



HYPERION ONE · APPLIED INTELLIGENCE

Informe Nacional de Inteligencia sobre Culminación de Educación Media en Colombia

14.512.497 adultos sin bachillerato (2026, estimado) · 1.122 municipios · 81 fuentes · 12.723 conversaciones públicas

Uno de cada 2,7 adultos colombianos no ha terminado el bachillerato

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

ADULTOS 18+ SIN MEDIA COMPLETA · 2026 [C]

DE LOS ADULTOS (45,4 % EN 2018)

COTAS CENSALES 2018 [A]

MUNICIPIOS ALTA NECESIDAD · ALTA OPORTUNIDAD

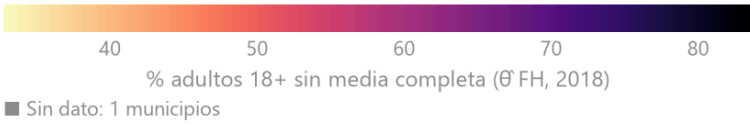
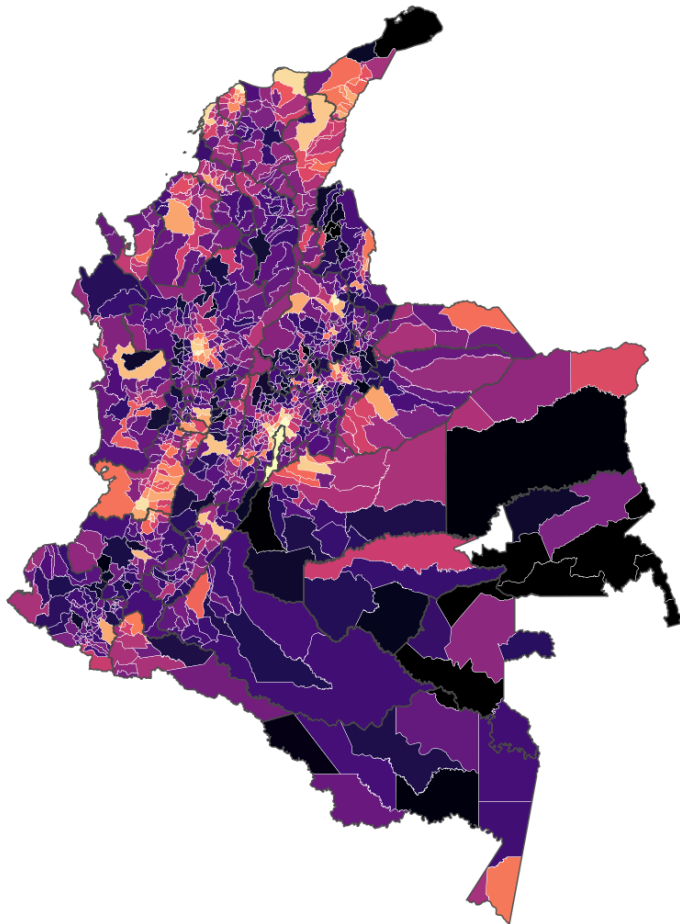
Censo 2018 a nivel de persona + GEIH 2025 + proyecciones DANE 2026, integrados con Small Area Estimation. La cifra no es un dato: es una estimación con intervalo (IC80 14.448.777–14.576.214) y con cotas censales observadas. El mercado no es homogéneo: volumen en el centro y la costa; tasa en la periferia.

El mapa nacional del rezago ($\theta^{\wedge}$ 2018 por municipio)

REZAGO EDUCATIVO ADULTO 2018

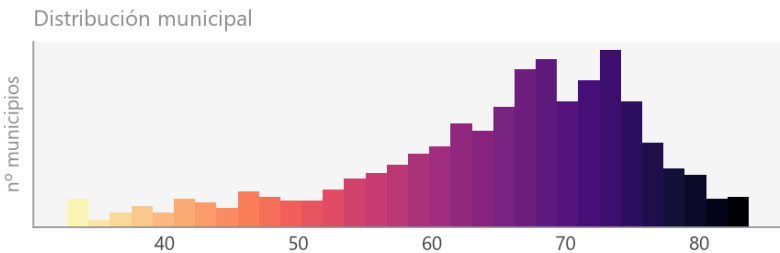
San Andrés y Providencia fuera del encuadre

$\theta^{\wedge}$: 991 de 1.122 municipios tienen más de la mitad de sus adultos 18+ sin media completa; promedio ponderado nacional 45,4 %



Mayores 8 municipios

1. Pana Pana (Guainía)	95,7 %
2. La Victoria (Amazonas)	92,5 %
3. Morichal (Guainía)	91,7 %
4. Puerto Colombia (Guainía)	90,3 %
5. Papunahua (Vaupés)	89,2 %
6. Cacahual (Guainía)	89,1 %
7. Hacarí (N. Santander)	84,8 %
8. Pacoa (Vaupés)	84,6 %



Fuente: Fay-Herriot espacial capa 2 sobre CNPV 2018 (S001) + proyecciones DANE 2018-2035 (S003) · Periodo: 2018

ESTIMADO [C]

Rezago educativo adulto 2018. 991 municipios con $\theta^{\wedge}$ > 50 %; promedio ponderado nacional 45,4 %

