trabajadores capacitados en una institución o los recursos financieros con los que cuenta un país o entidad para desarrollar un plan, un programa o un proyecto.

Actividades

Es relevante destacar que los indicadores relacionados con las actividades o el seguimiento y cumplimiento varían según si se aplican en procesos recurrentes o no recurrentes. De hecho, cada tipo de proceso tiene sus propios marcos normativos y metodológicos para su implementación. Por ejemplo, en el contexto de Colombia, la norma ISO 9000 es referente en el ámbito de los procesos recurrentes, mientras que los lineamientos del DNP (2018) son aplicados en los procesos no recurrentes⁶².

- *Indicadores de actividades:* se enfocan en el proceso *per se* y no en el resultado final. Permiten determinar el rendimiento de un proceso mientras se está llevando a cabo y, así, implementar diferentes estrategias durante su ejecución. Un indicador de proceso podría ser la cantidad de estudiantes que asisten a las sesiones del programa de capacitación en empleabilidad.
- *Indicadores de seguimiento o cumplimiento:* monitorean y miden el avance en el cumplimiento de las metas incluidas en un plan, programa o proyecto —es decir, en un proceso no recurrente— que

⁶² Es importante mencionar que tales normas y guías suelen estar pensadas desde y para el ámbito privado. Sin embargo, la gestión de lo público tiene unas características particulares, como mayor complejidad en la gestión o recursos limitados y, en general, interés por aspectos diferentes a la eficacia y el rendimiento, como podrían ser la equidad o la justicia social. De manera que la implementación de tales normas en el contexto de lo público debería involucrar un proceso de reflexión que redunde en su adaptación y contextualización. Por ejemplo, en Colombia se implementó la norma técnica NTC GP 1000, que buscaba establecer los lineamientos para la implementación de un sistema de gestión de calidad en entidades del sector público.

se ejecuta a lo largo de un periodo de tiempo que puede, o no, estar dividido en intervalos o periodos de tiempo menores. Por ejemplo, el cumplimiento de una meta propuesta en un grupo de pacientes que han mejorado su salud después de un mes de tratamiento médico en una población base intervenida. De acuerdo con el DNP (2018), estos indicadores se pueden clasificar a su vez en subtipos como *Stock*, Flujo, Acumulado, Capacidad, Reducción, entre otros.

Con el fin de exemplificar y entender⁶³, de manera general, los subtipos de indicadores de seguimiento o cumplimiento que se presentan a continuación y que permiten monitorear y evaluar el avance de las metas incluidas en planes, programas o proyectos a nivel estatal o institucional, se va a suponer que se dispone de un plan, programa o proyecto (proceso no recurrente) y que este se ejecutará en un intervalo de tiempo⁶⁴ *T*. Asimismo, se parte de que se hará seguimiento al cumplimiento de las metas incluidas en el plan, programa o proyecto de manera general y a través de subdivisiones del tiempo *T* en un número determinado de períodos⁶⁵ *P*. Las metas propuestas, por su parte, tienen unas líneas

-
- 63 La *Guía para la construcción y análisis de indicadores* (2018) del DNP de Colombia ofrece mayores detalles y ejemplos sobre los subtipos de indicadores de gestión para procesos no recurrentes que conforman el conjunto de indicadores de seguimiento o cumplimiento.
 - 64 El tiempo *T* para la ejecución de un plan, programa o proyecto en el contexto de lo público está altamente correlacionado con los períodos de tiempo de los gobiernos de turno, tanto a nivel estatal como institucional. Por ejemplo, en Colombia, los planes, programas o proyectos del Gobierno nacional se ejecutan en un lapso de cuatro años, que es el período presidencial vigente. En contraste, los planes, programas y proyectos de la UNAL, por ejemplo, se ejecutan durante un intervalo de tiempo de tres años, que equivale a los tiempos definidos estatutariamente en esta institución para ejercer las distintas rectorías.
 - 65 En lo público, el número de períodos *P* a lo largo de los cuales se ejecutará un plan, programa o proyecto usualmente hace referencia a años, que es la forma en que se definen y disponen los recursos financieros en las diferentes naciones y sus instituciones. Por ejemplo, en Colombia el número de períodos —también conocidos como vigencias— para el desarrollo de los planes, programas y proyectos a nivel nacional es de cuatro años, que equivale al tiempo constitucionalmente establecido para ejercer cada una de las presidencias del país.

base —es decir, de dónde se parte⁶⁶— y unas metas generales y por períodos —es decir, a dónde y cuándo se desea llegar— una vez se ha implementado una política pública incluida en un plan, programa o proyecto.

- *Stock*: los indicadores de *stock* miden los esfuerzos públicos por mantener un resultado constante, es decir, miden que el nivel alcanzado en determinado periodo P se mantenga. De modo que las líneas base deben ser iguales a las metas. La Ecuación 1 describe la representación formal de un indicador de *stock*.

ECUACIÓN 1. STOCK

$$\% \text{ de avance periodo}_t = \frac{\text{Avance}_t}{\text{Meta Periodo}_t}$$

$$\% \text{ de avance general} = \frac{\text{Avance}_P}{\text{Meta Final}}$$

$$\text{Metas Intermedias y Meta Final} = \text{Meta Periodo}_P$$

$$t = \{1, 2, 3, \dots, P\}$$

- Flujo: los indicadores de flujo miden el cumplimiento de las metas que, aunque están acotadas en el tiempo T de ejecución de un plan, programa o proyecto, se repiten periódicamente durante los períodos P que conforman los planes, programas o proyectos adelantados a nivel nacional o institucional. Los resultados sincrónicos, es decir, los de un periodo P , no afectan los resultados finales. A medida que avanza el tiempo de ejecución,

⁶⁶ Las líneas de base de un plan, programa o proyecto a nivel de los Estados y las instituciones deberían ser extraídas de las estadísticas oficiales o de los indicadores temáticos. Esto último garantizaría una alta correspondencia entre el sentir social, medido a través de las estadísticas e indicadores temáticos, y las apuestas de gestión de los gobiernos de turno. Lamentablemente, aún es muy común encontrar en los Estados y las instituciones planes, programas o proyectos cuyas metas no siempre se corresponden con los síntomas que reflejan las cifras oficiales disponibles.

y a diferencia de los indicadores de *stock*, en donde la meta se mantiene constante en los diferentes periodos de ejecución, en este subtipo de indicador la meta se incrementa periodo a periodo con el fin de que, al finalizar la ejecución del plan, programa o proyecto, se alcance el resultado final deseado. La Ecuación 2 describe la representación formal de un indicador de flujo.

ECUACIÓN 2. FLUJO

$$\% \text{ de avance periodo}_t = \frac{\text{Avance}_t}{\text{Meta Periodo}_t}$$

$$\% \text{ de avance general} = \frac{\text{Avance}_P}{\text{Meta Final}}$$

$$\text{Meta Final} = \text{Meta Periodo}_P$$

$$t = \{1, 2, 3, \dots, P\}$$

- Acumulado: los indicadores de acumulado miden el avance a través de datos de periodos de tiempo independientes que, aunque no se afectan entre sí, se suman para determinar el avance total. Dicho de otro modo, las mediciones de cada periodo de tiempo tienen una línea base de 0, y el avance final se calcula sumando los resultados obtenidos en cada periodo de medición. La Ecuación 3 describe la representación formal de un indicador de acumulado.

ECUACIÓN 3. ACUMULADO

$$\% \text{ de avance periodo}_t = \frac{\text{Avance}_t}{\text{Meta Periodo}_t}$$

$$\% \text{ de avance general} = \frac{\sum_{t=1}^P \text{Avance}_t}{\text{Meta Final}}$$

$$\text{Meta Final} = \sum_{t=1}^P \text{Meta Periodo}_t$$

$$t = \{1, 2, 3, \dots, P\}$$

- Capacidad: los indicadores de capacidad miden el avance acumulado —es decir, periodo a periodo— en el cumplimiento de las metas a partir de la disposición de una línea base y una meta final. La línea base es el punto de partida del plan, programa o proyecto y equivale al valor inicial existente en el proceso al momento de definir la meta final. Es decir, cuando aún no ha iniciado la ejecución del plan, programa o proyecto. La Ecuación 4 describe la representación formal de un indicador de capacidad.

ECUACIÓN 4. CAPACIDAD

$$\% \text{ de avance periodo}_t = \frac{\text{Avance}_t - \text{Línea Base}}{\text{Meta Periodo}_t - \text{Línea Base}}$$

$$\% \text{ de avance general} = \frac{\text{Avance}_P - \text{Línea Base}}{\text{Meta Final} - \text{Línea Base}}$$

$$\text{Meta Final} = \text{Meta Periodo}_P$$

$$t = \{1, 2, 3, \dots, P\}$$

- Reducción: los indicadores de reducción miden los esfuerzos para disminuir, a través de un plan, programa o proyecto, los factores que generan resultados negativos en el contexto de lo público. Se parte de una línea base conocida y se busca llegar a una meta definida previamente, la cual busca ser alcanzada en un periodo de tiempo determinado de antemano. Los resultados tienen que ver con la disminución lograda en el valor de la meta proyectada a lo largo del tiempo. La Ecuación 5 describe la representación formal de un indicador de reducción.

ECUACIÓN 5. REDUCCIÓN

$$\% \text{ de avance periodo}_t = \frac{\text{Línea Base} - \text{Avance}_t}{\text{Línea Base} - \text{Meta Periodo}_t}$$

$$\% \text{ de avance general} = \frac{\text{Línea Base} - \text{Avance}_P}{\text{Línea Base} - \text{Meta Final}}$$

$$\text{Meta Final} = \text{Meta Periodo}_P$$

$$t = \{1, 2, 3, \dots, P\}$$

Producto

• *Indicadores de producto:* se enfocan en la medición y evaluación de la cantidad y calidad de los bienes o servicios públicos resultantes de la transformación de los insumos a través de la ejecución tanto de los procesos recurrentes como de los no recurrentes. Según el DNP (2018), pueden ser de oferta o demanda. Los primeros se centran en medir la capacidad de lo público para suministrar bienes y servicios; los segundos, por su parte, buscan dar cuenta de los miembros de la comunidad que resultan beneficiados de los productos producidos a nivel estatal e institucional.

Resultado

• *Indicadores de resultado:* tienen como propósito medir y evaluar los resultados parciales de los procesos recurrentes o finales del plan, programa o proyecto a través de los productos inmediatos resultantes del proceso. Constituyen un elemento fundamental para determinar si los objetivos de los procesos recurrentes y no recurrentes se alcanzan. Por ejemplo, un indicador de resultado de procesos recurrentes es la cantidad de estudiantes graduados y uno de procesos no recurrentes puede ser el número de estudiantes recién graduados que han conseguido vincularse laboralmente después de una capacitación.

• *Indicadores de impacto:* a diferencia de los indicadores de resultado, los de impacto miden las consecuencias de la implementación de un plan, programa o proyecto a mediano y largo plazo. Se debe tener en cuenta que los indicadores de impacto constituyen, apenas, una estrategia de aproximación a la medición de los efectos de la implementación, sin embargo, el impacto propiamente dicho de un plan, programa o proyecto debe medirse a través de un proceso más amplio, sistemático y riguroso, a saber, la evaluación de impacto⁶⁷.

⁶⁷ El Departamento Nacional de Planeación (DNP) de Colombia, desde hace más de dos décadas y como se evidencia en el repositorio de evaluaciones (<https://sinergiapp.dnp.gov.co/#Evaluaciones/EvalFin>), ha adelantado más de 300 evaluaciones a diversas políticas públicas implementadas en el país.

Indicadores según eficiencia, eficacia y efectividad⁶⁸

Los indicadores de gestión en el ámbito de lo público se pueden clasificar de acuerdo con la eficiencia, la eficacia y la efectividad. Los indicadores de eficacia miden qué tanto se están cumpliendo las funciones u obligaciones recurrentes de un Estado o entidad pública, así como las metas contenidas en los planes, programas y proyectos asociados a estas. Por otro lado, los indicadores de eficiencia permiten medir la capacidad que tiene un país y sus instituciones públicas para cumplir con sus obligaciones recurrentes, así como sus metas, haciendo uso de la menor cantidad de insumos o recursos disponibles (indicadores de insumo). Finalmente, los indicadores de efectividad miden la relación o el equilibrio entre la eficacia y la eficiencia alcanzada por los Estados y las instituciones públicas en el ejercicio de sus obligaciones.

Es importante mencionar que, como se observa en las definiciones anteriores, los indicadores en cuestión resultan principalmente de la división/cociente entre los subtipos de indicadores de gestión que fueron presentados previamente y que conforman, como se señaló, los componentes de la cadena de valor de un proceso, sea recurrente o no. Lo anterior siempre y cuando sean de diferentes etapas del proceso. Por ejemplo, un indicador de eficacia podría ser el resultado de dividir el logro alcanzado por un proceso (indicador de resultado) entre los recursos empleados en dicho proceso para alcanzar ese logro (indicador de insumo).

