

DOCUMENTOS DE
TRABAJO SOBRE
**ECONOMÍA
REGIONAL
Y URBANA**

Categorías municipales en
Colombia: Avanzando hacia un
modelo de descentralización
asimétrica

Por:
Karina Acosta
Yuri Reina-Aranza

Núm. 321
Diciembre, 2023



Centro de Estudios Económicos
Regionales (CEER) - Cartagena

Categorías municipales en Colombia:

Avanzando hacia un modelo de descentralización asimétrica*

Karina Acosta[‡]

kacostor@banrep.gov.co

Yuri Reina-Aranza[♦]

yreinaar@banrep.gov.co

La serie **Documentos de Trabajo Sobre Economía Regional y Urbana** es una publicación del Banco de la República en Cartagena. Las opiniones contenidas en este documento son de exclusiva responsabilidad de los autores y no comprometen al Banco de la República ni a su Junta Directiva.

Resumen

La descentralización asimétrica se ha convertido en una discusión relevante en países en desarrollo como Colombia, donde las capacidades entre los gobiernos subnacionales difieren dramáticamente. Aunque estas discusiones han estado presentes desde la década de los sesenta, el éxito de la descentralización es aún una discusión relevante, donde la clasificación del territorio es fundamental. Este documento discute los posibles obstáculos de las clases territoriales usadas actualmente en Colombia. Asimismo, propone y discute la utilidad del uso de algoritmos recientes de la literatura de aprendizaje de máquinas no supervisados para la clasificación de los territorios subnacionales. Específicamente, este documento implementa *Clustering via Optimal Trees* (ICOT), un algoritmo que permite clasificar los territorios y, a la vez, identificar las reglas de las clases definidas. Este estudio además propone la creación de diferentes tipologías territoriales conforme a sus usos.

Palabras clave: descentralización, capacidades, resultados, aprendizaje de máquinas, ICOT, Colombia.

JEL: C45, R11, E62.

*Las autoras agradecen la valiosa ayuda de Jack Dunn en la instalación de la licencia académica de ICOT, así como al grupo Interpretable AI por el acceso a la licencia. Asimismo, agradecemos los comentarios recibidos en el 62nd European Regional Science Association Congress, la misión de descentralización del DPN, Jaime Bonet, Juliana Jaramillo y Luis Armando Galvis.

[‡] Investigadora junior del Centro de Estudios Económicos Regionales CEER – Sucursal Cartagena.

[♦] Profesional del Centro Regional de Estudios Económicos CREE – Sucursal Cartagena.

Municipal categories in Colombia:

Moving towards an asymmetric decentralization

Karina Acosta

kacostor@banrep.gov.co

Yuri Reina-Aranza

yreinaar@banrep.gov.co

The series **Documentos de Trabajo Sobre Economía Regional y Urbana** is a publication of Banco de la República in Cartagena. The opinions contained in this document are the sole responsibility of the authors and do not commit Banco de la República or its Board of Directors.

Abstract

Asymmetric decentralization has become relevant in developing countries like Colombia, where capabilities between subnational governments differ dramatically. Although these discussions have been present since the 1960s, the success of decentralization is still a relevant discussion in these contexts, where the classification of the territory is fundamental. This document discusses the possible obstacles of the territorial classes currently used in Colombia. Likewise, it proposes and discusses the usefulness of using recent algorithms from the unsupervised machine learning literature to classify subnational territories. Specifically, this document implements Clustering via Optimal Trees (ICOT), an algorithm that allows us to classify the territories and identify the rules underneath the defined classes. This study also proposes the existence of different territorial typologies according to their uses.

Keywords: decentralization, capabilities, results, unsupervised learning, ICOT, Colombia.

JEL: C45, R11, E62.

1. Introducción

La pertinencia de la descentralización, entendida como la transferencia de poderes y recursos de funciones de gobiernos centrales a locales, ha sido una discusión cardinal en la búsqueda de la igualdad en la eficiencia y distribución de los servicios y bienes públicos a lo largo del territorio. La literatura sobre descentralización sugiere que esta no es intrínsecamente beneficiosa ya que existen ventajas y desventajas de la gobernanza multinivel (OCDE, 2019). Las experiencias de los países hacia nuevos modelos de descentralización advierten que su efectividad depende de las características de cada país y, por lo tanto, se deben ajustar a las peculiaridades de cada territorio (Allain-Dupré, 2018; OCDE, 2019). Asimismo, una creciente literatura propone modelos de descentralización asimétrica al interior de cada país, donde las asimetrías conceptualizan tratos diferenciados a los gobiernos locales según sus capacidades (Tan & Avshalom-Uster, 2021). Aunque en principio las propuestas de descentralización favorecían una estructura regional, en la evidencia más reciente se hace más relevante la figura municipal (OCDE, 2019). Los procesos de aglomeración y urbanización han generado una cada vez más marcada diferenciación intrarregional que le otorga una mayor relevancia al papel de las ciudades principales.

Aunque Colombia tiene una prolongada historia de descentralización, cuyos rastros se manifiestan desde la década de 1960, un punto de inflexión trascendental se dio con la constitución de 1991, desde donde se comenzaron a otorgar importantes poderes y recursos a gobiernos subnacionales (Bonet et al., 2023). Como parte de este proceso de descentralización, se han propuesto y usado en Colombia unas categorizaciones del territorio con la finalidad de definir transferencias de recursos y responsabilidades territoriales. Entre las principales agrupaciones del territorio utilizadas en Colombia están las definidas en la Ley 617 de 2000, Ley 1551 de 2012, así como en ejercicios como los del Departamento Nacional de Planeación – DNP (2015, 2016). Luego de revisar las propuestas de agrupaciones, argumentamos que tales propuestas presentan fallas que pueden no estar alineadas con los objetivos de la descentralización municipal. En las propuestas identificamos amplias heterogeneidades dentro de grupos específicos definidos, carencia de interpretabilidad en las agrupaciones y una posible discordancia entre los objetivos de la descentralización y las clases de agrupaciones territoriales propuestas.

En esta investigación proponemos dos alternativas de categorizaciones municipales: tipologías por capacidades y por resultados; cada una de estas clases, potencialmente, con usos diferentes. En comparación con los estudios preliminares y los diseños actuales de los modelos municipales, este estudio utiliza avances recientes en aprendizaje de máquinas no supervisado (*clustering*) para definir grupos de municipios cuyas características son explícitamente definidas. Adicionalmente, este documento propone una reagrupación de las clases de municipios basadas en la naturaleza de las variables consideradas en combinación con métodos de reducción de la dimensionalidad, con lo cual se espera que se superen las limitaciones de las tipologías contempladas en la norma vigente. Los datos utilizados se limitan al 2019, pero la metodología se puede fácilmente expandir a otros años y contextos.

Este trabajo se divide en seis secciones. La segunda resume la evolución de la categorización municipal en Colombia, identifica sus propósitos y analiza su funcionalidad actual. El tercer capítulo revisa los ejercicios de categorización en países de América Latina, con el fin de identificar algunas referencias de variables y métodos usados. La cuarta parte incluye un análisis de la distribución de los municipios por las categorías vigentes en Colombia y sus principales características. La quinta sección desarrolla el ejercicio de categorización con nuevas variables propuestas y, posteriormente, el último capítulo realiza algunas reflexiones finales.

2. Evolución de la categorización municipal en Colombia

La primera referencia que se tiene sobre la categorización de municipios en Colombia está en el Acto Legislativo N° 1 de 1968 que reformó la Constitución Política de 1886. El artículo 63 del Acto Legislativo, además de permitir autónomamente la asociación de municipios dentro de un mismo departamento para mejorar la prestación de servicios públicos, indica que “La ley podrá establecer diversas categorías de municipios de acuerdo con su población, recursos fiscales e importancia económica, y señalar distinto régimen para su administración” (Acto Legislativo 1 de 1968).

Pero fue solo hasta 1988, y con el fin de determinar los límites de los salarios para los alcaldes, cuando por el Decreto 222 se crearon cinco categorías de municipios: distritos especiales, municipios mayores, intermedios, menores y básicos (Perea Gómez & Delúquez

Medina, 2020). Esta clasificación estaba determinada por el Índice de Categorización Municipal que tenía en cuenta tres variables: (1) La población municipal, de acuerdo con el último censo recolectado por el Departamento Administrativo Nacional de Estadística (DANE); (2) Los recursos fiscales certificados por la entidad de control fiscal; y (3) El índice de importancia económica certificado por el DNP. Este último podría ser calculado con base en variables como el promedio anual de los ingresos brutos declarados por industriales y comerciantes del municipio, el flujo anual de carga, los ingresos operacionales de las entidades financieras, el avalúo catastral y el consumo de energía, entre otros (Decreto 222 de 1988).

Sin embargo, fue a partir de la Constitución de 1991 cuando se empezó a fortalecer el proceso de descentralización y a dar mayor autonomía a los municipios. El artículo 320 autorizó la creación de categorías municipales, con base no solo en la población, recursos fiscales e importancia económica, sino que además incluyó la situación geográfica. Así mismo, avaló señalar un régimen distinto para estas categorías en organización, gobierno y administración, lo que implicaba una asignación diferenciada de competencias y funciones (Constitución Política de Colombia, 1991; Perea & Delúquez, 2020).

De acuerdo con la revisión realizada por Carrillo (2022), a partir de lo consignado en la Constitución de 1991, varias leyes han avanzado en el establecimiento de las categorías municipales. En primera instancia, la Ley 136 de 1994 estableció normas para modernizar la organización y el funcionamiento de los municipios, dentro de las cuales incluyó la distribución de competencias, funciones y la categorización de municipios. Esta última se hizo con base en la población y los ingresos corrientes de libre destinación (ICLD)¹ y dio como resultado siete categorías (Ley 136 de 1994; Carrillo, 2022).

Años después, esta categorización fue reformada con la Ley 617 de 2000, introduciendo cambios en los rangos poblacionales y de ingresos para las categorías 4, 5 y 6, pero sin modificar el número de categorías. Además, estableció que los municipios fronterizos con un

¹ ICDL son ingresos corrientes recaudados por las entidades territoriales (municipios o departamentos), que no se afectan por ninguna Ley o Acto Administrativo (e.j. Acuerdo Municipal u Ordenanza Departamental), lo cual le da libertad a las entidades para financiar actividades o sectores específicos (Contraloría General de la República, s.f.).

número de habitantes mayor a 70.000 se clasificarán mínimo en categoría 4 y sus gastos de funcionamiento no podrían superar el 100% de los ICLD. De igual manera, incluyó una norma de categorización para los municipios aledaños al Distrito Capital que tuvieran más de 30.001 habitantes, de manera que deberían ser clasificados en categoría 2 (Ley 136 de 1994; Carrillo, 2022).