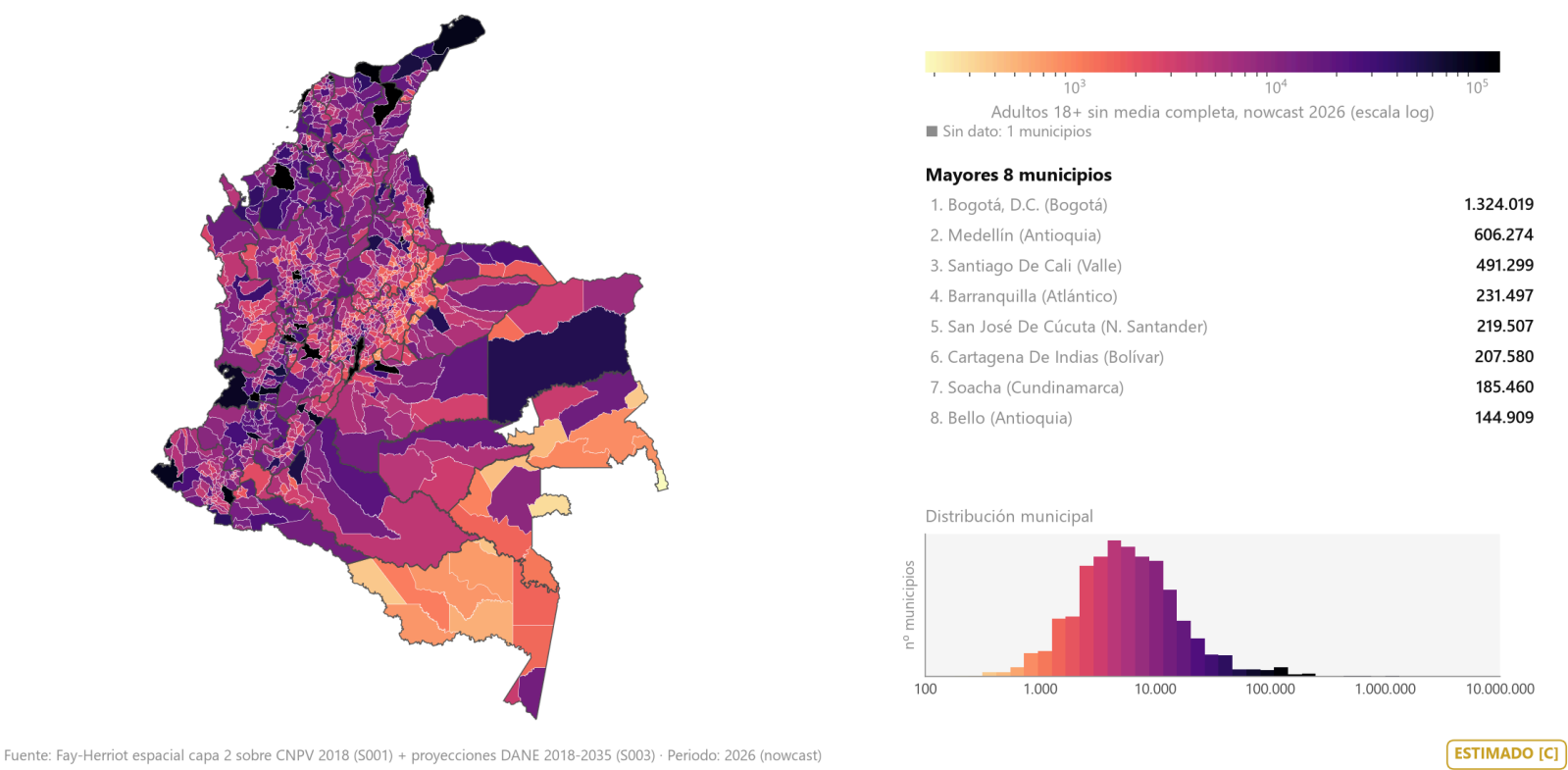
ESTIMADO ['S001', 'capa2_fh']

Dónde viven las personas objetivo en 2026

TARGET 2026: PERSONAS SIN MEDIA

San Andrés y Providencia fuera del encuadre

14,5 millones de adultos sin media completa en 2026; los 10 municipios mayores concentran 25,4 % del total

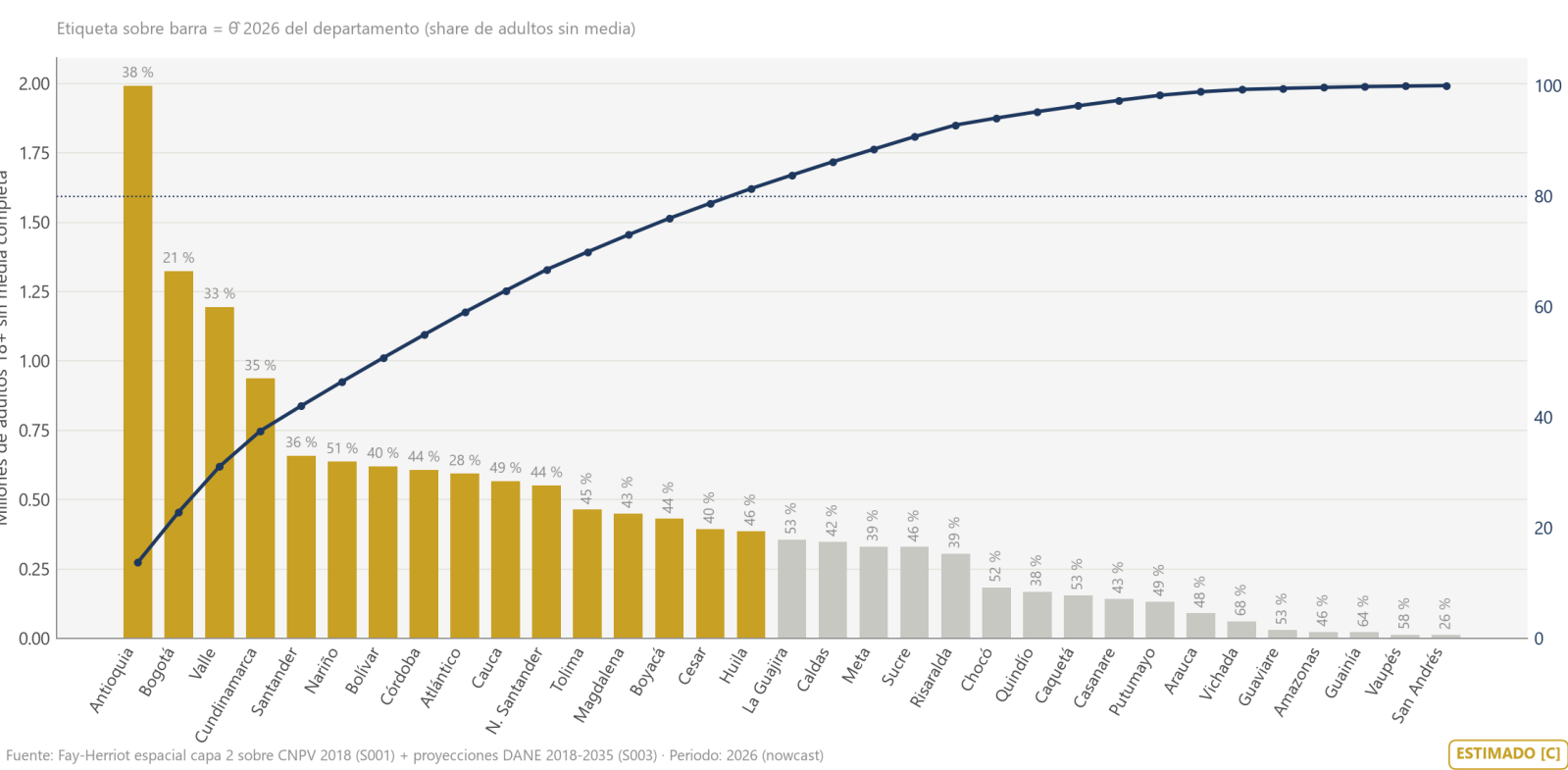


Target 2026: personas sin media. Target nacional 2026: 14.512.497 personas; top 10 municipios = 25,4 %