Llegados a este punto, es importante mencionar que los subtipos de indicadores de gestión son ampliamente reconocidos y solicitados en el ámbito de la gestión pública contemporánea. Sin embargo, resulta desafiante y poco común encontrar experiencias significativas de construcción y uso efectivo de estos indicadores. Existe una discrepancia notable entre la abundante literatura, la promoción y la demanda de indicadores de eficacia, eficiencia y efectividad, y su implementación real en la administración pública, así como los beneficios sociales obtenidos a través de su uso.

⁶⁸ Esta tipología de indicadores puede recibir varios nombres, por ejemplo, en la *Introducción al diseño, construcción e interpretación de indicadores*, el DANE (2011) se refiere a estos como “indicadores según calidad”. No obstante, para evitar confusiones y ambigüedades, en este libro se denominan indicadores de eficiencia, eficacia y efectividad.

Indicadores según jerarquía

Una manera de responder a la diversidad de tipologías y de estrategias de clasificación de indicadores es a través de la jerarquización, donde se le atribuye una cierta importancia, o pertinencia, a cada indicador. Vale la pena tener en cuenta que la jerarquización tiene, de fondo, una configuración ideológica, y que la atribución de importancia puede depender de las tendencias de gestión o ejecución del momento.

Por ejemplo, en términos del DANE (2011), la clasificación de los indicadores según su jerarquía tiene que ver con qué tan estratégicos o no estratégicos son. La clasificación que presentan señala que entre más relacionado esté un indicador con las dinámicas internas de gestión, como la relación entre insumos y procesos, es menos estratégico. Y entre más relacionado esté con los resultados y el impacto, es más estratégico.

Aunque para el DANE (2011) en el proceso de jerarquización se definen como menos estratégicos los procesos internos y como más estratégicos los elementos relacionados con los resultados y el impacto, la jerarquización y los términos de “más estratégico” o “menos estratégico” podrían emplearse de manera transversal para cualquier proceso de medición y evaluación que involucre indicadores.

Dicho de otro modo, desde la propuesta del presente libro, la jerarquización de indicadores no está relacionada exclusivamente con los indicadores de gestión y, dentro de estos, con los subtipos de indicadores de la cadena de valor de un proceso. Por el contrario, la jerarquía de los indicadores es transversal a las demás tipologías de indicadores y se puede hablar de ella en los indicadores cualitativos, cuantitativos, temáticos o de gestión y, dentro de estos, en cada uno de sus subtipos. Así, dependiendo de las características particulares de los procesos, en cada eslabón de la cadena de valor se pueden definir sus correspondientes indicadores como estratégicos o no estratégicos.

CARACTERÍSTICAS DE LOS INDICADORES

A diferencia de las estadísticas, donde se plantearon características específicas que se dividieron en extrínsecas e intrínsecas, con los indicadores

es difícil hablar de características puntuales porque pueden variar de acuerdo con cada tipología. Por ejemplo, en la tipología de indicadores temáticos no se habla de metas o niveles de desempeño, pero sí de las fuentes de los datos. En el mundo de los cuantitativos o cualitativos las reglas de juego también son diferentes, se puede hablar, por un lado, de percepciones y, por el otro, de cantidades. De manera que este libro no se concentra en formular características puntuales para las diversas tipologías de indicadores, pues excede sus alcances y objetivos⁶⁹.

Ahora bien, aunque no se presente una lista exhaustiva de características para los diferentes tipos de indicadores, hay una serie de condiciones para garantizar su construcción y uso. Por ejemplo, el enfoque SMART, donde los indicadores deben ser *específicos* (*Specific*), es decir, claros y precisos; *medibles* (*Measurable*) o cuantificables; *alcanzables* (*Attainable*), o sea, que se puedan realizar, que sean realistas; *relevantes* (*Relevant*), esto es, alineados con los objetivos; y, finalmente, *temporales*⁷⁰ (*Time-bound*), enmarcados en un periodo de tiempo específico.

NATURALEZA MATEMÁTICA DE LOS INDICADORES

La naturaleza matemática de los indicadores es inherente a su construcción y disposición. Sin importar su tipología (cuantitativos, temáticos, de gestión o jerarquía), los indicadores hacen uso de operaciones matemáticas, ya sean simples o complejas, como parte integral de su proceso de construcción. Es importante destacar que, si bien todos los indicadores comparten una fórmula o mecanismo matemático estadístico para su constitución, dichas fórmulas y mecanismos pueden variar significativamente y se pueden

-
- 69 De hecho, establecer las características y los parámetros de cada tipología debería ser una tarea de las instituciones que trabajan con indicadores.
 - 70 Habría que preguntarse si la acotación en el tiempo es una característica que, efectivamente, deben cumplir todos los indicadores o si, al contrario, es más bien solo para los indicadores de gestión no recurrentes, es decir, para los planes, programas y proyectos. Ahora bien, una característica que comparten los indicadores, ya sean aplicados a procesos recurrentes o no, es que deben tomar en cuenta el factor tiempo como un elemento influyente en su medición. Esto es válido tanto para procesos que buscan perdurar en el tiempo como para aquellos que son de duración acotada. El seguimiento temporal es relevante para evaluar el desempeño y los resultados de cualquier proceso.

clasificar en diversas tipologías según su naturaleza numérica. En el ámbito de lo público, se emplean principalmente las siguientes: medidas estadísticas, frecuencias relativas o absolutas, tasas, razones e índices.

Medidas estadísticas

Una de las estrategias de naturaleza matemática empleada para la conformación de indicadores son las medidas estadísticas que, como se mencionó en el capítulo II del presente libro, se emplean para la conformación de cifras agregadas, principalmente, aquellas derivadas de variables cuantitativas o continuas. Estas medidas se dividen en cuatro grandes grupos. En primer lugar, las medidas de tendencia central, que se utilizan para establecer los valores medios o centrales de una distribución, como por ejemplo, la media o la moda. En segundo lugar, se encuentran las medidas de dispersión, que indican el grado de variación presente en la distribución, tales como la desviación media, la varianza o la desviación estándar. En tercer lugar, se encuentran las medidas de localización, que dividen la distribución en partes iguales, como la mediana, los cuartiles, los deciles o los centiles. Finalmente, las medidas de forma o simetría, que permiten cuantificar el grado de apuntamiento o simetría de una distribución, como la curtosis o la asimetría.

Frecuencias

Las frecuencias son imprescindibles para la elaboración de indicadores, pues permiten conocer la cantidad de veces que ocurre un fenómeno o evento en una población o muestra determinada. Las frecuencias pueden ser de dos tipos, a saber, absolutas y relativas.

- *Frecuencia absoluta:* se refiere al número de veces que se observa una característica específica en una muestra o población de datos o, dicho de otro modo, representa el número de veces que una característica está presente en una muestra o población conformada por n individuos⁷¹. La Ecuación 6 describe la representación formal de la frecuencia absoluta.

⁷¹ El total de individuos que conforman una muestra se nota estadísticamente

ECUACIÓN 6. FRECUENCIA ABSOLUTA

$$\sum_{i=1}^n Z_i$$

En donde:

- $Z_i = 1$ si la característica deseada está presente en el individuo i.
 $Z_i = 0$ si la característica deseada no está presente en el individuo i.

Un ejemplo⁷² de frecuencia absoluta es el conteo que describe la cantidad de graduados en 2022 de la UNAL, es decir, 9057 estudiantes. Otro ejemplo, de frecuencias absolutas, dentro de esta misma población, es el total de mujeres graduadas en el año 2022 en la UNAL, que fue de 3867.

- *Frecuencia relativa:* se refiere a la cantidad de veces que se observa una determinada característica en relación con el total de observaciones que conforman una muestra o población de individuos. Vale la pena resaltar que el numerador de una frecuencia relativa está incluido dentro del denominador, es decir, el numerador es un subconjunto del conjunto total que se expresa en el denominador. El rango de una frecuencia relativa oscila entre 0 y 1. La Ecuación 7 describe la representación formal de la frecuencia relativa.

mediante la letra minúscula n; en contraste, si estos individuos conforman una población, se nota mediante la letra mayúscula N. En las fórmulas de este apartado, por simplicidad, cuando se haga referencia al total de individuos que conforman una población o muestra, se hará uso exclusivo de la letra minúscula n.

⁷² A lo largo de este apartado se hará uso de algunas cifras oficiales contenidas en la página de estadísticas de la UNAL (<http://estadisticas.unal.edu.co/home/>). Como se mencionó, los indicadores son estadísticas en potencia, por lo que estos se fortalecen si se construyen y monitorean a partir de cifras contenidas en sistemas de estadísticas nacionales o institucionales.

ECUACIÓN 7. FRECUENCIA RELATIVA

$$\frac{\sum_{i=1}^n Z_i}{n}$$

En donde:

$Z_i = 1$ si la característica deseada está presente en el individuo i.

$Z_i = 0$ si la característica deseada no está presente en el individuo i.

n = Total de individuos que conforman la muestra o población.

Por ejemplo, de 9057 estudiantes graduados en la UNAL en el año 2022, 5705 se graduaron en programas de pregrado. Es decir, la proporción de estudiantes graduados de pregrado en esta universidad durante el año 2022 fue

$$\text{Frecuencia relativa} = (5709 / 9057) = 0,63$$

Para facilitar el entendimiento, usualmente, las frecuencias relativas tienen un multiplicador M . El más común es el porcentaje, que no es otra cosa que la forma de expresar un número como partes de cada cien. Sin embargo, el multiplicador M podría ser diferente de 100. El rango de una frecuencia relativa cuando se usa un multiplicador oscila entre 0 y el multiplicador M . Así las cosas, cuando una frecuencia relativa hace uso de un multiplicador, su ecuación presenta el siguiente aspecto. La Ecuación 8 describe la representación formal de la frecuencia relativa con un multiplicador.

ECUACIÓN 8. FRECUENCIA RELATIVA CON MULTIPLICADOR

$$\text{Frecuencia relativa} = \left(\frac{\sum_{i=1}^n Z_i}{n} \right) * M$$

En donde:

$Z_i = 1$ si la característica deseada está presente en el individuo i.

$Z_i = 0$ si la característica deseada no está presente en el individuo i.

n = Total de individuos que conforman la muestra o población.

M = Multiplicador empleado.

Por ejemplo, al usar un multiplicador *M* de 100, es decir, al calcular un porcentaje, los 5709 estudiantes de pregrado graduados en 2022 de la UNAL representan el 63 % del total de 9057 graduados en esta institución.

$$\text{Frecuencia relativa} = (5709 / 9057) * 100 = 63 \%$$

Un ejemplo de frecuencia relativa cuyo multiplicador *M* es diferente a un porcentaje, por ejemplo 1000, podría ser la proporción de personas en una muestra o población que padecen determinada enfermedad⁷³. Supongamos que esa población o muestra se conforma de 950.000 personas de las cuales 1500 están contagiadas. La proporción de personas contagiadas, por cada 1000 habitantes de la muestra o población, es:

$$\text{Frecuencia relativa} = (1500 / 950000) * 1000 = 1,58 \text{ personas}$$

Razones

Una razón es una comparación o cociente entre dos o más cantidades donde se pretende señalar el número de elementos del numerador que existen por cada elemento del denominador. Numerador y denominador pueden, o no, provenir de muestras o poblaciones diferentes, es decir, ninguno está incluido dentro del otro, incluso si hacen parte de una misma población.

Por ejemplo, volviendo sobre los estudiantes graduados en 2022 de la UNAL, el número de estudiantes graduados de pregrado es 5709 y el número de estudiantes de posgrado es 3348. Los dos subconjuntos hacen parte de la misma población —es decir, la población del total de estudiantes graduados—, sin embargo, ningún subconjunto incluye al

⁷³ El uso de frecuencias relativas con multiplicadores fue empleado con regularidad para expresar, por ejemplo, las cifras de contagios o fallecimientos en la pandemia de covid-19. De hecho, este tipo de frecuencias se usa ampliamente para expresar la presencia de eventos relacionados con la salud en general, así como con fenómenos demográficos.

otro. Pues bien, la razón entre estudiantes graduados de pregrado por cada estudiante graduado de posgrado es la siguiente:

$$\text{Razón} = 5709 / 3348 = 1,7$$

Así, por cada estudiante graduado de posgrado hay 1,7 estudiantes graduados en pregrado. Es importante mencionar que una estrategia comunicativa, empleada sobre todo cuando se trabaja con elementos que usualmente no se presentan de manera decimal, es expandir por múltiplos de 10 el resultado de la razón hasta que este sea cercano a un número entero. Por ejemplo, para este caso, la razón se podría presentar de la siguiente manera: por cada 10 estudiantes de posgrado graduados hay 17 estudiantes de pregrado graduados.

Una razón en donde el numerador y el denominador provienen de poblaciones diferentes es la razón de estudiantes por profesor de tiempo completo equivalente en una universidad. La población de estudiantes, desde luego, no está incluida dentro de la población de profesores y viceversa; aun así, resulta útil en términos de calidad de la formación universitaria calcular un indicador que acuda al uso de esta razón. Por ejemplo, en la UNAL, en el segundo semestre del año 2022, hubo un total de 57.090 estudiantes matriculados y 2964 docentes de carrera en tiempo completo equivalente vinculados a la institución. En este sentido, la razón de estudiantes por profesor en esta institución, para el periodo 2022-2, fue de 19,3 estudiantes matriculados por cada profesor disponible.