Posteriormente, la Ley 1551 de 2012 realizó algunos ajustes a la categorización definida por la Ley 617, sin modificar el número de categorías ni los rangos de población establecidos en cada una. Los ajustes se concentraron en la adición de un indicador de importancia económica² dentro de cada categoría y el agrupamiento de las categorías así: i) Grandes municipios: pertenecientes a las categorías 1 y especial; ii) Municipios intermedios: clasificados en las categorías 2, 3 y 4; y iii) Municipios básicos: agrupados en las categorías 5 y 6. Frente a los grupos, el parágrafo 6 del artículo 7 de la Ley 1551 señala que el desarrollo de distinto régimen de organización, gobierno y administración para los municipios pertenecientes a cada grupo se haría en un plazo de dos años, desde la vigencia de esta ley. No obstante, de acuerdo con Carrillo (2022), las disposiciones acerca de los tres grupos de municipios no han sido desarrolladas ni llevadas a la práctica.

Finalmente, el Decreto 2106 de 2019 suprimió el uso del indicador de importancia económica para la categorización municipal, dejando tanto las categorías como los tres grupos sin cambios y bajó al 80% el límite de los gastos de funcionamiento sobre los ICLD para los municipios de frontera (Carrillo, 2022; Decreto 2106, 2019).

² Este indicador corresponde al peso relativo del PIB de cada municipio dentro de su departamento y es calculado por el DANE.

Cuadro 1. Evolución de la categorización de municipios en Colombia

	Decreto 222 de 1988	Ley 136 de 1994	Ley 617 de 2000	Ley 1551 de 2012
Categorías	- Distritos especiales - Municipios mayores - Municipios intermedios - Municipios menores - Municipios básicos	- Especial - Primera - Segunda - Tercera - Cuarta - Quinta - Sexta	- Especial - Primera - Segunda - Tercera - Cuarta - Quinta - Sexta	- Grandes municipios: especial y primera - Municipios intermedios: segunda, tercera y cuarta - Municipios básicos: quinta y sexta
Variables	- Población - Recursos fiscales certificados por la entidad de control - Índice de importancia	- Población - Ingresos corrientes de libre destinación (ICDL)	- Población - ICDL	- Población - ICDL - Índice de importancia Económica
Otras consideraciones		Municipios fronterizos >70.000 habitantes clasificarán mínimo en categoría 4	Municipios fronterizos >70.000 habitantes clasificarán mínimo en categoría 4 Municipios aledaños al distrito capital > 30.000 habitantes serán categoría 2	Municipios fronterizos >70.000 habitantes clasificarán mínimo en categoría 4

Fuente: Elaboración de las autoras con base en los decretos y leyes incluidos.

Así entonces, la categorización municipal vigente es la incluida en la Ley 1551 de 2012, pero sin el uso del indicador de importancia económica y sin el desarrollo ni implicaciones de política para los tres grupos establecidos. Esta categorización está determinada por los ICLD en términos de salarios mínimos mensuales legales vigentes (SMMLV) y la población, estableciendo siete categorías (Cuadro 2).

Paralelamente, hay una serie de normas que rigen la clasificación vigente. Específicamente, si un municipio según su población debe estar en una categoría distinta a la que estaría por el monto de sus ICLD, entonces se clasificará en la categoría correspondiente a sus ICLD. En la misma línea, si el municipio supera el límite de los gastos de funcionamiento como proporción de su ICLD, este municipio se clasificará en una categoría menor a la que estaba, independientemente de la categoría en la que se clasifique según población e ingresos.

Además, ningún municipio podrá avanzar o descender más de dos categorías de un año a otro. Por su parte, se mantiene la norma de que los municipios de frontera con población mayor a 70 mil habitantes serán clasificados como mínimo en cuarta categoría (Ley 1551, 2012, Decreto 2106 de 2019).

Cuadro 2. Categorías municipales vigentes en Colombia

	Tipo de municipio	Población (número de habitantes)	ICLD anuales (SMLV)
Especial	Grandes	≥ 500.001	Más de 400.000
Primera	Grandes	≥ 100.001 y ≤ 500.000	≥ 100.000 y ≤ 400.000
Segunda	Intermedios	≥ 50.001 y ≤ 100.000	≥ 50.000 y ≤ 100.000
Tercera	Intermedios	≥ 30.001 y ≤ 50.000	≥ 30.000 y ≤ 50.000
Cuarta	Intermedios	≥ 20.001 y ≤ 30.000	≥ 25.000 y ≤ 30.000
Quinta	Básicos	≥ 10.001 y ≤ 20.000	≥ 15.000 y ≤ 25.000
Sexta	Básicos	≤ 10.000	≤ 15.000

Fuente: Ley 1551 de 2012 y Decreto 2106 de 2019.

Para acercarnos a los objetivos o funciones de la categorización, la Ley 1551 de 2012 resalta la posibilidad que tienen los municipios de asumir atribuciones, funciones y competencias voluntarias, adicionales a las asignadas por ley, de acuerdo con sus capacidades fiscales, administrativas y técnicas. Además, resalta que, según su tipología, el Gobierno Nacional podrá delegar atribuciones y funciones a cada municipio. Sin embargo, esto aún no se ha llevado a la práctica, sino que el efecto de la categorización actual se ha limitado a las funciones como: establecer los salarios de los alcaldes y el valor máximo de los gastos de funcionamiento como proporción del ICDL, así como delimitar el gasto de transporte de concejales.

En este contexto, se entiende que la aplicación de la categorización municipal en Colombia se ha limitado a funciones organizativas y de racionalización del gasto (Ley 617 de 2000). En relación a la racionalización del gasto, se ha evidenciado un impacto positivo y exitoso en la disciplina fiscal subnacional (Pérez-Valbuena, *et al.*, 2021). Sin embargo, el artículo 320 de la Constitución Política de 1991 le otorga un mayor potencial de uso, que posiblemente no ha sido aprovechado completamente, por ejemplo, en asignación de

competencias y recursos diferenciados. Las categorías actuales no son utilizadas en la asignación de competencias voluntarias, como lo son las certificaciones en educación y salud. El proceso para asignar estas certificaciones es completamente independiente y para ello se genera una clasificación diferente, porque evalúa variables enfocadas en los sectores específicos.

Entre las críticas a las categorías municipales vigentes en Colombia se distinguen el débil impacto que han tenido en el proceso de descentralización y la insuficiencia de su actualización. Las clases vigentes dejan de lado especificidades de los territorios que son fundamentales para determinar las necesidades y capacidades del ente territorial. El uso exclusivo de dos variables (ICDL y población) oculta las heterogeneidades relevantes para la asignación de recursos y poderes a entidades subnacionales. Para ilustrar, las categorías 5 y 6, que incluyen municipios de menor población e ingresos, concentra más del 90% de los municipios (Hernández, 2015; Carrillo, 2022; Perea & Delúquez, 2020).

En los últimos años se han realizado algunos ejercicios que han dado como resultado otros tipos de categorización o clasificación de municipios, los cuales invitan a repensar las categorías municipales vigentes de forma que se pueda cumplir el objetivo constitucional.

Uno de estos ejercicios fue el elaborado por el DNP en 2015, el cual estableció siete componentes en el ejercicio de tipologías. Estos componentes fueron: funcionalidad urbano-regional, dinámica económica, calidad de vida, ambiental, seguridad e institucional. En total se incluyeron 18 variables para el ejercicio municipal, enmarcadas en los componentes mencionados. Entre las variables están: población rural, población total, índice de pobreza multidimensional, valor agregado municipal, disparidades económicas, penetración de internet, área de bosques naturales, desempeño fiscal, inversión en el sector ambiental, homicidios, secuestros y hurtos, entre otras. Utilizando el método de componentes principales llegaron a la conformación de siete categorías organizadas, a su vez, en tres grupos: desarrollo robusto, desarrollo intermedio y desarrollo incipiente. Con este ejercicio se logró establecer grupos más homogéneos en su interior, pero diferentes entre ellos, además de que ninguna categoría concentra más del 27% de los municipios (DNP, 2015).

Otro ejercicio es el del DNP (2016) que corresponde a la metodología para la nueva medición de desarrollo municipal (MDM). El primer paso dentro de esta nueva metodología es la “Definición de grupos de municipios según un indicador que se aproxime a las "dotaciones iniciales" para hacer la medición al interior de grupos comparables y tener en cuenta la heterogeneidad del país en términos de capacidades” (DNP, 2016). Para esto se incluyen tres variables: ingresos tributarios per cápita, densidad empresarial y ruralidad, esta última compuesta por el tamaño de la población, proporción de población en cabecera, densidad poblacional y pertenencia al sistema de ciudades. Las variables se estandarizaron con valores entre 0 y 1, donde los municipios con más bajas dotaciones iniciales tendrán un valor 0 y los de mejores dotaciones iniciales tendrán un valor de 1. De este ejercicio resultaron seis grupos, el primero con las 13 ciudades principales y que corresponde al grupo con mayores dotaciones iniciales en los aspectos evaluados. Los otros grupos fueron llamados: nivel alto, nivel medio alto, nivel medio, nivel medio bajo y nivel bajo.

3. Tipologías municipales en países de América Latina

Esta sección contiene una revisión de experiencias de clasificación de entidades territoriales, en varios países de Latinoamérica, con el fin de identificar potenciales variables y metodologías utilizadas para alimentar esta investigación.

En Chile existe un ejercicio de clasificación de entidades territoriales para la distribución del Fondo de Incentivo al Mejoramiento de la Gestión Municipal (FIGEM). Este fondo tiene como objetivo fomentar los progresos en la administración municipal otorgando recursos que el municipio puede autónomamente destinar a inversión (Ministerio del Interior y Seguridad Pública de Chile, 2015). Con el fin de determinar los municipios en los cuales se deben focalizar los recursos, el fondo utiliza una tipología comunal que agrupa los municipios de acuerdo con las características o parámetros estructurales que determinan las condiciones en las que cada municipalidad realiza sus funciones. Para determinar las tipologías se utilizan variables de condiciones de vida, agrupadas en dos dimensiones: demográfico y socioeconómico. La primera es la dimensión eje demográfico territorial que incluye el tamaño poblacional y el número de predios no agrícolas en cada comuna, la dispersión de la

población calculada a partir de la densidad poblacional, el nivel de ruralidad y la entropía de núcleos poblados. A su turno, la jerarquía político administrativa es determinada por la condición de ser capital regional o provincial y una ponderación por población y el tipo de localidad, el cual depende de un puntaje asignado a cada comuna, de acuerdo al Ministerio de Vivienda y Urbanismo (Ministerio del Interior y Seguridad Pública de Chile, 2015).