ESTIMADO ['S001', 'S003', 'capa2_fh']

TARGET 2026 POR DEPARTAMENTO

16 departamentos concentran el 80 % de los 14,5 M adultos sin media; Antioquia lidera con 1,99 M (37,6 % de sus adultos)



Target 2026 por departamento. 16 departamentos = 80 % del target; Antioquia 1.992.266

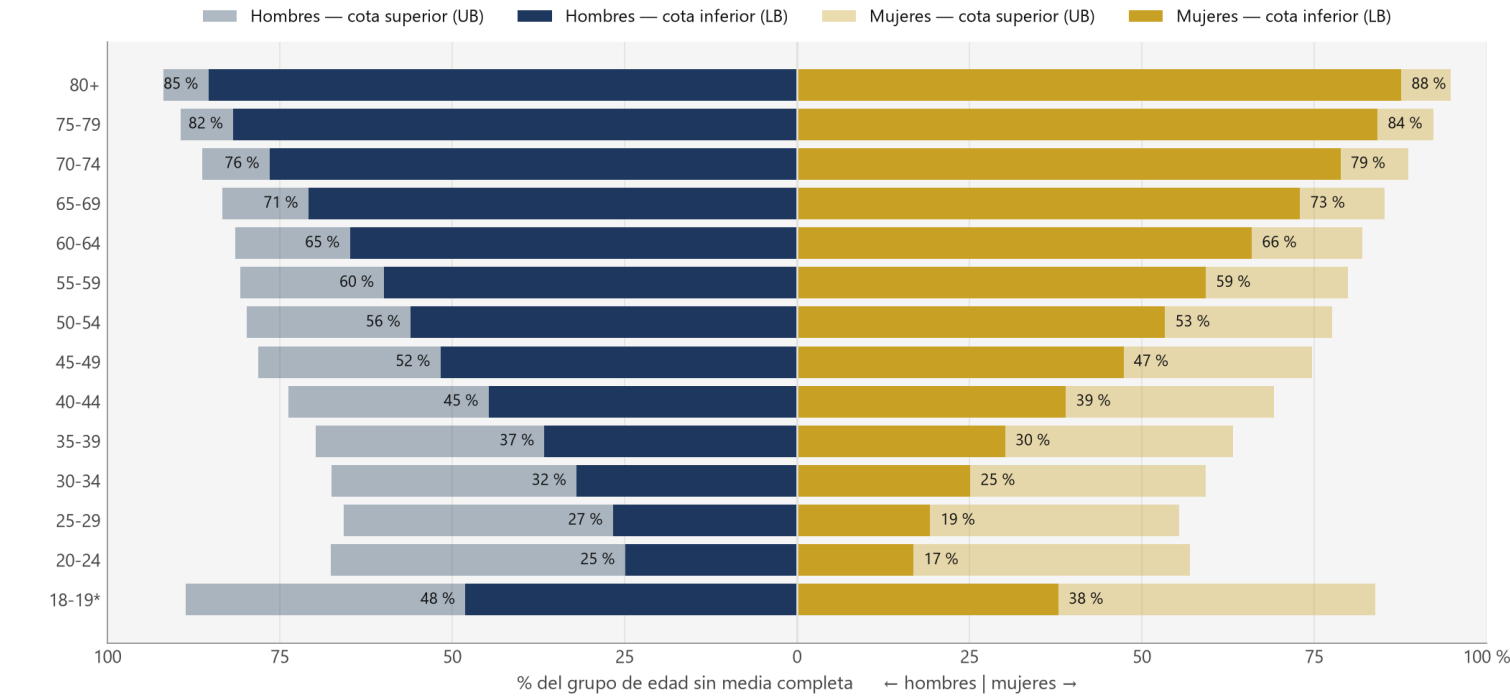
ESTIMADO ['S001', 'S003', 'capa2_fh']

Edad, sexo y campo/ciudad

PIRÁMIDE DEL REZAGO POR EDAD Y SEXO (2018)

* 18-19 aproximado desde grupo 15-19

El rezago crece con la edad: 26,0 % de los adultos de 20-39 no terminó la media (cota inferior) frente a 65,4 % a los 60-64; hombres 45,5 % vs mujeres 42,4 %