Una de las razones más empleadas en la conformación de indicadores en el contexto de lo público es la razón promedio. Esta razón, en general, es altamente utilizada para representar la distribución o participación de recursos financieros o presupuestos entre los individuos que conforman una población o muestra bajo observación. Por ejemplo, si se quiere conocer la razón promedio —también conocida como *per cápita*— de los aportes que una nación hace para la formación de 1000 estudiantes en una universidad que recibió un presupuesto de \$5000 millones, este sería de \$5 millones por estudiante.

Tasas

Dentro de las distintas tipologías matemáticas asociadas a las fórmulas de los indicadores, las tasas destacan por su amplia variedad de subtipos y aplicaciones en el ámbito público. Específicamente, las tasas se utilizan de manera cotidiana en diversas áreas de interés social, como la economía, la demografía, la educación, la salud pública y la epidemiología, entre otras. Sin embargo, debido a su amplio alcance y complejidad matemática, el estudio y la profundización de las tasas trascienden el alcance de este libro, el cual presenta apenas tres enfoques en el uso de las tasas en la construcción y aplicación de indicadores.

Pues bien, para empezar, una tasa es una razón o proporción entre cantidades en la que se tiene en cuenta el tiempo. En el primer caso, de carácter demográfico, el cociente resulta de dividir un número de acontecimientos sucedidos en un periodo de tiempo por la población media existente durante ese periodo. Por ejemplo, la tasa de mortalidad, que mide la proporción de personas que fallecen en un periodo de tiempo específico en relación con el total de personas⁷⁴. La Ecuación 9 describe la representación formal de la tasa de mortalidad.

ECUACIÓN 9. TASA DE MORTALIDAD

$$\frac{\text{Número de defunciones en un periodo de tiempo}}{\text{Población media en ese periodo}} \times 10^n$$

Además de la población media, en el segundo caso, el cociente de una tasa también puede tener un carácter económico, es decir, hacer referencia a una unidad monetaria. La tasa de cambio muestra la relación que existe entre dos monedas en un periodo de tiempo en particular⁷⁵. Dicha tasa es

⁷⁴ En una tasa, el valor 10 elevado a la n permite ampliar el resultado obtenido en el cociente mediante el uso de múltiplos de 10, así: 10 a la 0 es igual a 1; 10 a la 1 es igual a 10; 10 a la 2 es igual a 100; 10 a la 3 es igual a 1000, y así sucesivamente. Esta técnica de generación de múltiplos de 10 es ampliamente utilizada en tasas relacionadas con la demografía y la salud pública, especialmente cuando el número de casos que presentan una condición evaluada es bajo en comparación con las poblaciones expuestas a tales condiciones durante los períodos de medición.

⁷⁵ En el caso de la variación entre divisas, por ejemplo, las tasas se calculan generalmente en períodos de tiempo diarios.

un indicador que expresa cuántas unidades de una divisa se necesitan para obtener una unidad de la otra en un momento específico. La Ecuación 10 describe la representación formal de una tasa de carácter económico.

ECUACIÓN 10. TASA ECONÓMICA

$$\frac{\text{Cantidad o peso de la moneda local}}{\text{Una unidad de la moneda extranjera}} \times 10^n$$

En el tercer caso, las tasas pueden tener un carácter referente a diversas áreas de interés institucional o estatal. En este punto se habla de tasas de variación, que, según el Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval) de México (2013), expresan un cambio relativo en el tiempo a través del cociente de dos observaciones de una misma variable en diferentes períodos de tiempo (pasado $t-k$ /y presente t) para establecer si hubo un incremento o decremento de esta. Para su cálculo, el periodo más reciente (t) se coloca en el numerador, mientras que el menos reciente (pasado $t-k$) se ubica en el denominador (restando al cálculo del valor más reciente), así como en el denominador. La Ecuación 11 describe la representación formal de una tasa de variación.

ECUACIÓN 11. TASA DE VARIACIÓN

$$\frac{(\text{Variable en el periodo } t) - (\text{Variable en el periodo } t - k)}{\text{Variable en el periodo } t - k} \times 10^n$$

Un ejemplo de una tasa de variación resulta al comparar el total de aspirantes a la UNAL en el año 2022 respecto de la cifra reportada en el año 2010. En el año 2022 esta universidad tuvo un total de 96.095 inscripciones (variable en el periodo t) mientras que en el año 2010 a esta institución se inscribieron 117.325 aspirantes (variable $t-k$). Al calcular la tasa de variación de aspirantes haciendo uso de la Ecuación 11 se obtiene como resultado un valor de -0,181, lo que se traduce en una reducción de -0,181 o del 18,1% (multiplicar el cociente por 10 a la 2) en las inscripciones de esta universidad durante los últimos 13 años.

Las tasas se pueden dividir en tasas generales —también conocidas como brutas— y tasas específicas. Las generales tienen en cuenta a la

población en su totalidad y las específicas a una parte de esa población definida a partir de características como la edad, el sexo, la región, etc., si se trabaja con poblaciones humanas. Por ejemplo, en Colombia, la tasa de cobertura bruta de educación superior de pregrado⁷⁶ del año 2022, entendida como la participación de la población entre 17 y 21 años efectivamente matriculada a nivel universitario en las instituciones de educación superior (IES) del país, fue del 55 % ($2.284.637 / 4.159.592$)⁷⁷. Por otro lado, un ejemplo de tasa de cobertura específica podría ser la desagregación de la tasa bruta general de cobertura de educación superior de pregrado en el año 2022 según los subgrupos de hombres y de mujeres que se ubican entre los 17 y los 21 años a nivel poblacional en el país.

Índices

Dentro de los diversos subtipos de medidas matemáticas presentados en este apartado, los índices se destacan como el nivel más complejo en términos de construcción, medición e interpretación en el contexto de lo público. Debido a su complejidad, su uso se reserva principalmente para la medición de fenómenos de interés social a nivel multilateral entre Estados o a través de organismos encargados de las estadísticas oficiales a nivel nacional. En el ámbito de las entidades públicas, el uso de índices, comparado con otras formas de medición a través de indicadores, es escaso o excepcional.

Los índices son altamente utilizados para medir fenómenos de interés público que no pueden ser evaluados directamente, como la calidad de vida, la evolución de los precios de bienes y servicios, la inteligencia, el nivel de bienestar, entre otros constructos de relevancia social. Estos pueden ser simples o compuestos, basados en una o varias variables asociadas a un constructo o categoría conceptual de interés. Los índices simples se apoyan en un solo indicador o variable para evaluar una característica o fenómeno específico de interés. Un ejemplo claro de índice

⁷⁶ SNIES. Resumen Indicadores Educacion Superior. <https://snies.mineducacion.gov.co/portal/Informes-e-indicadores/Resumen-indicadores-Educacion-Superior/>

⁷⁷ A partir de esta tasa se concluye que, en Colombia, para el año 2022, de cada 100 personas que se ubicaban en el rango de edad comprendido entre los 17 y los 21 años, 54 de ellas se encontraban matriculadas en, al menos, un programa de pregrado ofertado por las instituciones de educación superior del país.

simple es la incidencia o la prevalencia de las enfermedades que hacen parte de la epidemiología o de la salud pública.

Por otra parte, se habla de índices compuestos cuando se basan en varias variables u otros índices. Su función principal es sintetizar o resumir la información que proporcionan otros índices simples o variables que hacen parte de su composición. Los índices compuestos pueden ser de dos tipos, no ponderados y ponderados. Un índice compuesto no ponderado suma los valores de cada índice, variable o indicador simple para obtener una puntuación total. Por otro lado, un índice compuesto ponderado asigna un valor particular a cada indicador simple que lo compone.

Un ejemplo de índice compuesto ponderado es el puntaje básico de matrícula (PBM) de la UNAL, que mide la capacidad económica de las familias de los estudiantes matriculados en esta institución y define el valor de la matrícula que cada uno de ellos debe pagar semestralmente. La metodología requerida para la construcción de este índice fue definida a través de la Resolución 2416 de 1993 de la Rectoría, y se basa en la información de siete variables con un puntaje, un peso ponderado y una función específica dentro de la fórmula asignada a cada una de ellas. La Ecuación 12 describe la representación formal del PBM de la UNAL.

ECUACIÓN 12. PBM

$$PBM = (0.4A_1 + 0.3A_2 + 0.3A_3) \times B_1 \times B_2 \times B_3 \times B_4$$

Donde las variables A₂ y A₁ constituyen al estrato⁷⁸ y el valor de la pensión escolar en el último año de secundaria respectivamente, y son indicadores del nivel de gastos en el hogar, mientras A₃ representa el nivel de ingresos. La variable B₁ es el tipo de colegio, B₂ es el lugar de residencia, B₃ propiedad de la vivienda familiar y B₄ es el número de hijos dependientes del ingreso del hogar y, son igualmente, indicadores del nivel de gastos⁷⁹.

-
- ⁷⁸ Según el DANE (s. f. a), la estratificación socioeconómica en Colombia se refiere a la clasificación de inmuebles residenciales con el propósito de cobrar de manera diferencial los servicios públicos domiciliarios.
- ⁷⁹ En el contexto de la educación superior pública de Colombia existen otras experiencias de índices como, por ejemplo, los de capacidad y de resultados del Modelo del Sistema Universitario Estatal (SUE), a través del cual se distribuían, por vigencias, recursos financieros entre las instituciones pertenecientes al Sistema. Tanto el índice de capacidad como el de resultados emplearon la metodología estadística de componentes principales para su construcción y sus cálculos.

CAPÍTULO IV

Modelo P2C-T

Antes de empezar con el capítulo IV de este libro, se hará un breve repaso por lo que se ha expuesto hasta ahora. Para empezar, en el capítulo I se abordó el concepto polisémico de información desde una perspectiva pragmática. Se enfatizó la idea de la información como cosa, susceptible de ser manipulada⁸⁰, analizada y procesada. En este sentido, se destacó la estrecha relación entre la información cuantitativa y la gestión pública, donde a través de los datos se implementan estrategias o políticas de interés para los miembros de una comunidad o un país.

Más adelante, en el capítulo II, se examinó el concepto de estadísticas como un conjunto estructurado de información derivada de cifras descriptivas, el cual desempeña un papel fundamental en la representación y conexión social de los “hechos notables de un estado”. Estas estadísticas son de gran importancia para la gestión pública, ya que permiten describir y analizar de manera significativa la realidad social. Las estadísticas, como se mostró, presentan rasgos o condiciones necesarias, o intrínsecas, sin las cuales no se podría hablar de estadísticas, así como condiciones complementarias, o extrínsecas, pertinentes para su construcción y disposición en el contexto de lo público.

Para continuar, en el capítulo III se abordó el concepto de indicador como una herramienta de medición empleada en la evaluación y el monitoreo de fenómenos de interés en el ámbito de lo público. Se examinó la estrecha relación entre las cifras estadísticas y los indicadores, así como los diversos usos prácticos de estos últimos. Además, se presentó una tipología de indicadores que abarca diferentes categorías, resaltando la importancia de comprender y utilizar en conjunto cada tipo para lograr buenos procesos de gestión en lo público. Por último, se subrayó la

⁸⁰ Es decir, que es susceptible de ser operada o intervenida.

naturaleza matemática que está a la base de la creación de indicadores, así como las diversas estrategias empleadas en su formulación.

Ahora bien, en este capítulo se exponen elementos y recomendaciones prácticas que facilitan la conformación de estadísticas e indicadores, así como de áreas de estadística o analítica de datos al interior de las instituciones. Con este objetivo en mente, se expone un modelo de carácter administrativo titulado Modelo P2C-T, compuesto por cinco dimensiones, a saber, *planeación estadística* (P), *conceptualización y metodologías* (C), *producción estadística* (P), *disposición y gestión de herramientas tecnológicas* (T) y, finalmente, *cultura estadística* (C). Asimismo, se resaltan tres soportes específicos con los que deben contar las organizaciones y entidades para que el modelo sea funcional: 1) *apoyo institucional*, 2) *acceso y disposición del talento humano* y 3) *acceso a recursos financieros*.

DEFINICIÓN

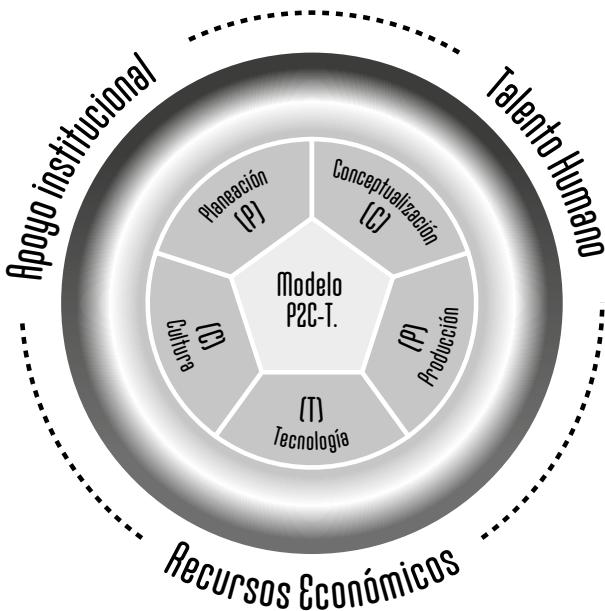
Como se ha visto hasta ahora, los datos⁸¹ que resultan de los procesos de gestión en las instituciones, además de aumentar en cantidad y variedad, se consolidan como un bien de alto valor que permite, entre otras, generar información útil para la toma de decisiones; la transparencia institucional; la rendición de cuentas; el fortalecimiento de entidades y gobiernos abiertos; el seguimiento a la implementación de planes, programas o proyectos, y la realización de evaluaciones a políticas públicas. Pues bien, para hacer uso de los datos, así como de las estadísticas, indicadores y demás estrategias de medición que surgen a partir de ellos, las entidades u organizaciones deben reconocer que tal ejercicio no es coyuntural o puntual, sino que exige una apuesta institucional decidida y sostenible que demanda la definición de un modelo que permita su desarrollo a nivel institucional.