La segunda dimensión es la del eje socioeconómico, el cual incluye variables como el patrimonio comunal, determinado con el avalúo promedio de inmuebles de la comuna y el recaudo per cápita de patentes municipales³. El capital humano, estimado a partir del nivel de escolaridad promedio de la población, el promedio ponderado del puntaje en la prueba de selección universitaria y el porcentaje de analfabetismo. La tercera variable es de características socioeconómicas y se construye con el porcentaje de pobreza y el promedio por hogar del ingreso monetario (Ministerio del Interior y Seguridad Pública de Chile, 2015).

Con estas variables y dimensiones se establecen cinco grupos: 1. Grandes comunas metropolitanas con alto y medio desarrollo; 2. Comunas mayores, con desarrollo medio; 3. Comunas urbanas medias, con desarrollo medio; 4. Comunas semi urbanas y rurales con desarrollo medio; y 5. Comunas semi urbanas y rurales con bajo desarrollo. Con esto, la distribución de los recursos FIGEM se concentrará en los grupos de menor desarrollo (Ministerio del Interior y Seguridad Pública de Chile, 2015).

Para el caso de Bolivia, existen cuatro categorías de municipios, distribuidos únicamente por criterio poblacional (Biblioteca Virtual de la Cooperación Alemana BIVICA, s.f). Sin embargo, Ayavari y Alarcón (2014) realizaron un ejercicio de tipologías aplicando técnicas multivariantes de análisis de clúster. El objetivo era lograr una clasificación municipal en grupos homogéneos a partir de características demográficas y socioeconómicas comunes de los municipios. Las variables que utilizaron fueron: tamaño de la población, grado de urbanidad (porcentaje de población urbana) y nivel de pobreza (población con Necesidades Básicas Insatisfechas -NBI-). Una vez aplicada la metodología de agrupamiento de tipo jerárquico aglomerativo (Algoritmo Howard – Harris), los autores obtuvieron seis grupos:

³ Una patente municipal es el permiso que otorga cada municipalidad para los negocios y es un elemento necesario para emprender una actividad comercial en determinado lugar. Existen patentes comerciales, profesionales, industriales y de alcoholes. <https://www.bcn.cl/leyfacil/recurso/patentes-municipales>

ciudades metropolitanas, urbanos, urbanos pobres, rural grande, rural pobre y rural pobre extremo.

En Argentina, Brasil y Ecuador no hay evidencia del uso de tipologías municipales o de ejercicios de clasificación de entidades territoriales. En la mayoría de los casos, la asignación de recursos (transferencias) se hace por población. En Brasil, además, se tiene en cuenta el ingreso per cápita, y en Ecuador se adiciona la medida de necesidades básicas insatisfechas, indicadores de calidad de vida, esfuerzo fiscal y administrativo, y cumplimiento de metas de planes de desarrollo (Presidencia de la República del Ecuador, 2010; Vivanco & Cordero, 2019; Ley XV – N°5 (antes Ley 257), s.f).

Por su parte, aunque en Costa Rica la distribución de transferencias también se basa en la población, Morales (2010) realizó un ejercicio para segmentar los cantones en tres grupos diferenciados. Para esto, construyó un indicador con las siguientes variables: tamaño poblacional del cantón, densidad poblacional, peso de los ingresos propios de cada gobierno local en el agregado nacional y en los ingresos fiscales de cada grupo. Como resultado se obtuvieron tres grupos, el primero compuesto por 14 cantones que fueron las diez municipalidades con mayor puntuación en el indicador y los cantones principales de las provincias; el segundo con 17 municipalidades, las cuales lograron un indicador por encima de la media nacional, y el tercero con 50 municipalidades que estuvieron Por debajo de la media nacional.

En Perú existe una tipología municipal para la distribución del programa de incentivos a la mejora de la gestión municipal, para la cual se usaron dos criterios: (1) sociales, demográficos y geoeconómicos, y (2) número de viviendas urbanas. El primer criterio fue utilizado para clasificar las ciudades principales a través del cálculo de un índice de potencialidades que suma la población, el monto del impuesto predial por vivienda y el Índice de Desarrollo Humano. De allí se seleccionaron las 50 municipalidades con un mayor resultado según el índice, de dónde se escogieron 40 que además estaban ubicadas en centros de desarrollo y dinamizantes locales. Estas 40 hacen parte del primer grupo, por lo que el segundo corresponde a los 209 restantes. Por su parte, los otros municipios que no son ciudades principales se dividieron en dos grupos de acuerdo al número de viviendas urbanas (más de 500 y menos de 500) (Ministerio de Economía y Finanzas de Perú, 2022).

En México los municipios están clasificados por población en cuatro categorías: rurales, semiurbanos, urbanos y metropolitanos. Orozco (1993) diseñó un método que parte de esta clasificación y le agrega otras variables socioeconómicas, para no basar la clasificación en un solo criterio. El método consiste en incluir 18 condicionantes (variables) y a cada una asignarle una ponderación y un puntaje. Las variables que tiene en cuenta son: población, déficit de vivienda, déficit de agua potable, cobertura de drenaje, cobertura de energía eléctrica, servicios médicos, ingreso promedio familiar, salario, población económicamente activa por rama de actividad, sistema de comunicaciones y transporte, indicadores de actividad industrial y educación. El trabajo de Orozco (1993) se concentró en proponer un método de clasificación de los municipios, pero no mostró resultados en cuanto rangos y categorías. Por otro lado, Pérez (2014) elaboró una tipología municipal por asentamiento humano utilizando correlaciones con variables estadísticas de población, infraestructura, servicios públicos y desempeño económicos, entre otros. Las categorías resultantes de este ejercicio fueron cinco: metropolitanos, urbanos, en proceso de transformación de lo rural a lo urbano, rurales y marginados.

Por su parte, para Uruguay, Nion y Feo (2017) utilizaron el análisis de clúster o conglomerados para crear los grupos de municipios. Las variables usadas están enmarcadas en tres dimensiones, socioeconómicas, demográficas y territoriales. Dentro de las primeras está el promedio de años de educación, el porcentaje de hogares con al menos una necesidad insatisfecha y tasa de actividad femenina. En la dimensión demográfica se incluyó el índice de envejecimiento, la población y el porcentaje de población que vive en área rural. Finalmente, dentro de las territoriales se usó la superficie en kilómetros cuadrados, la distancia entre la capital del departamento y la localidad principal del municipio, número de localidades censales y densidad de la población. Con este ejercicio, los 112 municipios fueron clasificados en 11 grupos, donde el grupo 1 se caracteriza tener municipios densamente poblados, pequeños en superficie y mayor nivel socioeconómico y educativo, mientras que el grupo 11 se concentran municipios de gran extensión, con un tamaño de población medio, alejados de la capital del departamento y con bajos niveles socioeconómicos.

4. Análisis de las categorías municipales vigentes en Colombia

La mayoría de los municipios colombianos se clasifican en las categorías 5 y 6 por criterios de población e ingresos, mientras que las otras categorías reúnen un menor número de municipios (Cuadro 3). En la categoría especial están las principales ciudades del país. Bogotá, capital del Colombia, es la que tiene un mayor número de habitantes y más altos ingresos corrientes de libre destinación (ICLD). Le siguen Medellín, Cali, Barranquilla y Cartagena. Aunque todas cumplen los requisitos para estar clasificadas en esta categoría, hay importantes diferencias dentro del grupo y entre ellas.

Cuadro 3. Municipios por categoría, 2022

Categorías	Número de municipios	Participación (%)
Especial	5	0,5
1	27	2,5
2	18	1,6
3	19	1,7
4	24	2,2
5	39	3,5
6	970	88,0
Total	1102	100,0

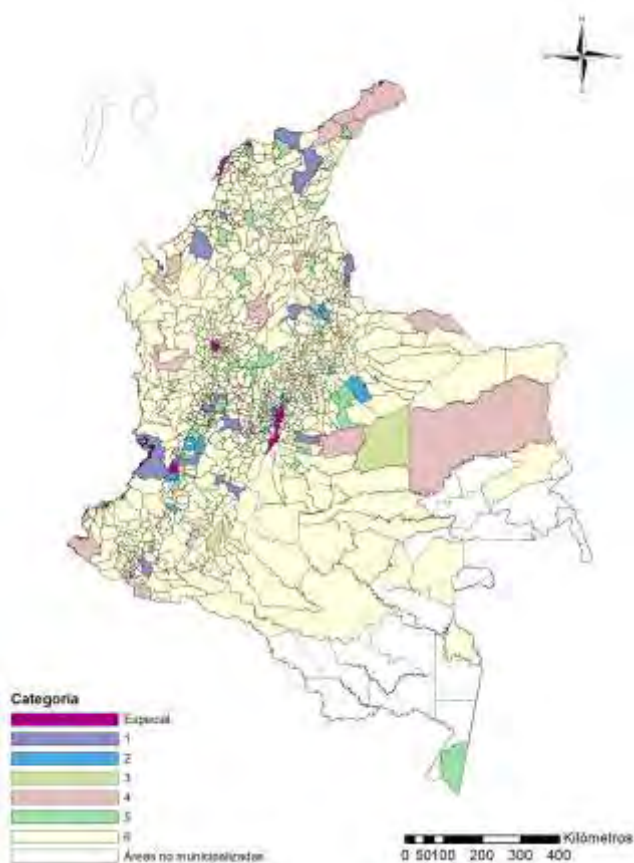
Fuente: Contaduría General de la Nación con base a la Ley 1551 de 2012 y sus ajustes posteriores con el Decreto – Ley 2106 de 2019.

En la categoría especial, compuesta por Bogotá, Cartagena, Barranquilla, Medellín y Cali en 2022, se observa que Bogotá tiene 7,5 y 6,5 veces más habitantes que Cartagena y Barranquilla, respectivamente. Adicionalmente, los ICLD de Cartagena y Barranquilla representaron solamente el 6,6% y 11,4% de los de Bogotá, respectivamente. También hay diferencias con Medellín y Cali, aunque no tan amplias. En general, las cifras de Bogotá se ubican lejos de las del resto de ciudades de su grupo. Cabe recordar que entre un año y otro los municipios pueden cambiar de categorías, de acuerdo con sus ICLD, por ejemplo, en 2021 Bucaramanga hacía parte de la Categoría especial, pero en 2022 estaba en la primera.