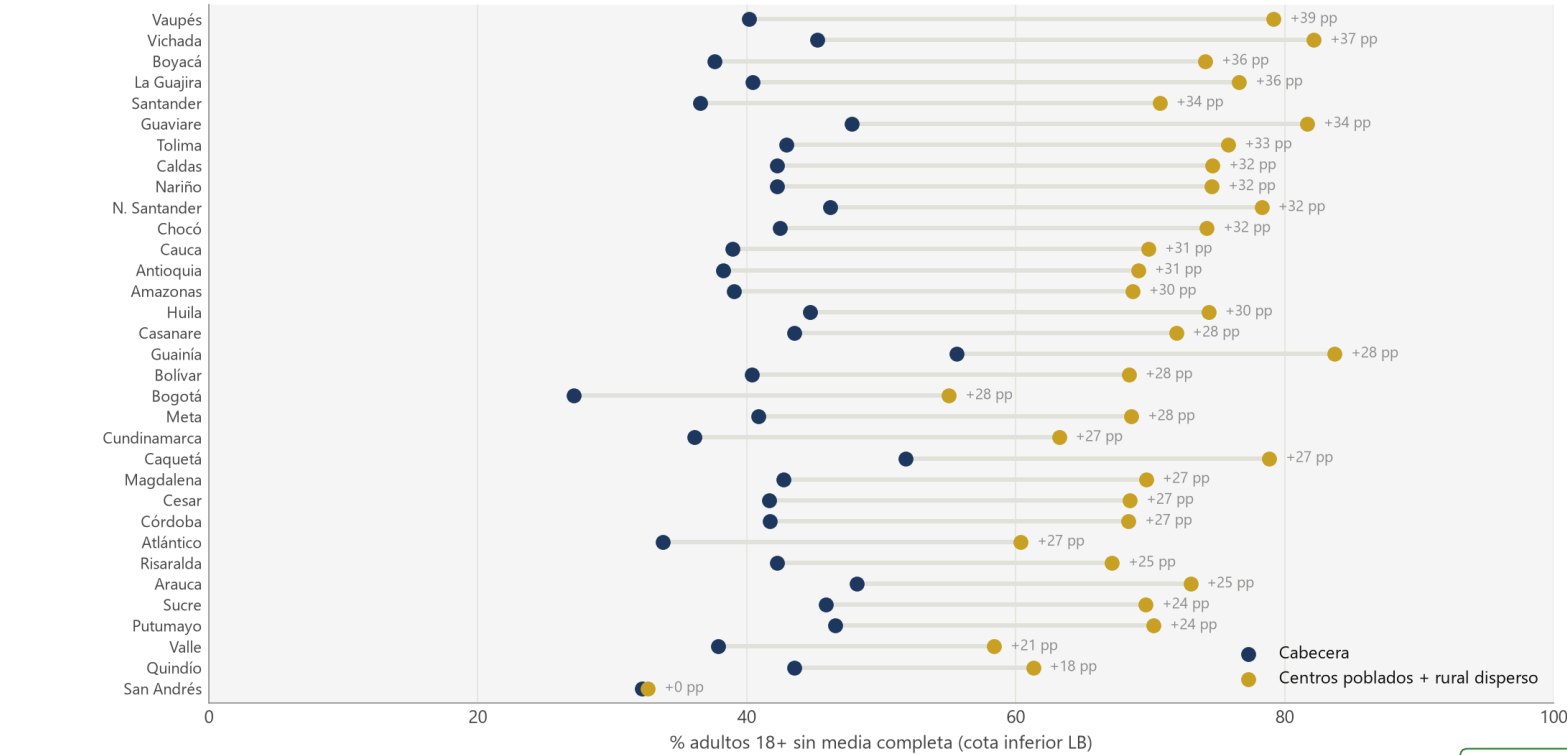


Fuente: DANE CNPV 2018 microdatos (S001) — capa 0 · Periodo: 2018

OBSERVADO [A]

BRECHA CABECERA VS RURAL POR DEPARTAMENTO

Nacional: 69,9 % de adultos rurales sin media frente a 36,9 % en cabeceras (+33 pp); la brecha máxima es Vaupés (+39 pp)



Fuente: DANE CNPV 2018 microdatos (S001) — capa 0 · Periodo: 2018

OBSERVADO [A]

Pirámide del rezago por edad y sexo (2018). LB 20-39: 26,0 %; hombres 45,5 % vs mujeres 42,4 %

OFICIAL ['S001']

Brecha cabecera vs rural por departamento. Rural 69,9 % vs cabecera 36,9 %; brecha máxima Vaupés +39 pp

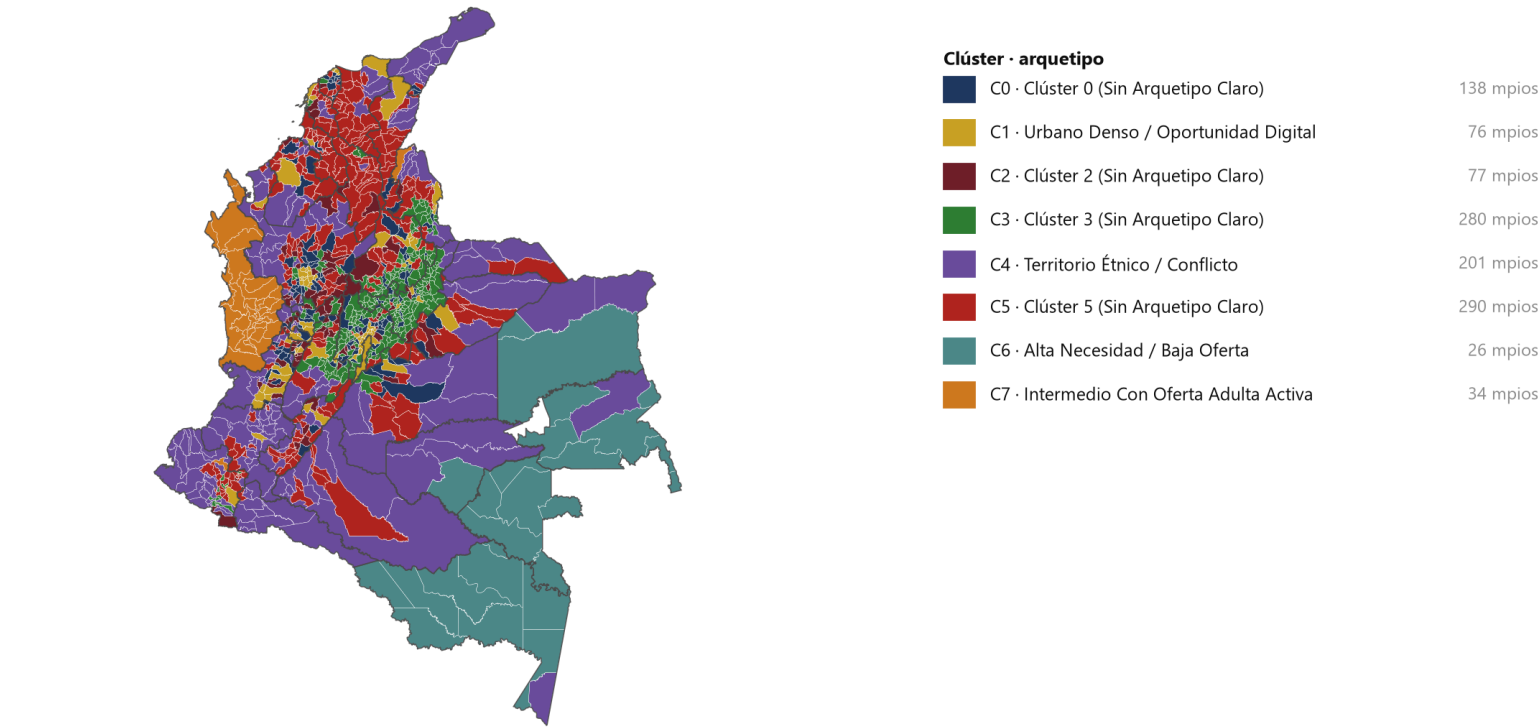
OFICIAL ['S001']

El machine learning encuentra tipos de territorio

ARQUETIPOS TERRITORIALES (K-MEANS, K=8)

San Andrés y Providencia fuera del encuadre

8 clústeres (silhouette 0.18); 4 tienen arquetipo interpretable, p. ej. 'Urbano Denso / Oportunidad Digital' con 76 municipios



Fuente: K-means (k=8) + UMAP sobre 19 features estandarizadas de municipal_master · Periodo: 2018-2024