En lo que sigue del presente libro, se define el Modelo P2C-T, que resulta de la experiencia acumulada por la UNAL durante los últimos años. El modelo se define con el propósito de fortalecer la conformación de estadísticas e indicadores, especialmente en el ámbito de la educación

81 Es decir, la información como cosa, que es susceptible de ser usada, entre otras, para la conformación de estadísticas e indicadores.

superior, en particular en las instituciones públicas. Así las cosas, el objetivo de este capítulo no es proponer una metodología para que sea adoptada por las instituciones, para ello se recomienda la consulta de los modelos establecidos a nivel estatal.

Gráfica 5. Representación gráfica Modelo P2C-T



Fuente: elaboración propia.

Como se observa en la Gráfica 5, el Modelo P2C-T está compuesto por cinco dimensiones y tres soportes que favorecen la implementación a nivel público de acciones orientadas al uso de los datos disponibles. Las dimensiones del modelo reconocen, primero, la importancia de la planeación estadística de corto, mediano y largo plazo (P); segundo, la disposición de elementos conceptuales y metodológicos que orienten el ejercicio cuantitativo (C); tercero, la disposición de datos y la producción con calidad y de manera recurrente de la información estadística requerida a nivel institucional (P); cuarto, el acceso y el dominio de las herramientas tecnológicas requeridas para la gestión contemporánea de los datos (T); quinto, el fomento de una cultura estadística en las entidades interesadas en la gestión y uso

de la información cuantitativa (C). Las dimensiones en consideración no se pueden desarrollar de manera adecuada sin la disposición de tres soportes requeridos a nivel institucional: primero, apuesta institucional para acceder al valor contenido en los datos disponibles; segundo, acceso y disposición de recursos financieros para el fortalecimiento de la actividad estadística y, tercero, acceso al talento humano especializado en la gestión de datos de naturaleza cuantitativa.

A continuación, se desarrollan cada una de las dimensiones del modelo propuesto, así como los soportes requeridos para su implementación y desarrollo.

PLANEACIÓN ESTADÍSTICA (P)

La planeación es necesaria en el ámbito de la gestión en general y de la pública en particular. Sin importar la entidad u organización, se busca establecer una serie de pasos y estrategias que permitan alcanzar determinados objetivos, es decir, se establece un plan. De acuerdo con Rodríguez y Bernal (2019):

La planeación en su forma más general puede ser entendida como la capacidad que tiene una entidad, a partir del comportamiento observado en el pasado, el presente y lo que ocurre en el contexto, de visualizar y formular un futuro deseado y fijar los objetivos, las metas y las estrategias requeridas para alcanzar el horizonte visualizado. Para ello, se apoya en políticas, recursos, instrumentos y mecanismos que garantizan el cumplimiento de los propósitos deseados. (§ 1.2.1)

Pues bien, la producción estadística se ve altamente favorecida si se hace uso de instrumentos de planeación general para el análisis y la proyección a nivel de las entidades. De modo que la primera apuesta que debe ser implementada en una entidad pública interesada en la producción estadística y la gestión estratégica de los datos disponibles es un ejercicio de planeación en general y de planeación estadística en particular.

La planeación permite definir una secuencia de acciones que buscan responder a unos propósitos particulares, en el caso de la planeación estadística, asociados a la consolidación de cifras, datos y, en general,

información resultante de estrategias cuantitativas. Así, la planeación estadística permite, por ejemplo, recopilar, analizar, presentar y gestionar los datos relevantes de manera efectiva y oportuna; satisfacer las demandas externas para que la información resultante de los procesos de gestión cumpla con los estándares establecidos; y, además, fortalecer la toma de decisiones basada en evidencias sólidas.

La importancia de la planeación estadística es tal, que en los Objetivos de Desarrollo Sostenible, el número 17 (PNUD, s. f), sobre el fortalecimiento de los medios de implementación y la revitalización de la Alianza Mundial para el Desarrollo Sostenible, se destaca la necesidad de mejorar las capacidades estadísticas y fortalecer la producción y el uso de datos confiables y desagregados, lo que muestra la relevancia de una planeación estadística sólida para abordar los desafíos socioeconómicos, evaluar el progreso, y orientar las políticas y acciones hacia el desarrollo sostenible.

Ahora bien, de acuerdo con las necesidades y los objetivos de cada entidad u organización, se pueden establecer diversas aproximaciones y metodologías para la creación de un plan. Para la elaboración y ejecución de un plan estadístico, en Colombia, el DANE (2020c) resalta las siguientes cuatro etapas: *formulación, implementación, seguimiento y evaluación*.

La primera etapa, de *formulación*, se caracteriza, en términos del DANE (2020c), por una fase de preparación, en la que se definen elementos como el alcance, la finalidad, los actores involucrados, entre otros. Luego, una fase de identificación, recolección y organización de información disponible que sirve para elaborar un diagnóstico de la actividad estadística vigente. Finalmente, se construye la parte estratégica del plan (objetivos, estrategias, acciones y metas) y se comunica el documento del plan estadístico resultante con los actores involucrados para la realimentación, aprobación y difusión. En esta etapa de formulación, se busca caracterizar los registros administrativos para que sirvan como fuente primaria de información para la consolidación de poblaciones y muestras.

La segunda etapa, de *implementación*, consiste, de acuerdo con el DANE (2020c), en la puesta en práctica de aquello que se definió en la etapa de *formulación* para el plan estadístico. La *implementación* involucra el trabajo con diferentes actores, de modo que, en primer lugar, es necesario un proceso de articulación y coordinación entre los actores a través de una estrategia de comunicación que defina el rol y las actividades de

cada uno. Para esta segunda etapa es importante programar y realizar las acciones definidas en el plan de acción y generar los productos acordados en los plazos establecidos.

La tercera etapa, de *seguimiento*, comprende, siguiendo al DANE (2020c), el monitoreo periódico que se hace de la implementación del plan estadístico. Para ello, se registran de manera sistemática los avances y las evidencias de la ejecución; se presentan los resultados a través de informes de seguimiento y se definen las acciones preventivas y correctivas.

La cuarta y última etapa, de *evaluación*, conforme con el DANE (2020c), pretende valorar si se alcanzaron los objetivos planteados y determinar los resultados. Para ello, la evaluación se desarrolla en dos momentos, una evaluación intermedia a la mitad de la vigencia del plan estadístico, y otra final, cuando ha concluido la vigencia. La intermedia pretende ajustar la estrategia para conseguir el efecto esperado, mientras que la final permite determinar qué estrategias necesitan un rediseño en el próximo plan estadístico. Finalmente, es importante señalar que con la etapa de *evaluación* se busca un mejoramiento de la actividad estadística en general, es decir, la producción, difusión y articulación entre los miembros del ecosistema de datos, así como en la calidad de la información, la accesibilidad y el uso por parte de los usuarios en el ecosistema.

El ejemplo por antonomasia de plan estadístico en Colombia es el denominado Plan Estadístico Nacional (PEN) (DANE, 2020e), que es el principal recurso estratégico en la política estadística de la nación, con el objetivo central de fomentar el desarrollo, la divulgación y la aplicación de datos estadísticos. Su función primordial radica en la definición de estrategias encaminadas a asegurar la disponibilidad de información estadística precisa y detallada, facilitando, de esta manera, la toma de decisiones basadas en evidencia.

Por otra parte, la UNAL cuenta con el Plan Estadístico Institucional de la UNAL (2018), que se realizó con el acompañamiento del DANE y para el cual se definieron cinco líneas de acción que se corresponden con las dimensiones del modelo P2C-T, sobre el cual versa este capítulo. El plan tiene como objetivo, entre otras cosas, fortalecer la producción estadística a nivel institucional y contribuir al desarrollo de la actividad estadística a nivel nacional, sobre todo, para el sistema de educación superior. Para ello se divide en cuatro líneas estratégicas: primero, la construcción de elementos conceptuales de referencia; segundo, la producción estadística;

tercero, la gestión de las capacidades técnicas y tecnológicas; y, cuarto, la promoción de la cultura estadística.

CONCEPTUALIZACIÓN Y METODOLOGÍA (C)

Como se ha visto hasta ahora, el mundo de los datos está en constante crecimiento y la información disponible es tan extensa como variada. Ya sea a través de redes sociales y medios digitales, bases de datos y registros administrativos o encuestas y cuestionarios, hay una riqueza en términos de datos que permite mejorar procesos clave para el ámbito de la gestión, como son la toma de decisiones, la optimización de procesos o la evaluación y el monitoreo. En el contexto de lo público, la gestión a través de los datos adquiere, quizá, una relevancia mayor. Los países y sus entidades tienen a disposición una gran cantidad de información que permite describir y comprender, entre otras, las necesidades puntuales de una población, el estado actual y las posibilidades de mejora de procesos de gestión, así como el impacto de las directrices políticas implementadas.

No obstante, la cantidad y diversidad de datos que pueden emplearse en la gestión no garantiza de por sí el éxito o el mejoramiento de los procesos de gestión. De hecho, para que en efecto haya un avance en dichos procesos, es necesario contar con metodologías y marcos conceptuales definidos y consensuados. Así las cosas, podría decirse que para propiciar una gestión efectiva y de calidad es tan importante basar los procesos de gestión en fuentes de datos, como gestionar los datos sobre los cuales se sostienen tales procesos.

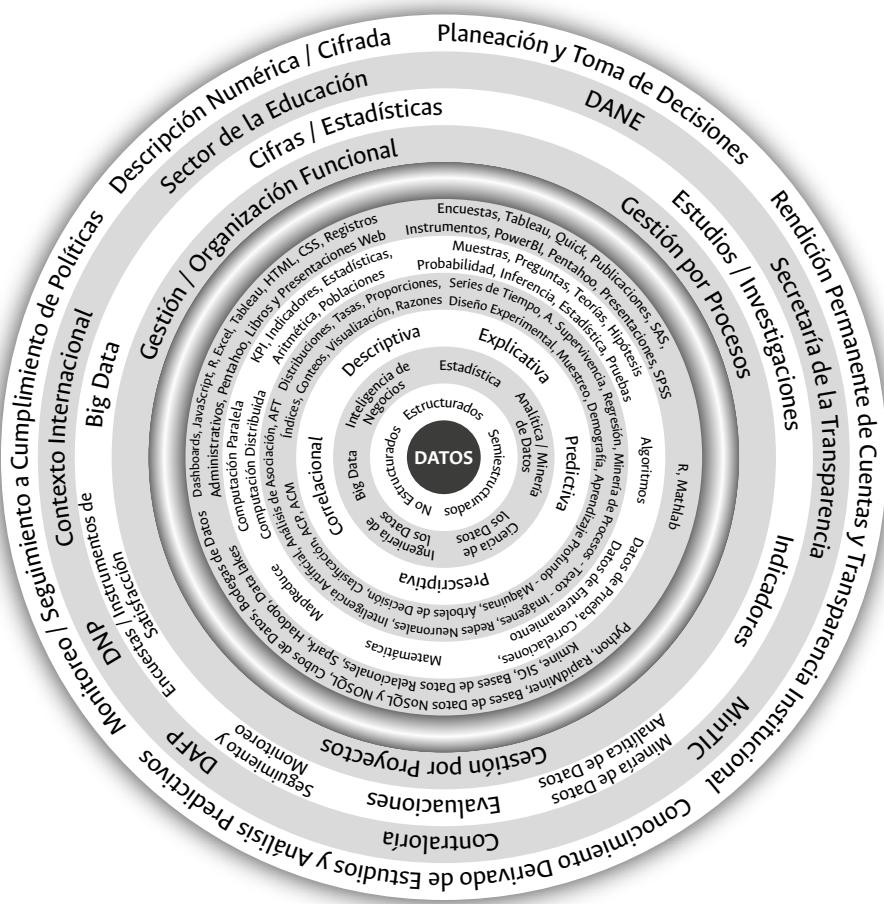
Pues bien, la gestión de los datos implica tener en cuenta que los diferentes enfoques estadísticos que se pueden aplicar a los datos⁸² y sus distintos usos en el ámbito público⁸³ llevan consigo aspectos conceptuales, metodológicos y procedimentales diversos. Por ejemplo, diversidad de directrices, lineamientos, normativas y políticas; variedad en los alcances y usos de los datos cuantitativos; multiplicidad de disciplinas y enfoques interesados en el abordaje de datos; diversidad de propósitos y

⁸² Por ejemplo, descriptivo, correlacional, explicativo, predictivo y prescriptivo.