En 2022, la categoría 1 reunió la mayoría de las ciudades capitales excluidas del grupo especial, como Pasto, Armenia, Cúcuta, Ibagué, Valledupar y Tunja, entre otras, y municipios grandes como Rionegro, Bello, Soacha, Chía, Yumbo y Buenaventura. La categoría 2 incluyó

otros municipios capitales como Popayán y Sincelejo, y municipios intermedios como Tuluá, Jamundí, Piedecuesta y Cajicá. En la categoría 3 y 4 se destacan capitales como Florencia, Riohacha y Quibdó, y municipios como Tumaco, Malambo y Duitama. Por su parte, las categorías 5 y 6 reúnen los pequeños municipios, donde se agrupa la mayoría de estos (Figura 1).

Figura 1. Distribución geográfica de los municipios por categoría



Fuente: elaboración de las autoras con base en Contaduría General de la Nación.
Nota: áreas no municipalizadas son áreas que no tienen definida una clase.

Para conocer qué tantas diferencias hay al interior de cada grupo se calculó la desviación estándar de cada variable (ICLD y población) en cada categoría. Los resultados muestran mayor dispersión en las categorías especial, primera y segunda, mientras que la categoría 6 es la de menor dispersión en todas las variables que se tienen en cuenta para la categorización (Cuadro 4). La mayor dispersión en la categoría especial es explicada porque en este grupo está la capital (Bogotá) que tiene amplias diferencias sobre el resto de las ciudades en el

mismo grupo como Cartagena, Barranquilla, Cali y Medellín. En la primera categoría hay municipios con ICLD tres veces mayores a otros como Bucaramanga, Envigado y Pereira que están en el límite superior de los ingresos, y contrastan marcadamente con Armenia, Mosquera y Pasto. En esta categoría también está incluida Buenaventura.

Cuadro 4. Desviación estándar por variable, 2022

Categorías	Población (personas)	ICLD (miles de pesos)	Gastos Funcionamiento (miles de pesos)
Especial	2.455.195	2.492.578.362	689.971.319
1	193.910	59.863.053	21.119.637
2	137.185	16.822.239	7.683.935
3	43.336	6.408.641	4.495.445
4	60.067	8.692.051	5.036.811
5	33.378	4.223.209	2.804.755
6	14.920	2.356.564	1.307.264

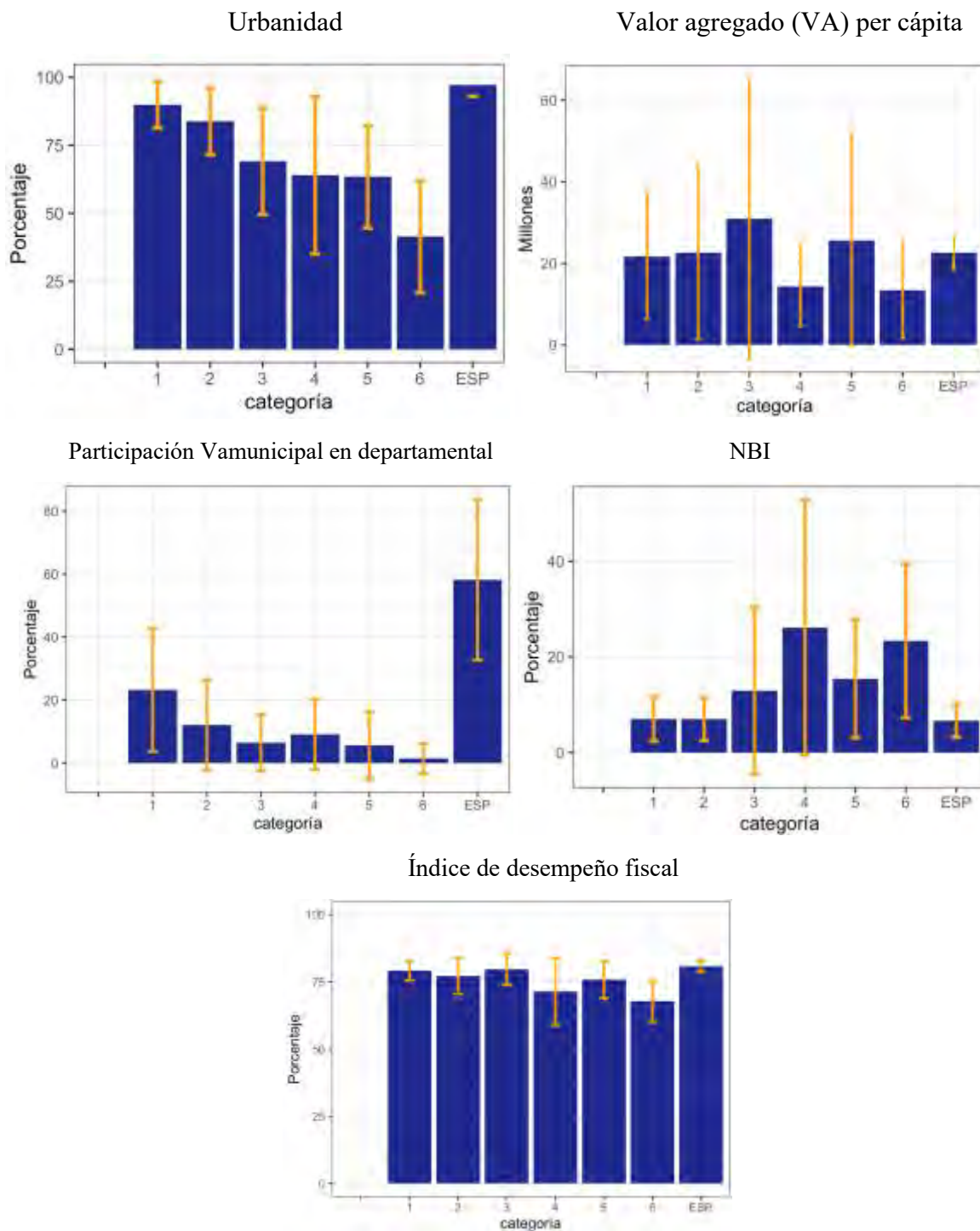
Fuente: elaboración de las autoras con base en DANE y Contaduría General de la Nación.

Estos resultados indican que la diferencia en la cantidad o proporción de municipios dentro cada categoría no es un indicador de una distribución necesariamente errónea. El grupo con mayor número de municipios no es el más disperso, al menos, en lo referente a población e ICLD. En este estudio proponemos que la discusión sobre tipologías no sólo se limite a estas variables ni a estas clases. Sería deseable tener en cuenta otras variables socioeconómicas y clases de territorios subnacionales, además de las ya definidas, con el fin de robustecer la categorización y su uso. Estos ejercicios posiblemente podrían generar otras distribuciones con más o menos grupos.

Como ejercicio preliminar a la propuesta que se planteará en la siguiente sección, se muestran algunas estadísticas descriptivas que permiten observar otras características de los municipios incluidos en la categorización vigente. Para esto, se escogieron cinco variables relacionadas con los aspectos que deben regir la clasificación de municipios de acuerdo con la Ley 1551 que son: población, ICLD, importancia económica y situación geográfica. Las variables son valor agregado per cápita, participación del valor agregado municipal en el total departamental, porcentaje de población con NBI, porcentaje de población urbana e indicador de desempeño fiscal. Las categorías incluidas en este análisis y las variables de valor

agregado, población e índice de desempeño fiscal corresponden al 2021, el NBI es del año 2018.

Figura 2. Promedio de las variables por categoría



Fuente: DANE, DNP, Contaduría General de la Nación. Cálculos de las autoras. Nota: las barras naranjas representan la desviación estándar (aproximación de la dispersión de las variables).

Como se observa en la Figura 2, los municipios de las categorías especial, 1 y 2 son los de mayor proporción de población urbana, mientras que los de sexta categoría tienen una población mayoritariamente rural. La participación del valor agregado municipal en el departamental también es mucho mayor en los municipios de categoría especial, lo que es coherente, dado que este grupo está compuesto por las principales capitales del país. La diferencia con el resto de categorías es alta, mientras que, en promedio, un municipio de categoría especial aportó el 60% del valor agregado de su departamento, uno de categoría 1 aportó aproximadamente el 20% y el resto, mucho menos de eso. Los de categoría 6 contribuyeron en promedio solamente con el 1,4%.

En cuanto al valor agregado per cápita, sorprende que el mayor promedio no se encuentre en la categoría especial, sino en la categoría 3, en donde el valor agregado per cápita de municipios como Puerto Gaitán (Meta) que fue cinco veces más que el de Bogotá, y otros como Acacias (Meta), La Jagua de Ibirico (Cesar), Tenjo (Cundinamarca) y Girardota (Antioquia) superan el de algunas de las principales ciudades del país. No se evidencian grandes diferencias entre las primeras categorías (especial, 1 y 2), pero sí entre la quinta y la sexta, dado que, aunque ambas reúnen municipios pequeños, la quinta tiene un promedio de valor agregado per cápita que duplica la categoría 6 y supera el de resto de categorías (excepto la 3). Los municipios de categoría 6 que mayor valor agregado per cápita tienen son Castilla La Nueva (Meta), Nobsa (Boyacá) y Tauramena (Casanare).

Por su parte, el porcentaje de población con NBI es mayor en las categorías 4 y 6. En la cuarta categoría los municipios con mayor NBI son Uribia (La Guajira) con un NBI de 88,7%, Cumaribo en Vichada (87,5%) y Manaure en La Guajira (81,1%). Mientras que en el mismo grupo están La Calera (Cundinamarca), Marinilla (Antioquia) y Cartago (Valle del Cauca) con cifras de 3,1%; 5,1% y 5,4%. En la categoría 6 también hay municipios con muy alto NBI como San Jacinto y Clemencia en Bolívar con 93,4% y 84,9% e Istmina en Chocó con 84,0% y otros con cifras bajas como Subachoque y Tabio en Cundinamarca (3,7% y 3,9%).

Por último, en cuanto al Índice de Desempeño Fiscal (IDF) no se observan marcadas diferencias, aunque los mejores resultados se concentran en las categorías especial, seguido de cerca por 3 y 1, indicando que estos municipios tienen una mejor gestión fiscal, con unas finanzas más saludables. Por su parte, las categorías 4 y 6 tienen resultados más deficientes,

lo que podría relacionarse con alta dependencia de las transferencias, de acuerdo a la descripción de los rangos del IDF.