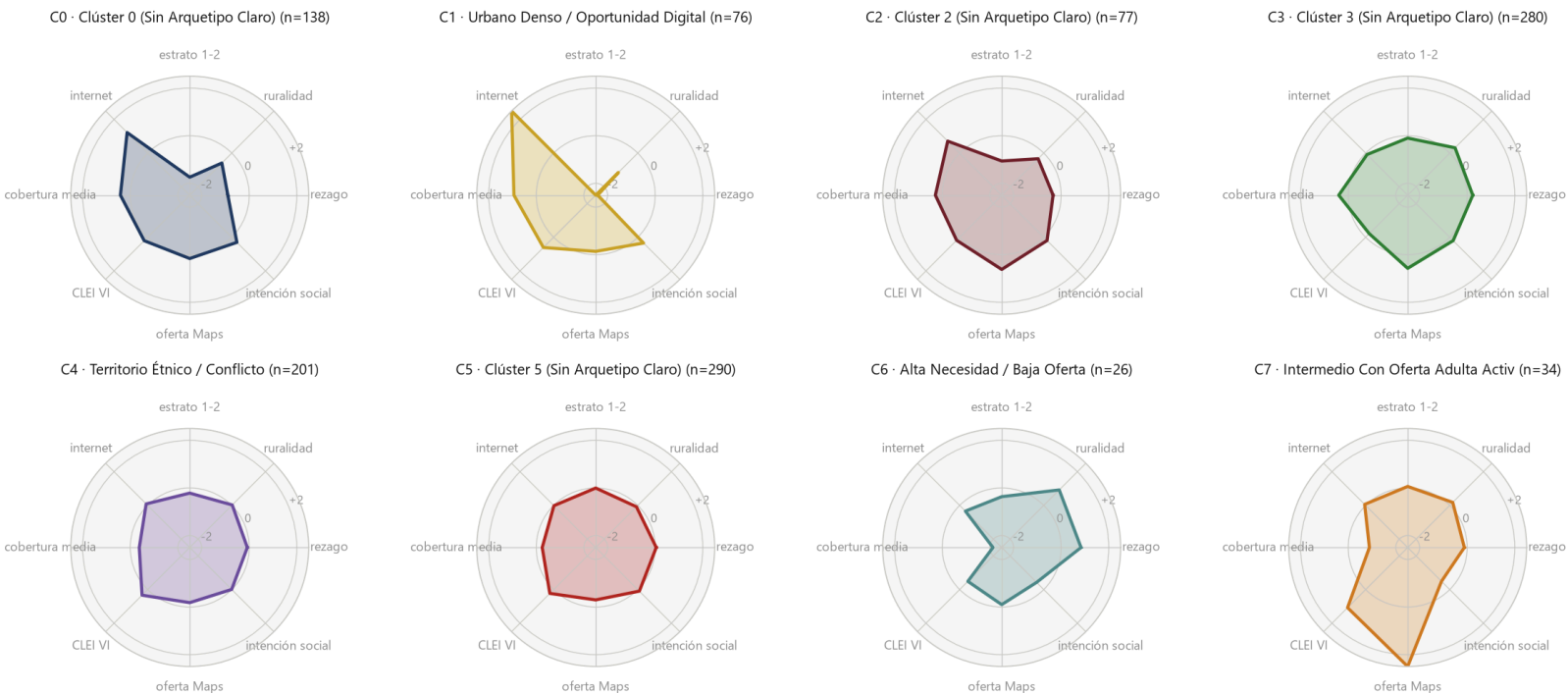
INFERIDO [G]

Arquetipos territoriales (k-means, k=8). 8 clústeres, silhouette 0.18; el mayor tiene 290 municipios

INFERIDO ['municipal_master']

PERFILES DE CLÚSTER (Z-SCORES)

Perfiles medios estandarizados: el clúster 'alta necesidad / baja oferta' (26 municipios) combina rezago y ruralidad altos con la menor oferta CLEI; el urbano denso es su imagen especular



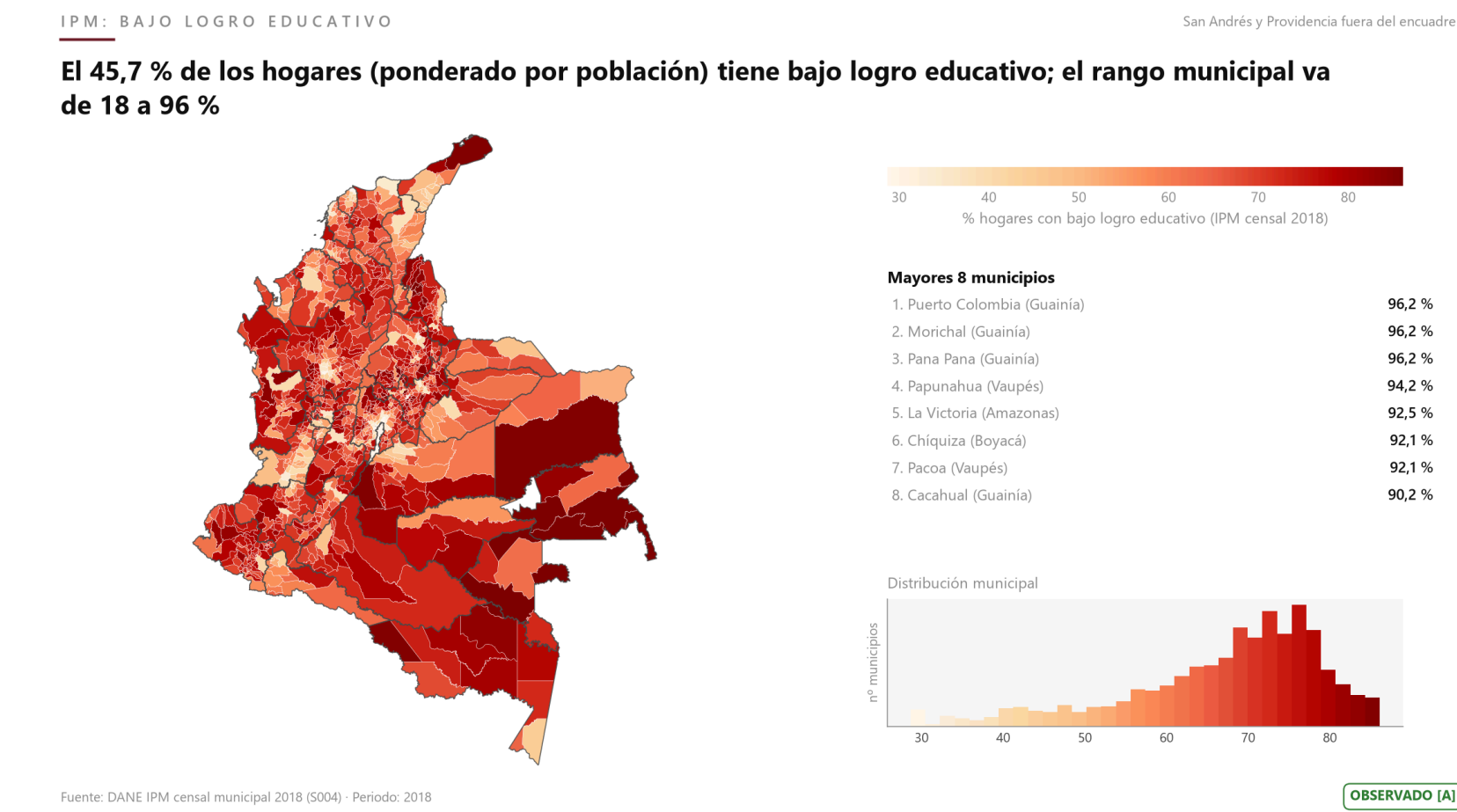
Fuente: K-means (k=8) + UMAP sobre 19 features estandarizadas de municipal_master · Periodo: 2018-2024