⁸³ Tales como la toma de decisiones, la implementación de evaluaciones de impacto, la rendición pública de cuentas o la transparencia institucional.

aplicaciones; amplia variedad de herramientas tecnológicas disponibles, entre otras. La unión de estos elementos crea un entramado complejo que orbita alrededor del mundo de los datos en la actualidad, como se muestra en la Gráfica 6.

Gráfica 6. Contexto contemporáneo de los datos



Fuente: Rodríguez y Bernal (2019).

De este modo, la segunda dimensión del modelo P2C-T, sobre *conceptualización y metodología*, tiene que ver con definir criterios para la clasificación y organización de los datos, así como normas y estándares para

su recopilación, almacenamiento y uso. Lo anterior teniendo en cuenta la amplia gama de tipos de datos disponibles, enfoques de análisis, estrategias estadísticas, entidades interesadas y demás componentes propios de la diversidad de la información cuantitativa. Si tales criterios y definiciones no son tenidos en cuenta, los datos se pueden usar de manera inadecuada y, así, en la disposición de cifras o análisis cuantitativos errados o con bajos niveles de calidad que generan desinformación y desconfianza.

En busca de establecer una suerte de consenso en términos de conceptualización y metodología, se han desarrollado guías y manuales que buscan facilitar la aplicación práctica del trabajo con datos. Desde la Oficina Nacional de Estadística de la Universidad Nacional de Colombia, por ejemplo, se han destinado esfuerzos importantes al establecimiento de lineamientos conceptuales para la gestión de estadísticas e indicadores⁸⁴.

Como resultado del esfuerzo, se presenta este libro con la intención de proporcionar apoyo conceptual a las entidades públicas en el estudio y la construcción de estadísticas e indicadores institucionales. Además de este texto, la UNAL pone a disposición de los interesados un documento conceptual previo titulado *Gestión de la información cuantitativa – Pistas para su abordaje en la era de la sobreinformación* (Rodríguez y Bernal, 2019), así como una guía metodológica para el diseño y la construcción de indicadores de procesos recurrentes (Universidad Nacional de Colombia, s. f.). Estos documentos, junto con futuros trabajos, pretenden contribuir al desarrollo de lineamientos que faciliten la comprensión de la dimensión conceptual y metodológica que constituye y debe ser parte integral de la gestión cuantitativa en el ámbito público, sobre todo, de la educación superior.

⁸⁴ El trabajo conceptual y metodológico adelantado hasta este momento por la Oficina Nacional de Estadística de la Universidad Nacional de Colombia se ha centrado principalmente en el estudio de la dimensión de las estadísticas y los indicadores en el escenario de lo público y, en especial, en el de las estadísticas oficiales. No obstante, falta aún profundizar a nivel conceptual en otros posibles usos de los datos disponibles en los Estados y sus entidades como, por ejemplo, la implementación de evaluaciones de impacto, la realización de estudios e investigaciones de naturaleza correlacional o explicativa, la implementación de predicciones, así como la implementación de análisis de naturaleza prescriptiva.

PRODUCCIÓN ESTADÍSTICA (P)

La tercera dimensión del modelo se refiere a la producción estadística, que es, tal vez, la más importante en el modelo, pues las demás dimensiones y soportes propuestos adquieren sentido en función de esta. El propósito de gestionar datos en una entidad pública es, en sentido estricto, el de extraer el valor contenido en estos para garantizar una producción de información estadística útil, de calidad, sostenible en el tiempo y al servicio de los intereses institucionales y de la sociedad en general. Esta actividad debe realizarse en el marco de un *proceso estadístico* bien definido. Es decir, que se dé a través de una serie de pasos organizados, secuenciales, regulares y recurrentes que se llevan a cabo con el fin de lograr un objetivo específico. La implementación de un *proceso* definido garantiza la estandarización de las operaciones, mejora la eficiencia y la calidad, y facilita el seguimiento de la producción estadística.

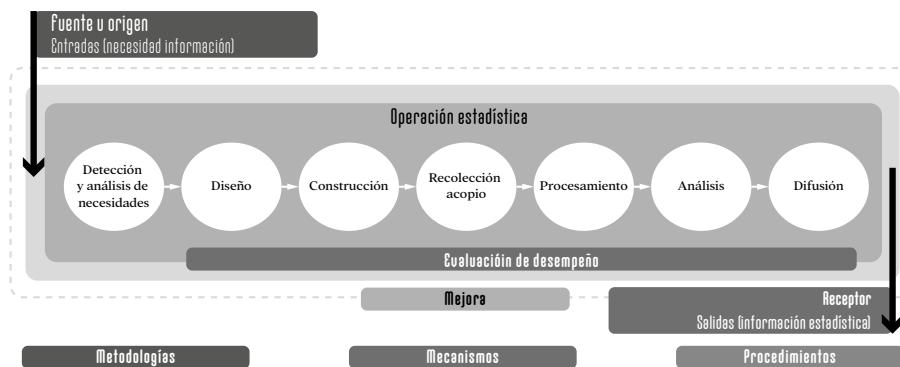
De acuerdo con el DANE (s. f. b), el proceso estadístico debe tener unos atributos particulares que garantizan su calidad, a saber: accesibilidad, claridad, credibilidad, coherencia, comparabilidad, oportunidad, puntualidad, precisión y relevancia. Asimismo, el DANE (s. f. b) resalta la importancia de que en el proceso estadístico se evalúen la planeación, el diseño, la producción, el análisis y la difusión de los resultados de la operación estadística.

En Colombia, el proceso estadístico se establece y se evalúa a través de la Norma Técnica de la Calidad del Proceso Estadístico⁸⁵ (NTC PE 1000) del DANE (2020d), que describe las siguientes fases constitutivas de la actividad de producción estadística: detección y análisis de necesidades, diseño, construcción, recolección o acopio, procesamiento, análisis, difusión y evaluación sistemática. La norma pretende establecer los requisitos, o pasos, de calidad que se deben cumplir en un proceso de generación de estadísticas e indicadores oficiales por parte de las entidades públicas que hacen parte del Sistema Estadístico Nacional (SEN).

En las gráficas 7 y 8 se describen las fases del proceso estadístico desde la norma NTC PE 1000:

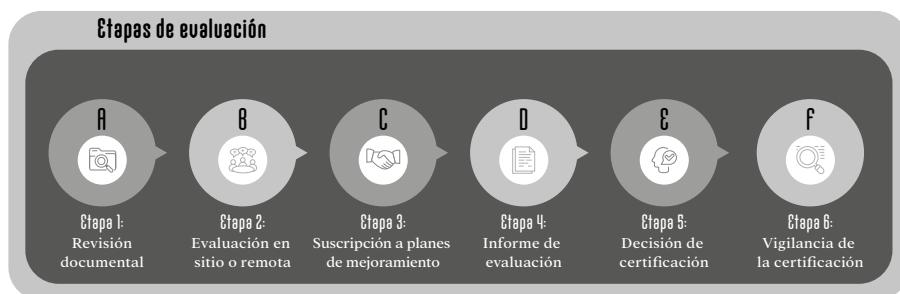
85 Con base en esta norma se construye el documento del DANE (2020b) que establece, de manera formal, los lineamientos para el proceso estadístico en el Sistema Estadístico Nacional.

Gráfica 7. Proceso estadístico desde la norma NTC PE 1000



Fuente: adaptado del DANE.

Gráfica 8. Etapas de evaluación NTC PE 1000



Fuente: adaptado del DANE.

Para el caso de la UNAL, la información estadística oficial se dispone y produce a través de un *proceso* que involucra las siguientes etapas: captura, importación, limpieza y transformación, producción y visualización, comunicación y difusión⁸⁶. Asimismo, desde el Sistema Integrado

⁸⁶ Las etapas, como es de esperar, se fundamentan en los lineamientos del DANE para el proceso estadístico en el Sistema Estadístico Nacional. No obstante, se adaptan a la dinámica interna de la creación de cifras estadísticas dentro de la UNAL. No se busca promover la transición a un nuevo modelo ni proponer alternativas a los documentos establecidos a nivel estatal en Colombia; simplemente,

de Gestión Académica, Administrativa y Ambiental (SIGA) se disponen, para la UNAL, una serie de procedimientos del Sistema de Gestión de Calidad que buscan, entre otras cosas, armonizar los sistemas de calidad internos con normas nacionales, como la NTC PE 1000 (DANE, 2020d).

A modo de exemplificación de lo abordado en esta sección, se resalta que la producción estadística de la UNAL en la actualidad —que se dispone en la página de estadísticas institucionales— está orientada a la producción de estadísticas institucionales, entre las que se destacan, por ejemplo, estadísticas de inscritos, admitidos, estudiantes matriculados, información de graduados, información de programas académicos, estadísticas de docentes y funcionarios, estadísticas de investigación y extensión, cifras financieras, entre otras.

TECNOLOGÍA (T)

Como se mencionó en este libro, tecnología y estadística han estado siempre vinculadas⁸⁷. Ya sea a través de recuentos a lápiz, codificaciones digitales, o *scripts* de R o Python, la producción de cifras estadísticas se ha visto permeada por herramientas que la simplifican y enriquecen. La cuarta dimensión del modelo P2C-T se refiere, pues, al uso e implementación de herramientas tecnológicas para la producción de estadísticas. Como se sabe, las TIC han transformado el proceso estadístico, lo que ha permitido una gestión más eficiente y efectiva de la información cuantitativa. Sin embargo, es crucial recordar que la tecnología debe estar al servicio de la calidad, la producción, la transparencia y la ética en la generación y el uso de las estadísticas. No basta con hacer uso de la tecnología, sino que, además, su uso debe ser adecuado.

A continuación, se exponen cuatro aspectos relacionados con el uso de las TIC en la gestión estadística. En primer lugar, se presenta una propuesta de arquitectura ideal para la gestión de datos en las entidades,

se quiere compartir la experiencia en generación de cifras estadísticas en la UNAL.

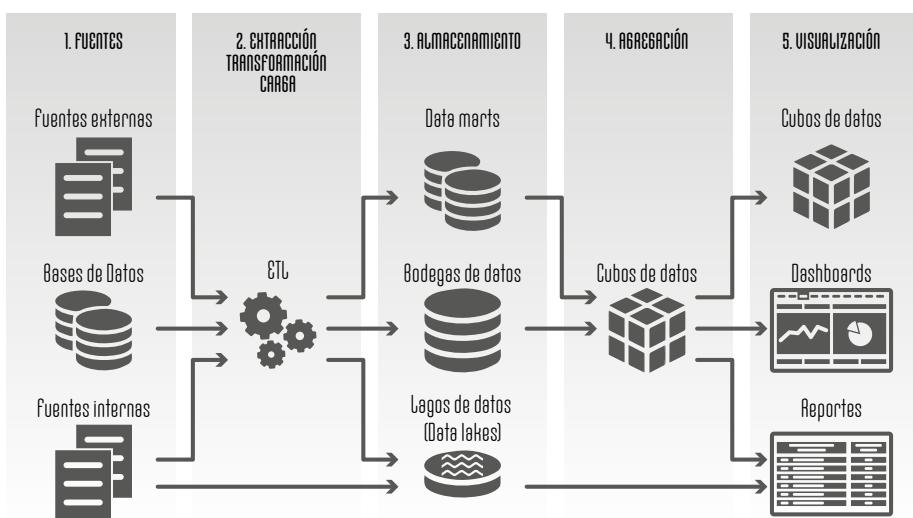
⁸⁷ En el capítulo II del presente libro, una de las características extrínsecas de las estadísticas alude al uso intensivo de las TIC, donde se menciona la relación entre tecnología y consolidación de cifras estadísticas.

sobre todo, de educación superior. Posteriormente, se describen los *softwares* disponibles para la gestión especializada de datos. Luego, se detallan las herramientas utilizadas por la UNAL en la gestión estadística. Por último, se destaca el aporte de la UNAL a través de la creación del paquete de R UnalR, diseñado para apoyar la gestión y producción de estadísticas institucionales.

Infraestructura tecnológica general para la gestión de datos

En el mundo de la gestión estadística, las herramientas tecnológicas se suelen emplear a través de una propuesta de arquitectura tecnológica sugerida para ser implementada en aquellas entidades interesadas en extraer de manera eficiente el valor contenido en los datos. Comúnmente, la arquitectura necesaria es proporcionada por la inteligencia de negocios.