5. Métodos

La clusterización (*clustering*) o agrupamiento consiste en el conjunto de métodos no supervisados para separar observaciones, patrones o características de objetos en clases. Sus usos se han extendido a múltiples áreas del conocimiento, incluyendo economía, marketing, medicina, entre otros (Bertsimas et al., 2021; Fraiman et al. 2013). El problema de la clusterización, a diferencia del aprendizaje de máquinas supervisado, es la identificación de las clases que asemejan unos objetos sin tener una variable objetivo evidente. El fin es generar particiones de unos datos con base en sus características, para hacer valoraciones de los mismos. En efecto, este es el problema al que se enfrenta la tipología territorial y, por ello, este estudio emplea métodos de clusterización.

Los métodos para la generación de clústeres no son nuevos, sus orígenes se remontan a los años sesenta con la propuesta *K-means* (MacQueen 1967). Aunque se han generado múltiples alternativas de algoritmos para generar clústeres (métodos jerárquicos, basado en modelos, basado en densidades, entre otros (Hastie et al., 2001)), el *K-means* se ha establecido como el método tradicional para el agrupamiento de datos, por lo que utilizaremos como punto de referencia su versión corregida, *K-means ++* (Arthur and Vassilvitskii, 2007), así como *K-medoids*, debido a la presencia de variables que no siguen una distribución normal (Struyf et al., 1997). Pese a los avances en los algoritmos para el agrupamiento de datos, la utilidad y fácil implementación de algoritmos como *K-means*, se han tenido pocos progresos en la interpretabilidad de las particiones resultantes y cómo evaluarlas (Doshi-Velez y Kim, 2017). En el contexto de este estudio, entendemos por interpretabilidad la habilidad para explicar explícitamente el agrupamiento de las unidades subnacionales en una clase específica.

Existen varias propuestas para la generación de agrupamientos interpretables. Dentro de las opciones en la literatura seleccionamos el algoritmo Clusterización Interpretable vía Árboles Óptimos (ICOT, por sus siglas en inglés) (Bertsimas, Orfanoudaki y Wiberg, 2021), que permite generar grupos interpretables con modelos basados en árboles y que, además, posibilitan la incorporación de variables numéricas y categóricas. Bertsimas et al. (2021)

demuestran que el algoritmo propuesto tiene un rendimiento superior a métodos no-supervisados en dos etapas utilizando métricas de pérdida. Asimismo, basado en criterios de validación interna, estos autores muestran que la calidad de los resultados de ICOT es comparable con la de otras técnicas modernas de clusterización como las basadas en modelos, jerárquicas o por densidades. Sin embargo, en comparación con estas estrategias de agrupación modernas, ICOT permite la interpretabilidad de los clústeres como resultado del mismo proceso de partición. Por las ventajas de ICOT frente a los algoritmos existentes, este ha sido seleccionado como el algoritmo base de esta investigación; trasladando su uso a las clasificaciones territoriales.

ICOT busca maximizar una métrica de calidad de los clústeres o minimizar la función de pérdida. Intuitivamente, la función de pérdida es una cuantificación de la diferencia entre cada observación (municipio) y las observaciones en el clúster más cercano corregido por una medida de distancia⁴. Cabe destacar que el concepto de distancia no hace referencia a distancias geográficas, sino al espacio multidimensional de características. Específicamente, ICOT busca maximizar la métrica de silueta o el índice Dunn de toda la partición de los datos, sujeto a una serie de restricciones que reproducen optimización con *modern mixed-integer optimization* (MIO). A diferencia de algoritmos de árboles de decisión pasados, este algoritmo permite que cada división tenga completa información de todas las otras particiones (Bertsimas y Dunn, 2019). El resultado final de ICOT es un árbol de particiones de datos, donde cada hoja representará un clúster definido por el algoritmo. Para esta aplicación se utilizó el índice de Dunn como función de pérdida. Este se define como la separación mínima inter-clúster y la separación máxima intra-clústeres.

$$DI = \frac{\min_{1 < j \leq m} \delta(C_i, C_j)}{\max_{1 < k \leq m} \Delta_k}$$

Donde la distancia entre clústeres se define por $\delta(C_i, C_j)$ y la distancia máxima de un clúster C se define por Δ_k . Superiores niveles de DI se asocian con mejores particiones de los datos, lo cual se puede interpretar como una mayor compactibilidad de las clases generadas. En

⁴ Una descripción detallada del funcionamiento de ICOT se encuentra en (Bertsimas, Orfanoudaki y Wiberg, 2021). Las estimaciones se obtuvieron usando Julia versión 1.1.0 y la licencia de Interpretable AI.

otras palabras, tener distancias más grandes entre clústeres, definida por el numerador, es indicador de mejores agrupaciones.

La implementación de ICOT requiere una serie de elecciones para su configuración inicial. Para este estudio utilizamos la métrica de Dunn como función de pérdida⁵. Adicionalmente, se establecieron los resultados del K-means ++ y K-medoids como punto de partida del algoritmo (*warm start*) para obtener estimaciones más eficientes. Finalmente, siguiendo la sugerencia de los desarrolladores del algoritmo, dejamos el nivel de complejidad en su valor predeterminado.

El algoritmo de ICOT ofrece dos alternativas de iniciación (*warm start*) que permite hacer las estimaciones más eficientes. Esto es, se puede iniciar en clústeres previamente definidos, lo cual mejora los tiempos de estimación. Para ello se utilizaron dos estrategias. En el caso de las tipologías por capacidades se establecieron los resultados de K-medoids como punto de iniciación de ICOT, debido a la presencia de variables que no siguen una distribución normal. Para las estimaciones de las tipologías por resultados, se utilizaron los resultados del algoritmo K-means ++ como opción de iniciación.

Entre algunas alternativas encontradas en la literatura con propuestas de tipologías territoriales, encontramos el uso de métodos como K-means estándar en combinación con Análisis por Componentes Principales (ACP). Si bien es cierto que las combinaciones de estos métodos tienen algunas ventajas en la agrupación de datos, como la reducción de la dimensionalidad de problemas complejos (e.j. múltiples variables), ACP requiere la creación de variables sintéticas a partir de la información disponible, lo que genera limitaciones adicionales en los análisis de las clases formadas a partir de la clusterización. Adicionalmente, esta reducción de la dimensionalidad presupone limitaciones en la implementación de políticas públicas, especialmente asociadas con transparencia.

⁵ Alternativamente, se puede usar el Índice de silueta. Las diferencias en los resultados no cambian significativamente.

5.1. Tipologías basadas en usos

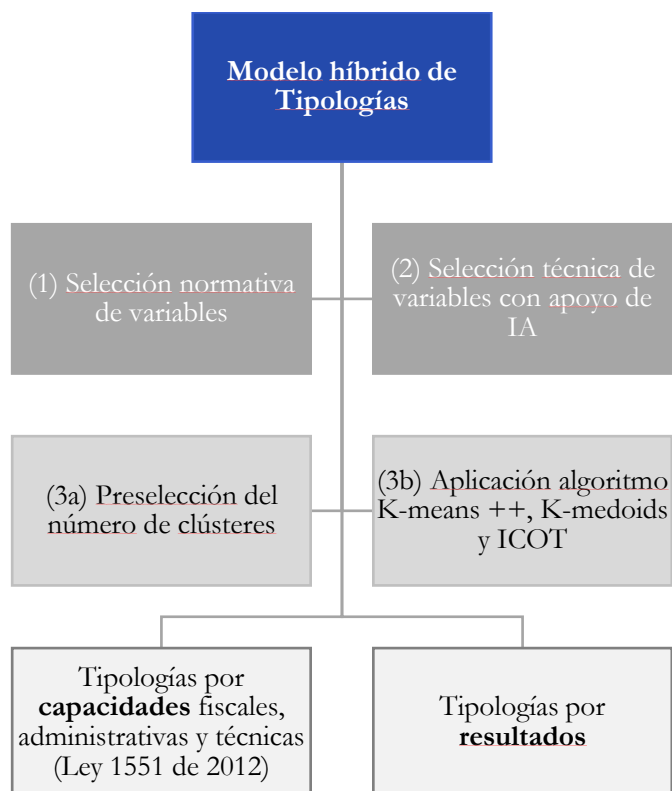
Además de proponer nuevos algoritmos para la definición de las clases, este documento plantea una discusión sobre la necesidad de crear diferentes agrupaciones de municipios con base en los objetivos mismos de las tipologías. En los estudios sobre clusterización del territorio identificamos la combinación de variables, indistintamente de su naturaleza, y no necesariamente están alineadas con lo demandado por la norma. Alternativamente, proponemos la creación de clústeres con base en dos tipos de características: capacidades y resultados. Las tipologías por capacidades se inspiran en la Ley 1551 de 2012 y buscan identificar clases de municipios según sus capacidades fiscales, administrativas y técnicas. Las clases por capacidades pueden usarse como referencia para la delimitación de autonomías por competencias que puedan adquirir los municipios. Por su parte, la propuesta basada en resultados está dirigida a la identificación de grupos de municipios con ventajas y deficiencias en su función esencial: suministrar bienes y servicios públicos a la población. La identificación de grupos por resultados puede ser útil para la generación de políticas de intervención nacional dirigidas a áreas específicas.

El esquema propuesto para la identificación de clases consta de tres pasos (Figura 3). En primer lugar, se emplea lo que denominaremos un criterio de identificación. Este consiste en la selección normativa de variables, limitada por su disponibilidad. Para este ejercicio identificamos un conjunto de variables de la literatura de las clasificaciones territoriales en Colombia y otros países. En un segundo paso, estas variables se delimitaron por su relevancia, disponibilidad y temporalidad. Se prescindió de variables con múltiples valores faltantes, con baja frecuencia de recolección (ej. censo de establecimientos económicos), y que no fueran relevantes para una clase específica, además de que tuvieran problemas de comparabilidad entre municipios (ej. Hurto a instituciones financieras). Asimismo, en el caso de las variables por capacidades, hicimos uso de bosques aleatorios para reducir la dimensionalidad del problema y permitir una mayor eficiencia en las estimaciones del algoritmo objetivo.⁶ Este paso estaría seguido del criterio de clasificación, que consiste en la

⁶ Estimamos un bosque aleatorio, donde la variable dependiente la determinamos como el ICLD. Delimitamos las variables por el nivel de importancia resultante. Previamente, excluimos una de las variables con altas correlaciones, debido a posibles problemas de los árboles aleatorios (Genuer *et al.*, 2010). Los bosques

categorización de la información con base en su facultad para definir capacidades o resultados.

Figura 3. Modelo híbrido de categorización municipal



Fuente: elaboración de las autoras.

En el último paso se emplean los algoritmos de *K-means ++*, *K-medoids* e *ICOT* como se definió anteriormente. Los algoritmos *K-means++* y *K-medoids* constan de un paso preliminar, la identificación del número óptimo de clústeres, para la cual empleamos conjuntamente el método de Elbow y Silueta. Este ejercicio se aplica a cada grupo de variables: capacidades y resultados, independientemente.