INFERIDO [G]

Perfiles de clúster (z-scores). 8 perfiles de clúster en 8 variables estandarizadas

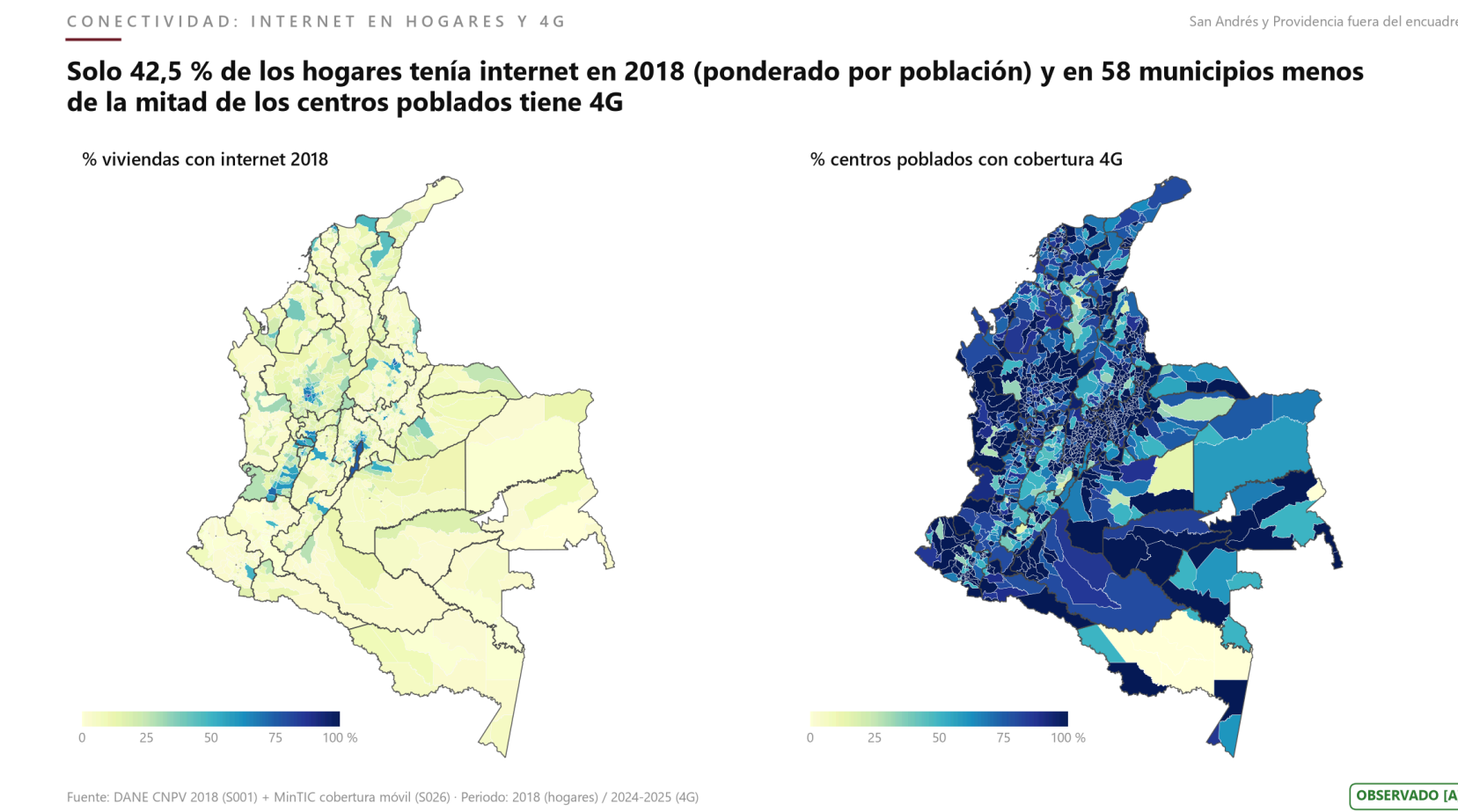
INFERIDO ['municipal_master']

Pobreza multidimensional y conectividad explican el mapa (asociación, no causalidad)



IPM: bajo logro educativo. Bajo logro educativo ponderado nacional 45,7 %

OFICIAL ['S004']