Gráfica 9. Arquitectura tecnológica típica de un modelo de inteligencia de negocios. Adaptación con base en arquitecturas semejantes



Fuente: elaboración propia con base en el libro de *Gestión de la Información*. https://estadisticaun.github.io/L_Conceptual/2-4-inteligencia-de-negocios.html

El primer componente de la arquitectura tecnológica propuesta en la Gráfica 9 son los datos contenidos en bases de datos internas de sistemas de información, así como en registros internos, fuentes externas, encuestas o censos. A partir de allí, el segundo componente de un modelo de inteligencia de negocios son las ETL, a través de las cuales se *extraen* (E) los datos disponibles en las fuentes; se *transforman* (T) según los requerimientos definidos a nivel funcional en las entidades, y se *cargan* (L), o disponen, en fuentes de almacenamiento orientadas a la disposición de los datos bajo estructuras que facilitan el uso de estos con propósitos analíticos. Los desafíos en este componente incluyen desconocimiento de la arquitectura de las bases de datos, duplicidad de acciones, tecnologías diferentes, ausencia de normalización de los datos, falta de interoperabilidad, baja calidad de datos, entre otros. En la tercera fase, los datos extraídos y transformados se almacenan en bodegas de datos, *data marts* o *data lakes*. Estas bases de datos están diseñadas para mantener datos históricos provenientes de diversas fuentes de información.

Los componentes siguientes de una arquitectura de inteligencia de negocios orientada a la gestión y disposición de los datos disponibles en las entidades son las herramientas y aplicaciones tecnológicas que suministran información a los usuarios finales. Estas incluyen, por ejemplo, cubos de datos, reportes y tableros de mando (*dashboards*). Los cubos de datos son estrategias tecnológicas útiles para la visualización y el análisis de datos desde diferentes perspectivas y dimensiones; los reportes suministran información tabular de interés institucional, y los *dashboards* son la herramienta gráfica más utilizada en la actualidad para gestionar la información cuantitativa en una organización.

Softwares especializados en gestión de datos

Dentro de las herramientas tecnológicas más empleadas en la actualidad para el proceso de producción de cifras estadísticas están los *softwares* especializados en gestión de datos. Tales *softwares* son aplicaciones informáticas diseñadas para organizar, almacenar, analizar, visualizar o modelar grandes volúmenes de datos de manera eficiente y efectiva.

Los *softwares* especializados en gestión de datos se pueden clasificar en dos grandes grupos: focalizados y diversificados. Los primeros hacen referencia a recursos cuya funcionalidad y capacidad de procesamiento

se centra en procesos particulares, como la visualización, y que permiten realizar operaciones matemáticas simples, así como análisis y modelamiento básico. Además, son herramientas que funcionan principalmente a través de interfaces gráficas de usuario (GUI), por lo que no requieren altos conocimientos en programación; sin embargo, su funcionalidad está limitada a las características incorporadas en el *software*. Estos *softwares*, además, son poco escalables en la medida en que tienen limitaciones en términos de capacidad de procesamiento y manipulación de grandes volúmenes de datos y requieren, generalmente, de licencias para su uso. Algunos ejemplos de este subgrupo de *softwares* son Excel, Tableau, Power BI o Data Studio.

Por su parte, los *softwares* diversificados ofrecen una funcionalidad y capacidad de procesamiento mayor, pues permiten realizar tareas complejas, como modelado estadístico, aprendizaje automático (*machine learning*) y visualización interactiva y estática de datos. Asimismo, no están restringidos a una interfaz como tal; al contrario, se pueden extender con bibliotecas y paquetes adicionales que permiten una mayor flexibilidad y personalización en el procesamiento de datos, análisis y creación de visualizaciones adaptadas a las necesidades específicas del usuario. Sin embargo, es necesario contar con conocimientos en programación para aprovechar la flexibilidad y personalización que ofrecen. Las herramientas en cuestión son altamente escalables en la medida en que permiten el análisis de datos a gran escala y no requieren, generalmente, licenciamiento. Algunos ejemplos son R, SQL, Julia o Python. La Gráfica 10 da cuenta de los *softwares* especializados en gestión de datos focalizados y diversificados, así como de sus principales características.

Herramientas tecnológicas utilizadas por la UNAL para la gestión de los datos

Hasta la fecha, en la UNAL, al igual que en muchas entidades públicas de este y otros países, no se ha implementado una arquitectura tecnológica sólida y enfocada en la gestión efectiva de datos, como la arquitectura basada en *inteligencia de negocios*, mencionada en la primera parte de este apartado. No obstante, esta situación no ha impedido el uso de diversas herramientas tecnológicas alternativas que facilitan la

Gráfica 10. Herramientas tecnológicas para la gestión de datos

	Ventajas	Desventajas	Herramientas
<i>Software focalizados</i>	<ul style="list-style-type: none"> • Amigables • Curva de aprendizaje rápida • Baja competencia académica y técnica • Bajos conocimientos en programación • Soporte técnico 	<ul style="list-style-type: none"> • Poco escalables • Requiere licenciamiento - altos costos • Descuidan la importancia académica y técnica de los datos • Ciencia cerrada 	   
<i>Software diversificados</i>	<ul style="list-style-type: none"> • Altamente escalables • No requieren licenciamiento • Importancia académica y técnica de los datos • Promoción de la ciencia abierta 	<ul style="list-style-type: none"> • Conocimientos en programación • No cuentan con soporte técnico • Curva de aprendizaje lenta 	   

Fuente: elaboración propia.

gestión de los datos disponibles y permiten, por ejemplo, disponer de cifras estadísticas oficiales como las que se encuentran en la página de estadísticas de la UNAL⁸⁸.

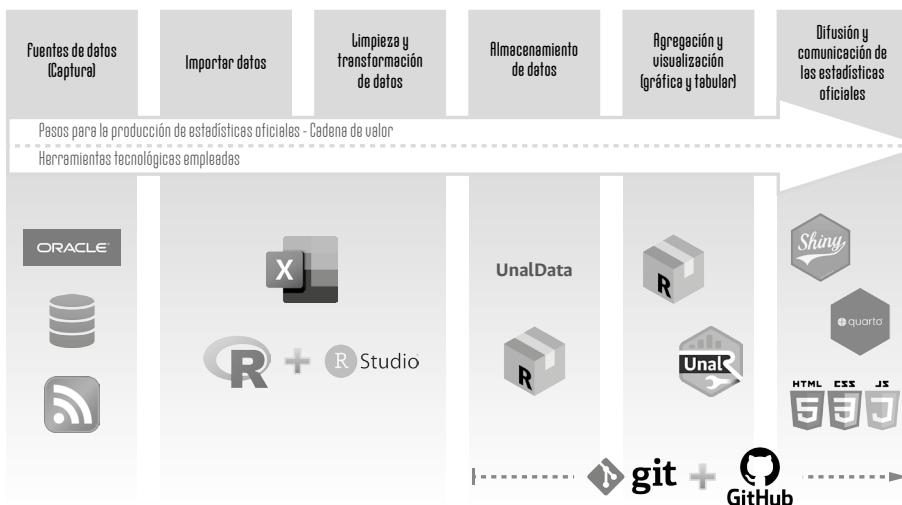
Con todo, en el ciclo de producción de las estadísticas institucionales de la UNAL, se emplean diversas herramientas tecnológicas a lo largo del proceso de producción. Como se muestra en la Gráfica 11, para la captura de datos se utilizan, principalmente, bases de datos Oracle; para la importación, limpieza y transformación preliminar, se recurre a Excel y otras herramientas similares; para la limpieza y transformación final de los datos, así como para la construcción de tablas y gráficas, se emplean, sobre todo, R y RStudio; en cuanto al control de versiones, disposición de repositorios, alojamiento de aplicaciones en la nube y trabajo colaborativo, se utilizan Git, GitHub y Amazon Web Services (AWS); y, por último, para la comunicación y difusión de los resultados se emplean lenguajes como HTML, CSS y JavaScript (JS) para la construcción y disposición de la información estadística en ambientes web, así como en aplicaciones haciendo uso de Shiny de R.

Es relevante destacar que las herramientas utilizadas en el proceso de producción estadística en la UNAL pueden tener aplicaciones diversas

⁸⁸ UNAL en un vistazo. <http://estadisticas.unal.edu.co/home/>

en otras entidades, adaptándose a las necesidades y los procesos específicos de cada una. Dicho de otro modo, la variedad de herramientas y posibilidades existentes permite que cada entidad pueda elegir su propio camino en función de sus necesidades y preferencias.

Gráfica 11. Herramientas tecnológicas empleadas por la UNAL en la producción de estadísticas



Fuente: elaboración propia.

Aporte de la UNAL a la producción de estadísticas institucionales. El paquete de R UnalR

La UNAL, durante los últimos años, ha realizado esfuerzos importantes para contribuir en la construcción y disposición de recursos tecnológicos que faciliten la gestión de cifras, especialmente en el ámbito de las estadísticas institucionales. En este sentido, se ha desarrollado el paquete UnalR⁸⁹ para el lenguaje de programación R, con el objetivo de facilitar y mejorar la gestión de microdatos, así como la visualización de cifras y estadísticas institucionales dentro del sistema de planeación institucional.

89 UnalR. <https://estadisticaun.github.io/UnalR/index.html>

Este paquete de R, de alto uso en la UNAL⁹⁰ y disponible de manera libre, agrupa un número importante de funciones orientadas a facilitar la visualización estática e interactiva de gráficos y tablas asociadas a estadísticas e indicadores institucionales, así como la disposición de algunas funciones auxiliares que facilitan la gestión estadística institucional, sobre todo, a nivel de educación superior. Además, no solo aborda los desafíos específicos en materia de gestión estadística en esta institución, sino que también se posiciona como una opción práctica para la gestión de datos y la consolidación de cifras cuantitativas en otras entidades a nivel nacional e internacional. El paquete, alojado actualmente en GitHub⁹¹, cuenta con una guía de instalación y ejemplos prácticos que ilustran cómo implementar las funciones de visualización gráfica y tabular a través de situaciones reales.

Invitamos a las entidades y los actores interesados en la gestión y disposición de estadísticas e indicadores institucionales a nivel nacional e internacional a hacer uso de este paquete, así como a colaborar y aportar en el desarrollo y crecimiento de este con el fin de contar a futuro con una herramienta tecnológica de acceso libre y construida en comunidad.

CULTURA ESTADÍSTICA (C)

Uno de los aspectos poco trabajados en la producción de cifras estadísticas, pero de máxima importancia, es el fortalecimiento de la cultura estadística al interior de los Estados y sus instituciones. Para lograr este objetivo, es necesario adelantar acciones en diferentes frentes.

En primer lugar, es importante fortalecer el conocimiento y la comprensión de los datos institucionales a nivel directivo y gerencial, así como la información derivada de estos, con el fin de impulsar la toma de decisiones fundamentadas. Los responsables de tomar decisiones deben ser conscientes de la importancia de utilizar datos confiables y actualizados como base para sus acciones estratégicas. En este sentido, es importante

⁹⁰ De hecho, la manera en la que se construyen y visualizan en la web las estadísticas oficiales de esta Universidad es a través del paquete de R en mención. estadisticas.unal.edu.co/menu-principal/cifras-general/cifras-generales/

⁹¹ UnalR. <https://github.com/estadisticaun/UnalR>

brindar capacitación y recursos adecuados que les permitan acceder y comprender la información institucional relevante. De hecho, el conocimiento y uso de la información cuantitativa contenida, por ejemplo, en estadísticas o indicadores debería ser una característica necesaria del nivel directivo, ya que su desconocimiento puede llevar a tomar decisiones basadas en creencias personales en lugar de evidencia sólida.

Otro frente importante es el fortalecimiento del conocimiento sobre la ubicación y disposición de las cifras oficiales. Es fundamental que todos los miembros de las entidades estatales, las instituciones y la sociedad en general conozcan dónde y cómo acceder a las cifras oficiales, ya sea a través de portales web, bases de datos internas u otros medios. Esto facilita la consulta y el uso de la información estadística en diferentes contextos y contribuye a una mayor transparencia y acceso a los datos.

Además, es importante fortalecer las competencias numéricas y la alfabetización estadística de los miembros de las entidades estatales, las instituciones y la sociedad en general. Esto implica mejorar la comprensión de conceptos estadísticos básicos, la interpretación de cifras y la capacidad de evaluar la calidad de los datos. El desarrollo de estas competencias permite una mejor comprensión de las cifras institucionales y una utilización más efectiva de la información cuantitativa en la toma de decisiones, por ejemplo.

Asimismo, es necesario fortalecer las competencias académicas y técnicas de los equipos encargados de la producción estadística. Estos profesionales deben contar con conocimientos actualizados en metodologías estadísticas, herramientas de análisis de datos y buenas prácticas en la gestión estadística. Esto garantiza la calidad y precisión de las estadísticas producidas, así como la aplicación de enfoques y técnicas adecuadas para abordar los desafíos específicos de los países y sus instituciones.

Para llevar a cabo estas acciones se pueden implementar diversas estrategias. Por ejemplo, se pueden utilizar correos institucionales para compartir información relevante sobre datos y estadísticas, así como para fomentar la participación y el intercambio de conocimientos entre los miembros de los Estados y sus instituciones. También se pueden realizar campañas en redes sociales para promover la importancia de la cultura estadística y compartir datos relevantes de manera accesible y atractiva.

Además, se puede disponer de videos educativos que expliquen conceptos estadísticos clave y su aplicación. Esto facilitará el aprendizaje

y la comprensión de la información estadística. De igual manera, se pueden ofrecer cursos tanto virtuales como presenciales para capacitar a los miembros de las entidades estatales, las instituciones y la sociedad en general en temas estadísticos y promover el uso efectivo de los datos.