Entre la información usada para este estudio encontramos algunos valores faltantes. En el caso de que el municipio careciera completamente de la información, estas observaciones

aleatorios definen un algoritmo que combina múltiples árboles en un único resultado. Para una explicación de bosques aleatorios, revisar Biau y Scornet (2016).

fueron excluidas. Cuando se tenían pocos valores faltantes, utilizamos tres medidas de imputación de información: *predicting mean matching* (pmm), clasificación y árbol de decisión (cart) y lasso⁷. Entre estas opciones, optamos por pmm para definir los valores faltantes debido a que para nuestra base de datos se obtuvieron mejores predicciones.

6. Datos

En 2019, Colombia contaba de 1123 municipios. Para este conjunto de divisiones territoriales, obtuvimos un grupo de variables para las clases de capacidades y resultados. Las variables utilizadas se describen en los Cuadros 5 y 6. Aquellas incluidas en las clases por desempeño fiscal siguen las recomendaciones del DNP para evaluar el desempeño fiscal agregado (primer componente del Cuadro 5).

La principal fuente de información de todas las variables es TerriData (Sistema de Estadísticas Territoriales de Colombia)⁸, una plataforma que compila información territorial de diferentes fuentes. Las variables identificadas a partir de la literatura y no disponible en TerriData fueron estimadas a partir de múltiples fuentes, entre estas variables se encuentran: (1) el porcentaje de tierras con catastro se estimó a partir de información suministrada por el departamento de descentralización del DNP, (2) el número de años de educación promedio de las personas con 15 años o más se calculó con información del Censo poblacional de 2018, (3) el índice de penetración de internet se obtuvo del Ministerio de Tecnologías de la Información y Comunicaciones de Colombia (MinTIC), (4) la distancia al mercado más cercano se calculó como la distancia euclidiana del centroide de cada municipio al centroide de los municipios cabecera del país, (5) los resultados de las pruebas escolares estandarizadas se estimaron como el promedio simple municipal de las evaluaciones aplicadas por el Instituto Colombiano de la Evaluación de la Educación (ICFES) a todos los estudiantes de colegios en el último grado de escuela secundaria, y (6) el ICDL se obtuvo de la Contraloría General de la Nación.

⁷ Las estimaciones se realizaron usando el paquete *mice* en R.

⁸ <https://terridata.dnp.gov.co/index-app.html#/acercade>

Para efecto de las estimaciones por capacidades, Bogotá se excluyó de la muestra debido a que representaba un valor atípico muy diferente al resto de unidades de análisis. No obstante, este sí se incluyó en la clase definida por el árbol de decisión que se ajusta a sus características.

Cuadro 5. Variables de tipologías por capacidades

Fiscales
Indicador de desempeño fiscal. Componentes:
Porcentaje de ingresos corrientes destinados a funcionamiento respaldo de la deuda
Porcentaje de ingresos que corresponden a transferencias
Porcentaje de ingresos corrientes que corresponden a recursos propios
Porcentaje de gasto total destinado a inversión
Capacidad de ahorro
Ingresos tributarios/ ICDL
Ingresos tributarios per cápita/ICDL per cápita
Transferencias nacionales per cápita
Administrativas
Población total
Índice de urbanidad del municipio
Total de unidades económicas
Porcentaje de tierras con catastro
Técnicas
Índice de penetración a internet
Años de educación en personas de 15 y más
Distancia al mercado más cercano

Nota: Elaboración de las autoras.

Cuadro 6. Variables de tipologías por resultados

Fiscales
Indicador de desempeño fiscal. Componentes:
Educativas
Cobertura neta de preescolar
Cobertura neta de primaria
Deserción escolar
Resultado de pruebas estandarizadas
Condiciones de vida
Índice de pobreza multidimensional
Índice de calidad el agua
Tasa de homicidios
Salud
Tasa de afiliaciones al sistema de salud

Nota: Elaboración de las autoras.

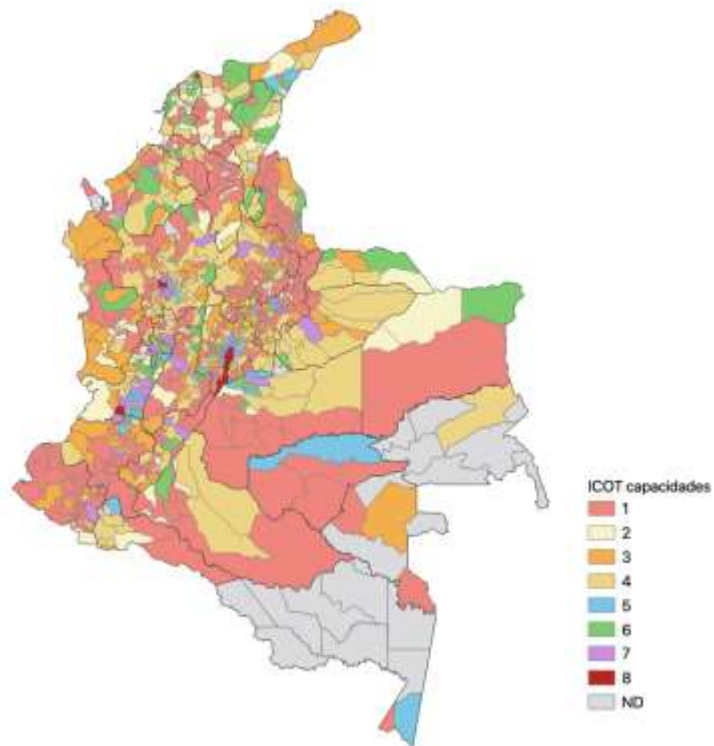
7. Resultados

Las tipologías por resultados tienen tres características distintivas cuando se comparan con las clases municipales vigentes en la Ley. En primer lugar, del grupo de variables utilizadas, las variables seleccionadas por el modelo como relevantes en la clasificación fueron: el desempeño fiscal, el nivel de urbanidad, la penetración de internet, el tamaño de la población y la educación. Estas muestran un conjunto de características y variabilidad que trascienden las observadas en las clases actuales: ICDL y población. En segundo lugar, a diferencia de las clases actuales, ICOT identifica la alta heterogeneidad observada en la agrupación de la categoría especial. En este caso se distinguen sólo cuatro áreas en la clase especial: Cali, Medellín, Barranquilla y Bogotá, cuatro de los municipios principales de Colombia.

Por último, como se ha descrito, ICOT tiene la ventaja de poder describir explícitamente las características de las clases creadas, evidentes en el árbol de decisión a partir del cual se crean las particiones (Figura 5). De este se concluye que entre los componentes fundamentales para clasificar el territorio con base en sus capacidades multidimensionales son sus niveles de urbanidad y su desempeño fiscal. Los resultados muestran que la mayor parte de municipios de Colombia son rurales y tienen un relativamente bajo desempeño fiscal (clúster 1). Cabe resaltar que el número de clústeres finales corresponden a los nodos terminales de los árboles (Figuras 4 y 5) que para las clases por capacidades corresponden a 8 nodos. Estos números no están pre definidos, como en el caso de los algoritmos de la familia K, sino que dependen de la profundidad del árbol.

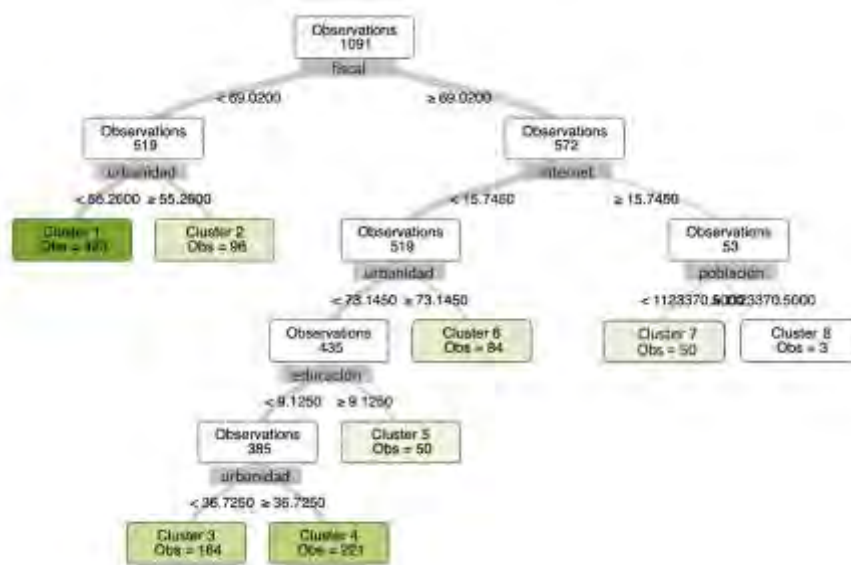
Simulando el ejercicio de la sección 4 para el caso de los resultados de este documento, utilizamos las mismas variables para analizar su comportamiento entre los grupos generados por la clasificación basada en resultados (Anexo A1). A diferencia de lo que se observa en las clases municipales vigentes, los cambios en las categorías (de menor a mayor) reflejan una tendencia relativamente constante para casi todas las variables. Por ejemplo, en la medida en que se avanza de la categoría 1 a la 8, se observa un crecimiento gradual del promedio del valor agregado per cápita.

Figura 4. Tipologías por capacidades



Fuente: Estimaciones de las autoras. Nota: ND son áreas sin información.

Figura 5. Árbol de decisión para tipologías por capacidades

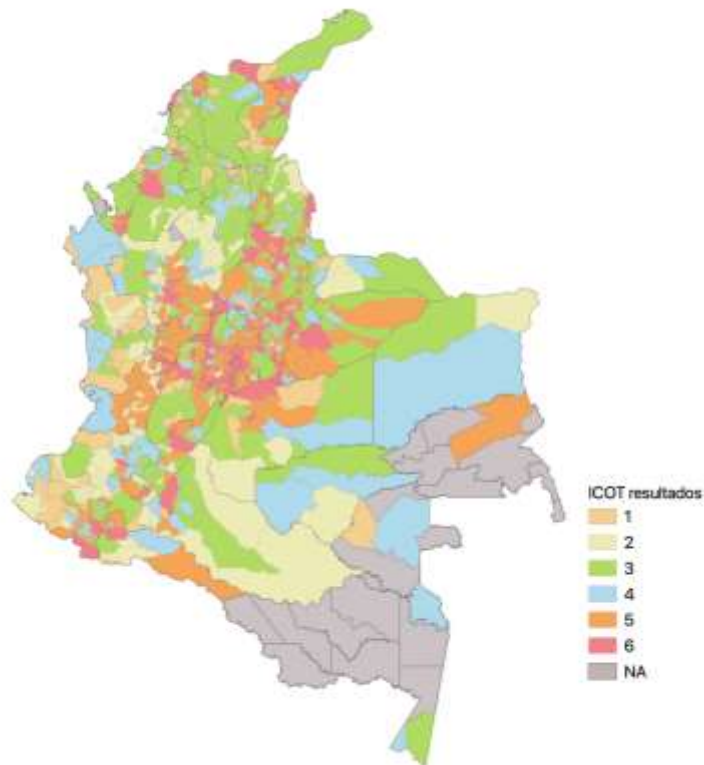


Fuente: Elaboración de las autoras.