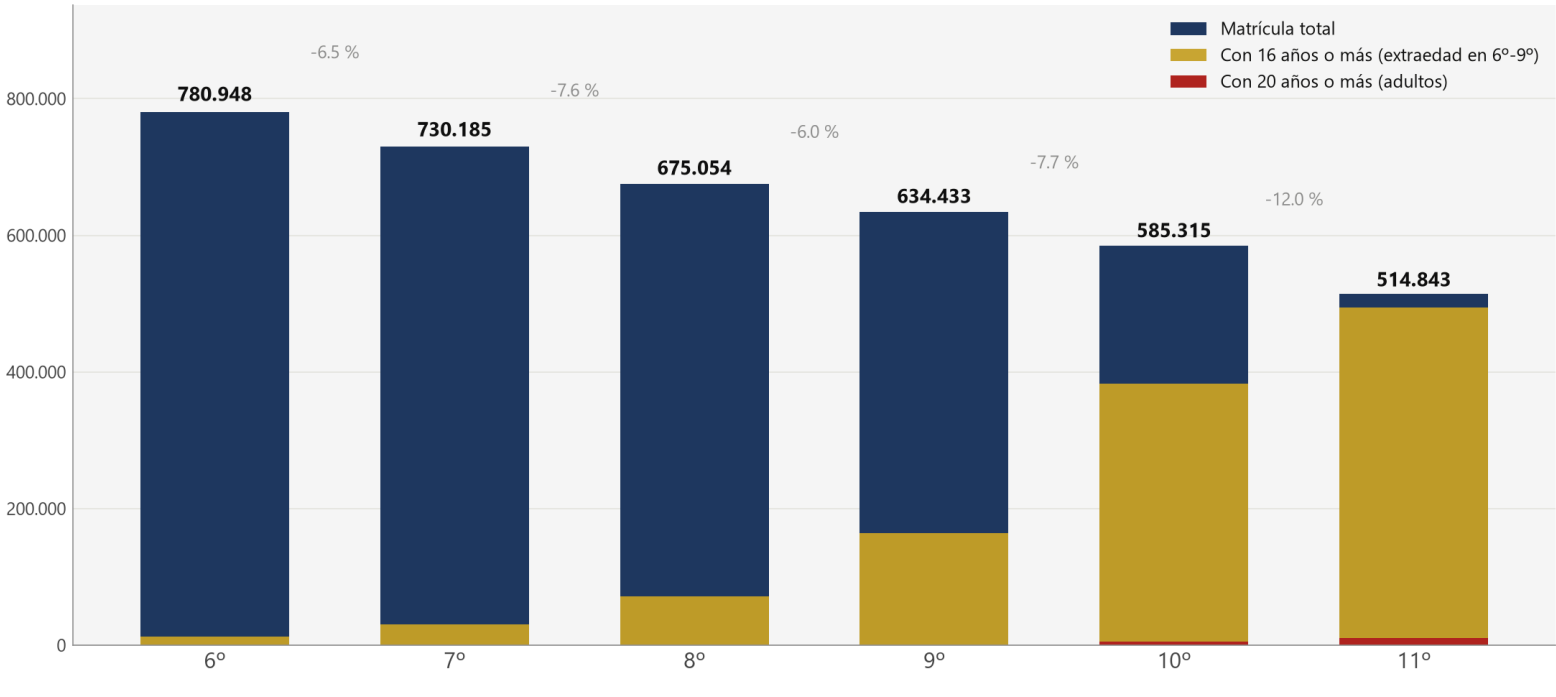
Conectividad: internet en hogares y 4G. Internet hogares 2018 ponderado 42,5 %; 58 municipios con < 50 % de CP con 4G

OFICIAL ['S001', 'S026']

El sistema pierde estudiantes entre 9º y 11º

EMBUDO ESCOLAR 6º → 11º (CORTE 2023)

Por cada 100 estudiantes matriculados en 6º hay 66 en 11º; la mayor fuga ocurre entre 9º y 10º (7,7 % menos matrícula)



Fuente: DANE EDUC 2023 microdatos C-600 (S080) · Periodo: 2023

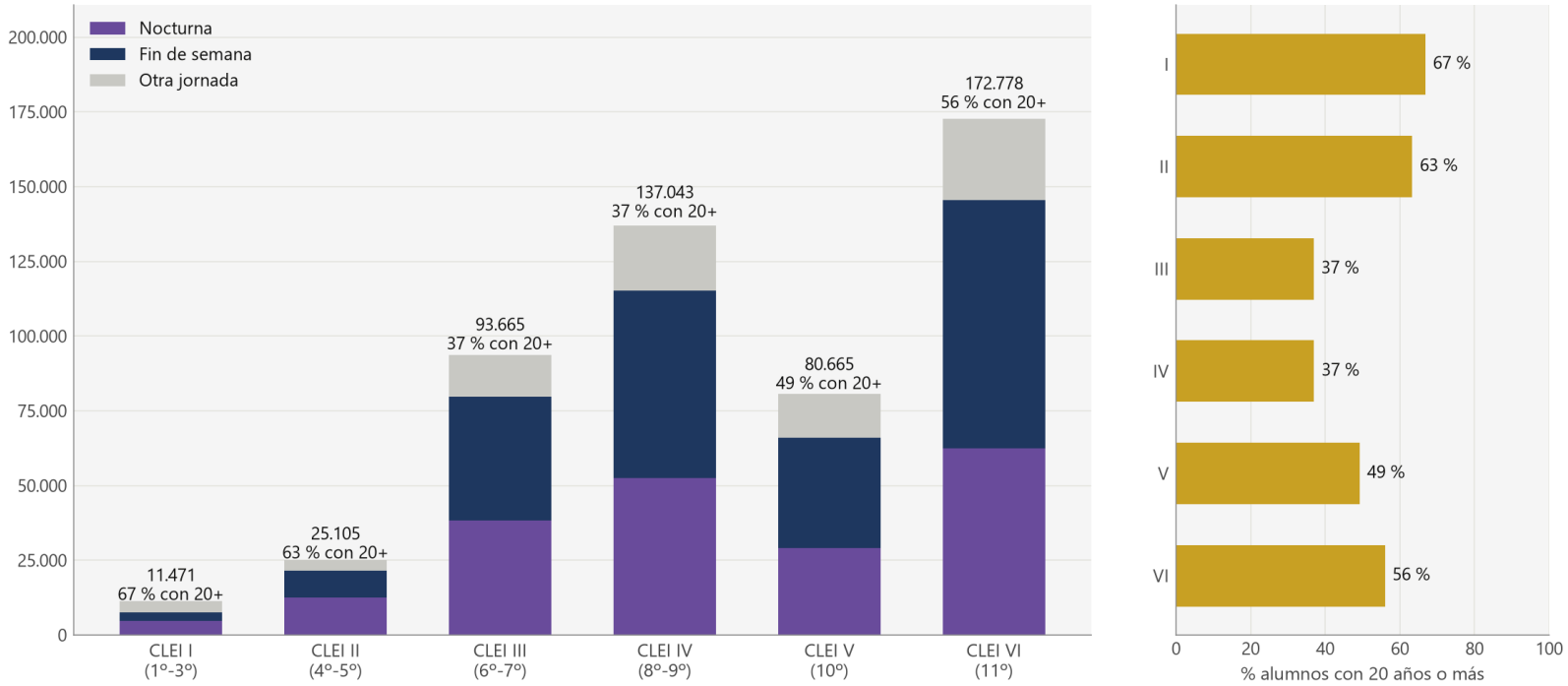
OBSERVADO [A]

Embudo escolar 6º → 11º (corte 2023). Retención transversal 11º/6º = 65,9 %; caída 9º→10º 7,7 %

OFICIAL ['S080']

CLEI I → VI: MATRÍCULA Y JORNADA (2023)

253.443 personas cursan CLEI V-VI (media para adultos); 53,8 % tienen 20 años o más y 83,4 % estudian de noche o fines de semana



Fuente: DANE EDUC 2023 microdatos C-600 (S080) · Periodo: 2023

OBSERVADO [A]