Es importante también fomentar la participación en conferencias y eventos relacionados con la gestión de datos y estadísticas. Estas instancias brindan oportunidades de aprendizaje, intercambio de experiencias y establecimiento de redes de trabajo colaborativo con otras instituciones y profesionales del campo.

SOPORTES DEL MODELO P2C-T

Hasta ahora hemos explorado cómo la planeación estadística, el establecimiento de conceptos y metodologías orientadoras, la producción estadística enmarcada en un proceso organizado y recurrente, el aprovechamiento de herramientas tecnológicas, y la promoción de una cultura estadística se combinan para establecer un modelo que impulse la consolidación de cifras y el uso adecuado de datos disponibles a nivel estatal e institucional. Sin embargo, la producción estadística requiere de otros elementos que brinden apoyo y sustento a la continuidad del proceso, pues es, sobre todo, una tarea de perseverancia y permanencia, no de velocidad. A continuación, se destacan tres soportes del modelo P2CT.

Apoyo institucional

Las dinámicas en la esfera de lo público se dan dentro de un marco normativo y una estructura administrativa establecida. Para fomentar la gestión de datos en el ámbito de las entidades públicas, es necesario contar con normativas y una estructura administrativa adecuada que facilite su implementación y asegure su continuidad.

Sin embargo, resulta sorprendente que, aunque a lo largo de la historia se haya evidenciado la necesidad de contar con cifras y estadísticas en los Estados e instituciones, la promoción de una institucionalidad que favorezca y permita la gestión de los datos haya sido escasa, e incluso inexistente. Sin embargo, para el caso de la UNAL se pueden destacar

algunos hitos normativos que han favorecido la gestión estadística de la institución.

En primer lugar, se destaca la creación del componente estadístico en lo que anteriormente se conocía como la Oficina Nacional de Planeación, la cual pasó a llamarse Dirección Nacional de Planeación y Estadística a partir de 2013. Este cambio reflejó el reconocimiento de la importancia de la estadística en el ámbito institucional.

En el Acuerdo 113 de 2013 del Consejo Superior Universitario (CSU) (Universidad Nacional de Colombia, 2013), se estableció la estructura interna académico-administrativa a nivel nacional de la UNAL. En el artículo 13 de este acuerdo se asignaron funciones específicas a la Dirección Nacional de Planeación y Estadística, entre las que se encuentran el monitoreo de la ejecución de la planeación y la recopilación de información académica y administrativa para producir datos estadísticos que orienten la gestión de la UNAL y sean reportados a las instituciones competentes.

Otro hito relevante fue la creación de la Oficina Nacional de Estadística, como se establece en la Resolución 383 de 2014 (Universidad Nacional de Colombia, 2014). Esta resolución definió la estructura interna de la Dirección Nacional de Planeación y Estadística, así como las funciones específicas de la Oficina Nacional de Estadística.

Finalmente, se destaca la creación del Régimen de Planeación y de Evaluación Permanente de Resultados de la UNAL, establecido en el Acuerdo 238 de 2017 del CSU (Universidad Nacional de Colombia, 2017). En el capítulo VI de este acuerdo se menciona el Sistema Integrado de Estadísticas e Indicadores como una herramienta fundamental para el monitoreo, el seguimiento y la evaluación de la UNAL. Este sistema, construido en fases, se basa en planes estadísticos institucionales que establecen la visión estadística a mediano y largo plazo, los proyectos requeridos y las fuentes de financiamiento.

Talento humano

El segundo elemento que garantiza el respaldo y la continuidad del modelo se relaciona directamente con las habilidades, los conocimientos y la experiencia de las personas responsables de recolectar, procesar, visualizar y analizar los datos, así como de llevar a cabo los procesos

asociados. Es a través del talento humano que se asegura, en gran medida, la precisión y confiabilidad de las estadísticas generadas.

En un mundo donde las nuevas tendencias están orientadas hacia la gestión de datos, como la estadística, la inteligencia de negocios, la analítica de datos, la ingeniería de datos y el *big data*, es fundamental comprender el propósito detrás del uso de estos. Así, en el campo de la gestión contemporánea de los datos, dependiendo de si se buscan análisis descriptivos, correlacionales, predictivos o prescriptivos, se deben aplicar enfoques disciplinarios y perfiles especializados más adecuados. Por ello es que muchos de los programas educativos actuales⁹² se centran en formar profesionales con conocimientos en matemáticas y estadísticas, habilidades avanzadas en tecnología y programación, así como una comprensión del entorno empresarial y una base teórica sólida en la medición. Sin embargo, en la práctica, se enfrentan a un desafío considerable, ya que es difícil para un único profesional desempeñar todas las tareas y poseer todos los conocimientos requeridos debido a la complejidad y diversidad de competencias exigidas en este campo.

En lugar de depender exclusivamente de un profesional especializado, es preferible conformar equipos de trabajo interdisciplinarios para abordar los desafíos de la gestión de datos. La colaboración entre expertos en diferentes áreas, como matemáticas, estadística, tecnología, programación e, incluso, economía, administración, derecho, antropología, historia o lingüística, permite aprovechar la diversidad de habilidades y perspectivas para obtener resultados más completos y sólidos. Al trabajar en conjunto, estos equipos pueden combinar sus conocimientos y enfoques disciplinarios para enfrentar los retos complejos que surgen en el análisis y uso de los datos. Esto fomenta la sinergia y la complementariedad entre

⁹² En Colombia se puede observar este fenómeno en la promoción de estrategias estatales como los *bootcamps* del Ministerio de Tecnologías de la Información y Comunicaciones (MinTIC), que buscan fomentar la formación en temas relacionados con los datos. Además, instituciones de educación formal como la Universidad Externado de Colombia y la Universidad Javeriana, por ejemplo, han desarrollado programas académicos de pregrado en Ciencia de Datos. Asimismo, existe una amplia oferta de plataformas de educación en línea, como Coursera, Udemy, Edx, DataCamp, Platzi, Digital House y Soy Henry, que ofrecen cursos en esta área.

los profesionales, lo que maximiza el potencial de los datos y su aplicación en diferentes contextos estatales o institucionales.

En este sentido, a lo largo del trabajo desarrollado por la Oficina Nacional de Estadística de la UNAL, se ha contado con el apoyo de diversos profesionales provenientes de diversas disciplinas. Así, se ha trabajado en colaboración con diseñadores gráficos, ingenieros industriales y de computación, expertos en ciencias sociales como la sociología, la antropología, la historia y la lingüística, así como con matemáticos, estadísticos y profesionales de áreas administrativas como la economía, la administración y la contaduría, entre otros. Todavía más, la conformación del presente libro refleja la fusión de dos áreas del conocimiento: las ciencias exactas, como la estadística y la ingeniería, y las humanidades, con la presencia de la lingüística y la sociología. Aunque estas áreas pueden parecer inicialmente distintas en la forma de entender y explicar el mundo, en realidad se complementan mutuamente en este libro para enriquecer la comprensión de la gestión de la información cuantitativa en el ámbito de lo público.

Recursos económicos

Los recursos económicos desempeñan un papel protagónico para el apoyo y sustento del proceso estadístico. La manera en la que los recursos económicos soportan dicho proceso se puede dividir en dos grandes grupos. En primer lugar, dentro de las entidades hay costos fijos necesarios para mantener las operaciones y garantizar la sostenibilidad y regularidad de la producción estadística. Esto incluye, por ejemplo, la disposición de personal especializado; la adquisición, el mantenimiento y la actualización de herramientas tecnológicas para la gestión de datos; los servicios de almacenamiento local o en la nube, bases de datos y difusión de resultados a través de informes, publicaciones y plataformas en línea. Aparte de los costos fijos, es importante contar, en segundo lugar, con recursos destinados a la inversión en innovación, proyectos, mejoras e investigación para estar a la vanguardia en el mundo de los datos.

A partir de la experiencia de la UNAL, desde la creación de la Oficina Nacional de Estadística de la institución, se ha asignado a un reducido grupo de funcionarios (pagados con recursos de funcionamiento), tanto a nivel nacional como en sus sedes, para liderar y mantener el ejercicio

de la actividad estadística. Esta estrategia ha sido clave para sostener la base de la práctica estadística institucional a lo largo de los últimos años. Además, se han destinado recursos para la inversión a través de la formulación y el desarrollo de tres proyectos específicos, incluidos en planes de desarrollo de tres gobiernos institucionales. Estos proyectos de inversión se han diseñado para cumplir con lo establecido, principalmente, en el Plan Estadístico Institucional.

Conclusiones

A lo largo de este libro se exploró cómo el término “información” puede ser empleado en una variedad de contextos, remitiendo a diferentes significados. De hecho, es un término común en las interacciones sociales debido a los avances en telecomunicaciones, computación y el surgimiento del mundo digital. Además, se resaltó que la información puede entenderse como un proceso, como conocimiento y como cosa. Esta última concepción, la de información como cosa, alude a los objetos, procesos, documentos y eventos que pueden servir como evidencia de algo en una situación dada. La información como cosa es manipulable, susceptible de ser procesada, se puede trabajar con ella; por tanto, es sobre la cual se conforman las estadísticas y los indicadores.

También se señaló que la información resulta de bastante utilidad para la gestión pública, entendida esta última como la puesta en marcha de iniciativas, la administración y organización, o el manejo de situaciones problemáticas en el marco de una dimensión social, comunitaria y ética. Asimismo, se destacaron los enfoques de la gestión pública y se subrayó la relación de estos con la información, donde se mencionaron cuatro grandes usos: comunicativo, donde se emplea la información para comunicar mensajes a la comunidad o a un público ideal; de registro, donde se menciona cómo se utiliza la información para describir, representar e inventariar hechos, objetos, actividades y características; valorativo, donde se exploró el uso de la información para establecer juicios sobre el quehacer de las entidades; finalmente, de procesamiento, donde se abordó cómo la información se procesaba para transformarla en nueva información con el propósito de influir en la práctica.

Luego llegó el turno de las cifras estadísticas que, como se sabe, son de gran utilidad para los procesos de gestión, pues proporcionan información necesaria para la toma de decisiones, la evaluación de políticas y la rendición de cuentas. De hecho, el término “estadísticas” proviene de la palabra en latín *statisticus*, que se refiere a lo que es relativo

a un Estado. Esta concepción de las estadísticas como una herramienta de conocimiento que permite describir el Estado y los fenómenos de la sociedad resalta su importancia en la gestión pública.

Con el fin de dar cuenta de las estadísticas como objeto de estudio, se describieron algunas características intrínsecas, es decir, necesarias, y otras extrínsecas, es decir, recomendables. Las características intrínsecas señalan que las estadísticas se basan en poblaciones y muestras; hacen uso de variables cualitativas (nominales u ordinales) y cuantitativas (discretas o continuas); implican la agregación de datos a través de conteos o medidas descriptivas de tendencia central, dispersión, localización y simetría; desagregan los datos en subgrupos temporales, geográficos y temáticos; se representan de manera gráfica y tabular para facilitar su visualización; y, por último, requieren de mecanismos de divulgación y comunicación que permitan darlas a conocer.

Las características extrínsecas, por su parte, advierten de la importancia de ser incluyentes, de tener en cuenta la diversidad de poblaciones y velar por que todas las personas tengan acceso a la información estadística; de ser comparables a nivel institucional, nacional e internacional; de cumplir con las leyes institucionales, estatales o internacionales, así como de tener en cuenta cuestiones éticas; de procurar que la información estadística sea pública, transparente y abierta; de hacer uso intensivo de las TIC; de describir los datos a través de metadatos para reducir malentendidos o tergiversaciones; y, por último, de la importancia de conformar estadísticas a través de un proceso estadístico que abarque diversas fases, desde la detección de necesidades hasta la evaluación.

Después de las estadísticas, el presente libro se centró en los indicadores. Se mostró cómo, desde una perspectiva etimológica, el término “índicador” alude a señalar o mostrar algo. Se explicó cómo, en la práctica, los indicadores se utilizan para focalizar, medir y describir fenómenos. Asimismo, se abordó la dificultad de precisar una tipología para los indicadores y se discutió, entre otras, la tipología de indicadores cuantitativos y cualitativos, donde los primeros se basan en mediciones numéricas, mientras que los segundos se centran en características no numéricas, como percepciones y actitudes. Además, se abordó la tipología de indicadores temáticos, que desempeñan un papel esencial en la medición y evaluación de diversas dimensiones de interés en la sociedad, como la economía, la salud, la educación, el medio ambiente y la demografía,

mientras permiten la comparación y el contraste entre diferentes países, regiones e instituciones, lo que resulta valioso en el ámbito de la toma de decisiones a nivel estatal e institucional.

Después de los indicadores temáticos, se abordó la tipología de indicadores de gestión. Estos siguen una lógica de procesos que pueden ser recurrentes, es decir, que no están acotados en el tiempo, y no recurrentes, que están acotados en el tiempo. Además, los procesos siguen una lógica de estructura de cadena de valor, es decir, en su interior se agrupan en varias tipologías que involucran insumos, actividades, productos y resultados.