Aunque variables como el valor agregado per cápita, la participación del valor agregado municipal en el departamental y el NBI no fueron utilizadas en el árbol de decisión final, las tendencias suaves de la media a través de las categorías de estas variables proveen elementos para evaluar el desempeño de los modelos y las variables utilizadas. Específicamente, el cambio relativamente suave entre las categorías para variables no consideradas nos puede confirmar que la presencia de más variables en los modelos no es fundamentalmente necesaria, ya que las incluidas en el análisis pueden indirectamente incorporar otras características no observadas en el proceso de clasificación. Adicionalmente, aunque se pueden notar grandes variabilidades dentro del valor agregado per cápita de algunas categorías como la 4 y 5, esto se debe a la presencia de municipios mineros dentro de estas clases.

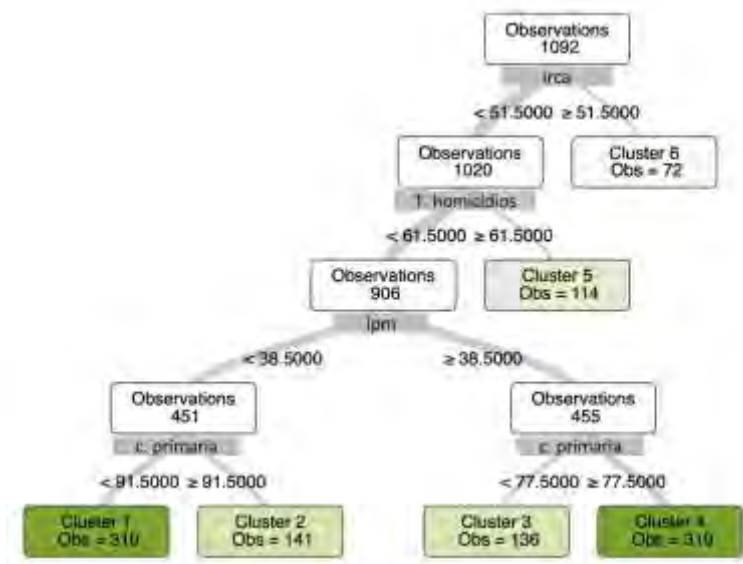
Los resultados por desempeño tienen una distribución ligeramente diferente a los obtenidos por las tipologías de capacidades. Consideramos que parte de esta diferencia se debe a que las variables incluidas en resultados están altamente correlacionadas con indicadores de desarrollo de largo plazo. El mapa que obtenemos por la tipología basada en resultados se asemeja a los de la céntrica e histórica concentración del desarrollo económica en Colombia y a la presencia de instituciones históricas extractivas en las regiones (Cepeda y Meisel, 2013) (Figura 6). Según el árbol de decisión resultante, los clústeres ubicados en la parte central de Colombia (3 y 4) se caracterizan por tener relativamente bajos niveles de pobreza multidimensional y una comparativamente alta cobertura de educación (Figura 7). Por ello, sugerimos que la descentralización territorial esté orientada hacia tipologías que consideren variables de capacidades en lugar de variables por resultados. Asignar responsabilidades con base en tipologías por resultados puede desconocer resultados vinculados a características que trascienden variables con efectos inmediatos de políticas de iniciativa local. En su lugar, las tipologías por resultados pueden considerarse para iniciar puentes de intervención entre el gobierno nacional y los gobiernos locales, especialmente en áreas con menor desarrollo.

Figura 6. Tipologías por resultados



Fuente: Estimaciones de las autoras. Nota: NA son áreas sin información.

Figura 7. Árbol de decisión para tipologías por resultados



Fuente: Elaboración de las autoras. Nota: irca: índice de riesgo de la calidad del agua.

Este documento también se ocupa de evaluar si existe correspondencia entre los clústeres identificados en el grupo de capacidades y resultados. El Cuadro 7 resume las correspondencias entre las clases definidas por el algoritmo para capacidades y resultados. Cada columna suma 100 y corresponde a la proporción de municipios de cada categoría de capacidades en la de resultados. Este cuadro exhibe dos patrones importantes. En primer lugar, en la medida en que nos desplazamos de menores a mayores clases de capacidades, asociadas con mejores indicadores, se observa una mayor concentración de intersecciones con las clases 5 y 6 de clases por resultados. Asimismo, se distingue una mayor variabilidad en la asignación de los municipios en la clase 1 de capacidades a la clase de resultados. No obstante, una mayor densidad de este grupo se acumula en el grupo 3 de la clase por resultados. La presencia de similitudes y discrepancias entre los dos grupos generados pueden aportar elementos de juicio para considerar la importancia de la separación de las variables con base en su naturaleza para la generación de clases territoriales.

Cuadro 7. Tabla de distribución de municipios de categoría capacidades en resultados

6	3	7	7	12	36	33	66	67
5	18	25	30	41	42	38	30	33
4	20	7	20	4	2	0	0	0
3	38	41	27	27	8	11	0	0
2	10	9	7	13	12	15	4	0
1	10	10	7	2	0	2	0	0
	1	2	3	4	5	6	7	8
	Clases de capacidades							

Fuente: Estimaciones de las autoras.

8. Discusión

Este documento estudia las clasificaciones territoriales actuales de Colombia y sugiere dos nuevas clasificaciones de municipios con base en la naturaleza de las variables consideradas. El análisis de las clases municipales vigentes sugiere que estas cumplen, aparentemente, uno de los propósitos de la Ley: control de los gastos y deuda públicos locales. Sin embargo, los ejercicios realizados en este estudio se direccionan hacia la necesidad de crear nuevas agrupaciones municipales orientadas hacia el propósito de descentralización asimétrica, como se discute en la actualidad para países como Colombia, donde el proceso de descentralización no muestra evidencia de la convergencia de indicadores territoriales (OCDE, 2019).

Este estudio sugiere que, en principio, cualquier ejercicio tendrá un componente subjetivo asociado a la necesaria selección de variables, el número de clústeres (en algunas aplicaciones) y el algoritmo. No obstante, independientemente de la estrategia utilizada, proponemos que se consideren cuatro componentes principales en la generación de las tipologías territoriales cuyo objetivo sea la descentralización. En primer lugar, es fundamental generar clústeres interpretables o con transparencia en la asignación de las clases, para lo cual el uso de algoritmos recientes como ICOT pueden cobrar importancia. El uso de componentes principales o algoritmos convencionales pueden generar clústeres adecuados, pero estos tendrán limitantes en la práctica.

Segundo, las variables consideradas en la generación de los clústeres deben estar alineadas con un propósito claro. En la literatura sobre clasificación subnacional encontramos propuestas que combinan variables que pueden no medir la capacidad de los territorios, objetivo clave de las clases territoriales. En tercer lugar, se deben considerar variables con la mayor disponibilidad y comparabilidad a través del espacio. Por ejemplo, variables como tasa de hurtos a instituciones financieras, disponibles en la información de Colombia, no fue utilizada debido a que no es posible la comparabilidad entre municipios urbanos y rurales o esta medida indirectamente por el nivel de urbanidad. Por último, los ejercicios de este documento sugieren que la generación de clústeres más rigurosos no necesariamente está

asociada con una mayor inclusión de variables. En contraste, los resultados sugieren que modelos más parsimoniosos tienen a generar clústeres de mejor calidad.

Bibliografía

- Allain-Dupré D (2018) Assigning responsibilities across levels of government: Trends, challenges and guidelines for policy-makers. *OECD Working Papers on Fiscal Federalism no. 24*, 27 September. Paris: OECD Publishing.
- Anja Struyf, Mia Hubert, Peter J. Rousseeuw, (Feb. 1997), Clustering in an Object-Oriented Environment, *Journal of Statistical Software*, Vol 1, Issue 4.
- Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. In Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms (pp 1027–1035). *Society for Industrial and Applied Mathematics*.
- Acto Legislativo 1 de 1968. *Por el cual se reforma la Constitución Política de Colombia*. Colombia: Sistema Único de Información Normativa. Recuperado el 02 de 02 de 2023, de https://www.suin-juriscol.gov.co/viewDocument.asp?id=1825156#ver_1825220
- Ayavari Nina, D., & Alarcón Lorenzo, S. (2014). Clasificación socioeconómica de los municipios de Bolivia. *PERSPECTIVAS*(33), 29-55. Obtenido de <http://www.redalyc.org/articulo.oa?id=425941263002>.
- Bertsimas, D., & Dunn, J. (2019). *Machine Learning under a Modern Optimization Lens*. Waltham: Dynamic Ideas Press.
- Bertsimas, D., Orfanoudaki, A., Wiberg, H. (2021). Interpretable clustering: an optimization approach. *Machine learning*, 110: 89-138.
- Biau., G & Scornet., E (2016). A random forest guided tour. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*. Springer; Sociedad de Estadística e Investigación Operativa, vol. 25(2), pages 197-227, June
- Biblioteca Virtual de la Cooperación Alemana BIVICA*. (s.f). Recuperado el 3 de junio de 2022, de Biblioteca Virtual de la Cooperación Alemana BIVICA: <https://www.bivica.org/files/municipios-Bolivia.pdf>
- Bonet, J., Acosta, K., Pérez-Valbuena, G.J., Ayala-García, J. (2023). Is There a Case for Regional Policies in Colombia?. In: Haddad, E.A., Bonet, J., Hewings, G.J.D. (eds) *The Colombian Economy and Its Regional Structural Challenges*. Advances in Spatial Science. Springer, Cham. https://doi.org/10.1007/978-3-031-22653-3_2
- Carrillo, L. S. (2022). *Inventario y reseña de clasificaciones de entidades territoriales, como base para avanzar en diseño de tipologías para descentralización asimétrica*. Bogotá: Documento inédito.
- Cepeda Emiliani, L. y Meisel Roca, A. 2014. ¿Habrà una segunda oportunidad sobre la tierra? Instituciones coloniales y disparidades económicas regionales en Colombia. *Revista de Economía Institucional*. 16, 31 (dic. 2014).