En el libro también se destacó cómo a la base de los indicadores hay una naturaleza matemática, es decir que, sin importar su tipología (cuantitativos, temáticos, de gestión o jerarquía), se basan en operaciones matemáticas, que pueden ser simples o complejas. En este contexto, las medidas estadísticas, como las medidas de tendencia central, dispersión, localización y forma, se utilizan para comprender la distribución y las características de los datos; las frecuencias, tanto absolutas como relativas, proporcionan información sobre la cantidad de veces que ocurrió un fenómeno en una población o muestra; las razones permiten comparar cantidades provenientes de diferentes poblaciones o muestras, o diferentes subconjuntos de una misma población o muestra; las tasas, por otro lado, son relaciones que incluyen el factor tiempo, y se aplican en diversos campos, como la demografía y la economía; los índices, por su parte, son indicadores más complejos que a menudo se utilizan para medir fenómenos complejos y multidimensionales, como la calidad de vida.

Ya hacia el final del libro se presentó el denominado Modelo P2C-T, que se enfoca en la conformación de estadísticas e indicadores, así como en la creación de áreas de estadística o analítica de datos dentro de instituciones. Este modelo surge de la experiencia institucional de gestión en la UNAL y consta de cinco dimensiones: planeación estadística (P), conceptualización y metodologías (C), producción estadística (P), disposición y gestión de herramientas tecnológicas (T), y cultura estadística (C). Además, se destacan tres soportes necesarios para que el modelo funcione: apoyo institucional, acceso y disposición del talento humano, y acceso a recursos financieros.

Como hemos visto hasta ahora, el presente libro proporciona a los lectores herramientas para mejorar la calidad de la gestión pública

a través de la reflexión de los conceptos que hay a la base de la conformación de cifras estadísticas e indicadores. Asimismo, destaca cómo una gestión informada y basada en datos, que tenga en cuenta las precisiones teóricas de los conceptos que la componen, puede conducir a mejores procesos, como la toma de decisiones o el seguimiento más preciso de los resultados, que, en última instancia, beneficia a los Estados, sus instituciones y a las comunidades y la sociedad en general.

A continuación, se presentan algunas conclusiones específicas que resultan de este proceso de investigación. En primer lugar, se destaca que la información desempeña un papel fundamental en la conformación de estadísticas e indicadores. Esta concepción se caracteriza por su enfoque cuantitativo y estructurado. No obstante, se abre un espacio para la exploración de cómo otros modos de entender la información, por ejemplo, como conocimiento o como información cualitativa no estructurada, podrían influir en la producción de estadísticas e indicadores, así como en su gestión en el ámbito público.

En segundo lugar, las estadísticas se presentan como una herramienta esencial para la representación y comprensión de los acontecimientos significativos de un Estado. Sin embargo, rara vez se las aborda como un objeto de estudio en el ámbito de la gestión pública. En su lugar, se suelen establecer planes y directrices para la recopilación de datos estadísticos. No obstante, es útil y necesario caracterizar las estadísticas como una entidad independiente.

En tercer lugar, cuando se habla de indicadores es común que su comprensión se limite a los indicadores de gestión, aunque en realidad son considerablemente más diversos. Es esencial comprender y utilizar una variedad de indicadores de manera adecuada en la gestión pública. Además, es importante destacar la utilidad de distinguir entre indicadores y estadísticas, a pesar de que esto no siempre se haga.

En cuarto lugar, el modelo P2C-T se revela como un recurso valioso que refleja la experiencia de la UNAL el cual, aunque se basa en las directrices establecidas a nivel estatal, permite concluir que ningún modelo o propuesta es definitivo, y la capacidad de adaptar y afinar las directrices a las necesidades específicas de una entidad se considera un ejercicio apreciable.

Ahora bien, este libro no solo proporciona un valioso apoyo para quienes trabajan en la gestión pública, sino también involucra cuestiones

sociales subyacentes. En ocasiones, el proceso de investigación adelantado en este documento se concentró, por ejemplo, en reflexiones al respecto de las implicaciones éticas asociadas a la clasificación de poblaciones, inherente a cualquier proceso de generación de datos estadísticos, o en el carácter polisémico de los conceptos relacionados con las estadísticas e indicadores en la gestión pública. De esto se concluye que la temática abordada a lo largo de estas páginas fomenta —e, incluso, exige— la discusión y la reflexión. Por lo tanto, resulta interesante y útil explorar la comprensión de términos como información, gestión, público, estadísticas e indicadores a la luz de varios enfoques y áreas del conocimiento.

Referencias

- Borges, J. L. (1974). El idioma analítico de John Wilkins. En *Obras Completas (1923-1972)* (pp. 706-709). Emecé Editores.
- Bozeman, B. (1998). *La gestión pública: su situación actual*. Fondo de Cultura Económica. <http://biblioteca.ufm.edu/opac/record/89143>
- Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, 42(5), 351-360.
- Castells, M. (1995). *La ciudad informacional: tecnologías de la información, re-estructuración económica y el proceso urbano-regional*. Alianza. <https://dialnet.unirioja.es/servlet/libro?codigo=17506>
- Coneval (2013). *Manual para el diseño y la construcción de indicadores. Instrumentos principales para el monitoreo de programas sociales de México*. https://www.coneval.org.mx/informes/coordinacion/publicaciones%20oficiales/manual_para_el_diseño_y_contrucción_de_indicadores.pdf
- Constitución Política de Colombia [Const]. Artículo 15. Título II. 1991. <https://www.corteconstitucional.gov.co/inicio/Constitucion%20politica%20de%20Colombia%20-%202015.pdf>
- DANE (s. f.a). *Estratificación socioeconómica*. <https://www.dane.gov.co/index.php/servicios-al-ciudadano/servicios-informacion/estratificacion-socioeconomica>
- DANE (s. f.b). *Requisitos de calidad*. <https://www.dane.gov.co/index.php/estadisticas=por=tema/construcciónvivienda-vis-y-no-vis/138-espanol/1900-requisitos-de-calidad>
- DANE (2011). *Introducción al diseño, construcción e interpretación de indicadores*. DANE.
- DANE (2018a). *Metodología de Diagnóstico de los Registros Administrativos para su aprovechamiento estadístico*. DANE, Dirección de Regulación, Planificación, Estandarización y Normalización.
- DANE (2018). *Censo Nacional de Población y vivienda 2018*. DANE. <https://www.dane.gov.co/index.php/estadisticas=por=tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivienda-2018>

- DANE (2020a). *Guía para la inclusión del enfoque diferencial e interseccional*. DANE.
- DANE (2020b). *Lineamientos para el proceso estadístico en el sistema estadístico nacional versión 2.0*. DANE.
- DANE (2020c). *Metodología para el desarrollo de planes estadísticos*. DANE.
- DANE (2020d). *Norma técnica de la calidad del proceso estadístico*. DANE.
- DANE (2020e). *Plan Estadístico Nacional*. DANE. <https://www.sen.gov.co/files/PEN/PEN%202020.pdf>
- DANE (2021). *Sistema de ética estadística*. DANE.
- DNP (2018). *Guía para la construcción y análisis de indicadores*. DNP.
- DNP (2020). *Documento CONPES 3995*. DNP. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3995.pdf>
- Datos abiertos Colombia (s. f.). Divipola - Códigos municipios. <https://www.datos.gov.co/widgets/gdxc-w37w>
- Echeita, G. (2013). *Inclusión y Exclusión Educativa. De Nuevo, “Voz y Quebranto”*. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 11(2), Article 2. <https://doi.org/10.15366/reice2013.11.2.005>
- García Marco, F. J. (1998). El concepto de información: una aproximación transdisciplinaria. *Revista General de Información y Documentación*, 8(1), 303-326.
- González, R. (2021). *Una sociedad mejor informada, es una sociedad más libre*. Unesco. <https://www.unesco.org/es/articles/una-sociedad-mejor-informada-es-una-sociedad-mas-libre>
- Goodrow, C. (s.f.). *You know what's cool? A billion hours*. Blog. Youtube. <https://blog.youtube/news-and-events/you-know-whats-cool-billion-hours/>
- Greenfield, P. (2018, marzo 25). The Cambridge Analytica files: The story so far. *The Guardian*. <https://www.theguardian.com/news/2018/mar/26/the-cambridge-analytica-files-the-story-so-far>
- Gutiérrez, D. (2009). La construcción de indicadores como problema epistemológico. *Cinta de Moebio*, 34, 16-36.
- Huff, D. (2023). *How to Lie with Statistics*. Penguin UK.
- Ingwersen, P. (1992). *Information and Information Science in Context*, 42(2), 99-135. <https://doi.org/10.1515/libr.1992.42.2.99>
- Ley 1581 de 2012. Ley de protección de datos personales (2012). <https://www.funcionpublica.gov.co/eva/gestornORMATIVO/norma.php?i=49981>

- Ley 1712 de 2014 - Ley de transparencia y del derecho de acceso a la información pública nacional (2014). <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=56882>
- Masuda, Y. (1984). *La sociedad informatizada como sociedad post-industrial*. Tecnos. <https://dialnet.unirioja.es/servlet/libro?codigo=235313>
- Mayer-Schönberger, V. y Cukier, K. (2016). *Big data: la revolución de los datos masivos*. Turner.
- Mèlich, J.-C. (2014). *Lógica de la残酷*. Herder Editorial.
- MinTIC (2020). Resolución 1519. Por la cual se definen los estándares y directrices para publicar la información señalada en la Ley 1712 del 2014 y se definen los requisitos materia de acceso a la información pública, accesibilidad web, seguridad digital, y datos abiertos. https://normograma.mintic.gov.co/mintic/docs/resolucion_mintic_1519_2020.htm
- Moore, N. (1997). La sociedad de la información. En *Informe mundial sobre la información* (pp. 287-300). Unesco. <https://unesdoc.unesco.org/ark:/48223/pf0000110221>
- Muñoz, P. (2020). La gestión pública: de los modelos al territorio. En F. Sánchez y N. Liendo, *Manual de ciencia política y relaciones internacionales* (pp. 95-124). Universidad Sergio Arboleda.
- Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web is Changing what we Read and How we Think*. Penguin.
- Platón (1998). Libro VII. En *República* (pp. 338-348). Gredos.
- Programa de las Naciones Unidas para el Desarrollo (PNUD) (s. f.). *Objetivo 17: Alianza para lograr los objetivos. Objetivos de Desarrollo Sostenible*. <https://www.undp.org/es/sustainable-development-goals/alianza-para-lograr-los-objetivos>
- Rendón Rojas, M. A. (2017). *El concepto de información desde una óptica de la filosofía de la Bibliotecología y Estudios de la Información. Significados e interpretaciones de la información desde el usuario*. UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información, 29-76.
- Rodríguez, A. y Bernal, E. (2019). *Gestión de la información cuantitativa en las universidades: pistas para su abordaje en la era de la sobreinformación*. Universidad Nacional de Colombia. https://estadisticaun.github.io/L_Conceptual/index.html
- Rodríguez, P. E. (2012). *Historia de la información*. Capital Intelectual. <https://ri.conicet.gov.ar/handle/11336/187340>

- Sarrionandia, G. E. (2013). Inclusión y exclusión educativa. De nuevo “voz y quebranto”. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 11(2), 99-118.
- Shannon, C. E. y Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.
- Universidad Nacional de Colombia (2013, 8 de octubre). Acuerdo 113. Por el cual se establece la estructura interna académico-administrativa del Nivel Nacional de la Universidad Nacional de Colombia. Acta 17. http://www.legal.unal.edu.co/rlunal/home/doc.jsp?d_i=58926
- Universidad Nacional de Colombia (2014, 08 de abril). Resolución 383. Por la cual se establece la estructura interna de la Dirección Nacional de Planeación y Estadística, y se determinan sus funciones. http://www.legal.unal.edu.co/rlunal/home/doc.jsp?d_i=64488
- Universidad Nacional de Colombia (2017, 28 de marzo). Acuerdo 238. Por medio del cual se adopta el Régimen de Planeación y de Evaluación Permanente de Resultados de la Universidad Nacional de Colombia. Acta 04. http://www.legal.unal.edu.co/rlunal/home/doc.jsp?d_i=88708
- Universidad Nacional de Colombia (2018). *Plan Estadístico Institucional*. <https://estadisticaun.github.io/PlanE/>
- Universidad Nacional de Colombia (s. f.). Guía para la cuantificación, medición y seguimiento a la gestión de los procesos. https://unal.softexpert.com/se/generic/gn_linkshare/linkshare.php?key=1vpood2gmv2s&cdisosystem=73
- Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59, 1-23. <https://doi.org/10.18637/jss.v059.i10>
- Wiener, N. (1961). *Cybernetics or Control and Communication in the Animal and the Machine*. MIT Press.
- Wolf, C., Joye, D., Smith, T. W. y Fu, Y. (2016). Non-probability Sampling. En *The SAGE Handbook of Survey Methodology* (pp. 329-345). SAGE.
- Zuboff, S. (2020). *La era del capitalismo de la vigilancia: la lucha por un futuro humano frente a las nuevas fronteras del poder*. Paidós.

Estadísticas e Indicadores en la Gestión Pública

Aproximación teórica y práctica

En esta publicación se utilizaron caracteres
Perlu y Addington, papel Bond de 90 g para las páginas internas
y papel Propalmate de 240 g para la cubierta.