Constitución Política de Colombia (1991). *Constitución Política de Colombia, 1991*. Colombia.

Contraloría General de la República (s.f). *Método de calculo de ICDL e indicadores Ley 617 de 2000*. Recuperado el 23 de 08 de 2023, de https://portalcef.contraloria.gov.co/assets/docs/539-d622c1a-598c-4480-b74f-8851833dcb37/202001-M%C3%A9todo-de-c%C3%A1lculo-ICLD-e-indicadores-CEF_94502.pdf

Decreto 222 de 1988. *Por el cual se establecen categorías de municipios*. Colombia: Sistema Único de Información Normativa. Recuperado el 02 de 02 de 2023, de <https://www.suin-juriscol.gov.co/viewDocument.asp?id=1050271#:~:text=DECRETA%3A,fiJar%20para%20los%20respectivos%20Alcaldes>.

Decreto 2106 de 2019. “*Por el cual se dictan normas para simplificar, suprimir y reformar trámites, procesos y procedimientos innecesarios existentes en la administración pública*”. Colombia: Función Pública.

Departamento Nacional de Planeación DNP. (2015). *Tipologías Departamentales y Municipales: una propuesta para comprender las entidades territoriales colombianas*. Bogotá: DNP.

Departamento Nacional de Planeación. (2016). *Nueva Medición de Desempeño Municipal MDM Boletín no. 2*. Bogotá: DNP.

Departamento Nacional de Planeación. (s.f). Sistema de Estadísticas Territoriales de Colombia -Terridata. <https://terridata.dnp.gov.co/>

Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. arXiv: Machine Learning. Preprint.

Fraiman, R., Ghattas, B., & Svarc, M. (2013). Interpretable clustering using unsupervised binary trees. *Advances in Data Analysis and Classification*, 7(2), 125–145

Genuer, Robin., Poggi, Jean-Michel., Tuleau-Malot, Christine. (2010). Variable selection using random forests. *Pattern Recognition Letters*. Volume 31, Issue 14 (2010), Pages 2225-2236. ISSN 0167-8655, <https://doi.org/10.1016/j.patrec.2010.03.014>.

Giopp, M. (1993). Sobre la tipología municipal. *La Gaceta Mexicana de Administración Pública Estatal y Municipal*(42,43 y 44), 23-32.

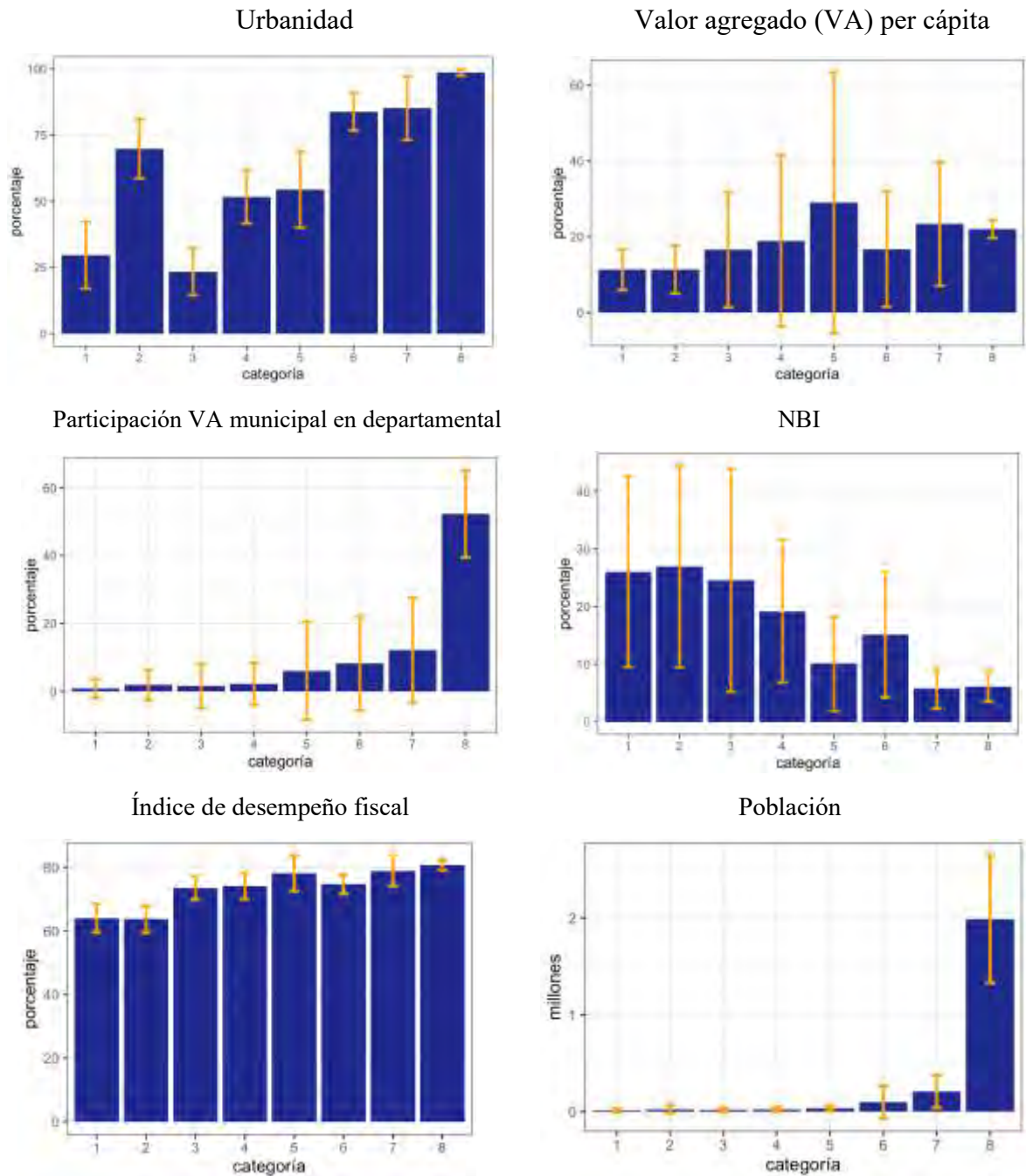
Hernández Becerra, A. (2015). Objetivos Inéditos de la Categorización Municipal. (P. A. Acosta, Ed.) *Serie Documentos de Trabajo n° 56 Departamento de Derecho Constitucional*, 1-19.

- Ley 136 de 1994. *Por la cual se dictan normas tendientes a modernizar la organización y el funcionamiento de los municipios*. Colombia: Departamento Administrativo de la Función Pública.
- Ley 617 de 2000. *Por la cual se reforma parcialmente la Ley 136 de 1994*. Colombia.
- Ley 1551 de 2012. *Por la cual se dictan normas para modernizar la organización y el funcionamiento de los municipios*. Colombia.
- Ley XV - N°5 (antes Ley 257). (s.f). Ley Orgánica de Municipalidades.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *In Proceedings of the 5th Berkeley symposium on mathematical statistics and probability*, volume 1: Statistics (pp. 281–297). Berkeley, CA: University of California Press.
- Ministerio de Economía y Finanzas de Perú. (25 de 04 de 2022). *ANEXO N° 01: CLASIFICACIÓN MUNICIPAL*. Obtenido de https://www.mef.gob.pe/es/?option=com_content&language=es-ES&Itemid=101547&lang=es-ES&view=article&id=2221
- Ministerio del Interior y Seguridad Pública de Chile. (2015). *Resolución 178*. Biblioteca del Congreso Nacional de Chile BCN. Obtenido de <https://www.bcn.cl/leychile/navegar?idNorma=1079104&idVersion=2021-07-07&idParte=>
- Morales, M. (2010). *Las Municipalidades en Costa Rica: Evaluación del sistema tributario y la administración presupuestaria y financiera de los gobiernos locales. Alternativas para su fortalecimiento*. Washington, D.C., U.S.A.: Banco Interamericano de Desarrollo.
- Nion, A., & Feo, Y. (2017). *Clasificación socioterritorial de los municipios - Reporte 6*. Montevideo: Observatorio Territorio Uruguay Dirección de Descentralización e Inversión Pública - OPP.
- Organización para la Cooperación y el Desarrollo Económicos OCDE. (2019). *Asymmetric decentralisation: Policy implications in Colombia*. Multi-Level Governance Studies. OECD.
- Orozco Barbosa, J. (1993). Tipología de los municipios de México. *La Gaceta Mexicana de Administración Pública Estatal y Municipal*(42,43 y 44), 41-45.
- Perea Gómez, J. M., & Delúquez Medina, K. M. (2020). *La categorización de municipios en Colombia, un esfuerzo inconcluso y errático (Tesis Maestría)*. Bogotá: Pontificia Universidad Javeriana.
- Pérez Rasgado, F. (2014). *Tipología del Municipio Mexicano para su Desarrollo Integral*. México: Instituto Nacional de Administración Pública, A.C.

- Pérez-Valbuena, G.J & Ricciulli-Marín, D & Bonet-Morón, J & Barrios, P. (2021). "Reglas fiscales subnacionales en Colombia: desde su concepción hasta los resultados frente al COVID-19," Documentos de trabajo sobre Economía Regional y Urbana 297, Banco de la República de Colombia.
- Presidencia de la República del Ecuador. (19 de octubre de 2010). Código Orgánico de Organización Territorial, Autonomía y Descentralización. *Registro Oficial Órgano del Gobierno del Ecuador*.
- Sánchez Martínez, L. A. (1993). Apuntes para el estudio de las tipologías municipales. *La Gaceta Mexicana de Administración Pública Estatal y Municipal*(42,43 y 44), 15-17.
- Struyf, Anja & Hubert, Mia & Rousseeuw, Peter J., 1997. "Integrating robust clustering techniques in S-PLUS," *Computational Statistics & Data Analysis*, Elsevier, vol. 26(1), pages 17-37, November.
- Tan Evrim & Avshalom-Uster Anna (2021) How does asymmetric decentralization affect local fiscal performance?, *Regional Studies*, 55:6, 1071-1083, DOI: 10.1080/00343404.2020.1861241
- Vivanco, L., & Cordero, F. (2019). Descentralización de la planificación del ordenamiento territorial en Ecuador, 2008-2018. *Ecuador Debate* 108, 73-100.

Anexo

Figura A1. Promedio de las variables por categoría basada en capacidades



Fuente: DANE, DNP, Contaduría General de la Nación. Cálculos de las autoras. Nota: las barras naranjas representan la desviación estándar.