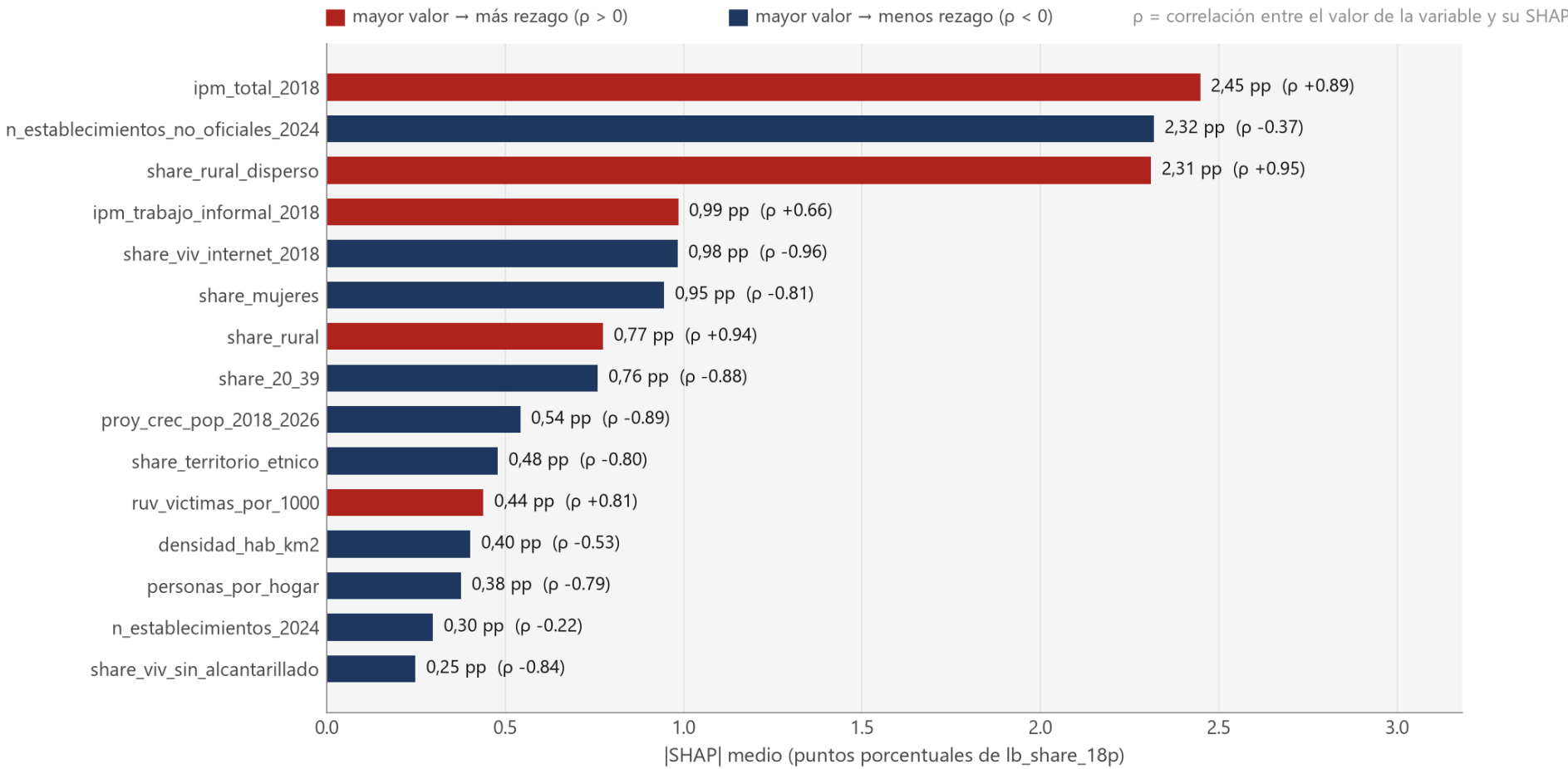
CLEI I → VI: matrícula y jornada (2023). 253.443 alumnos CLEI V-VI; 53,8 % con 20+ años

OFICIAL ['S080']

SHAP: qué variables explican un territorio de alto rezago

QUÉ EXPLICA EL REZAGO ADULTO (SHAP GLOBAL)

XGBoost explica R² = 0.78 del rezago en validación espacial por bloques; pobreza multidimensional, oferta privada y ruralidad dispersa dominan la importancia



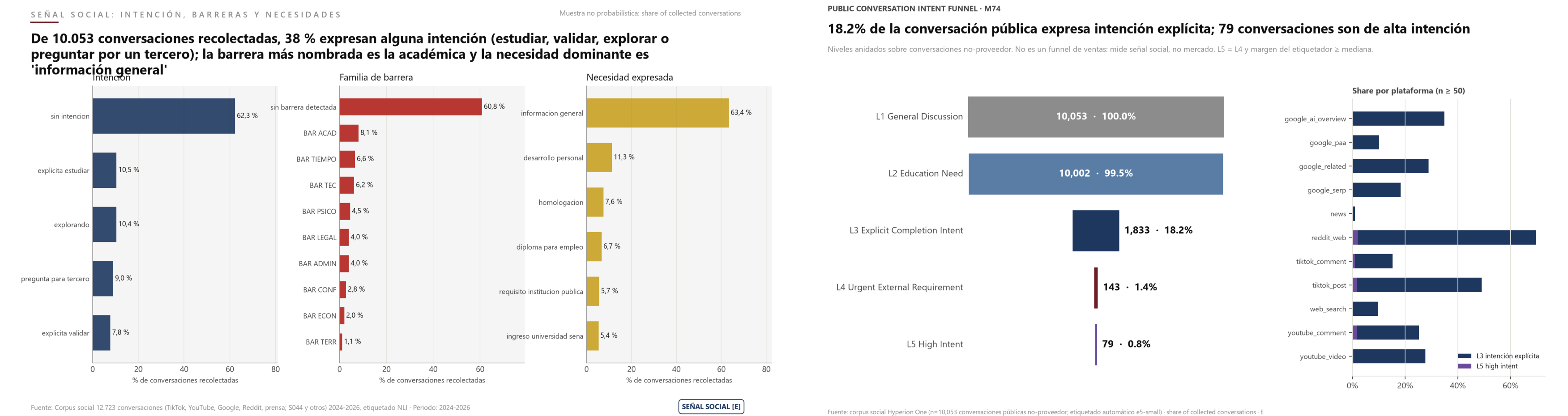
Fuente: XGBoost + SHAP TreeExplainer sobre municipal_master (26 covariables), validación espacial por bloques · Periodo: 2018-2024

INFERIDO [G]

Qué explica el rezago adulto (SHAP global). R² espacial 0.78; top features: ipm_total_2018, n_establecimientos_no_oficiales_2024, share_rural_disperso

INFERIDO ['municipal_master', 'models/ml']

Lo que la gente dice: necesidades, barreras y sentimiento (share of collected conversations)

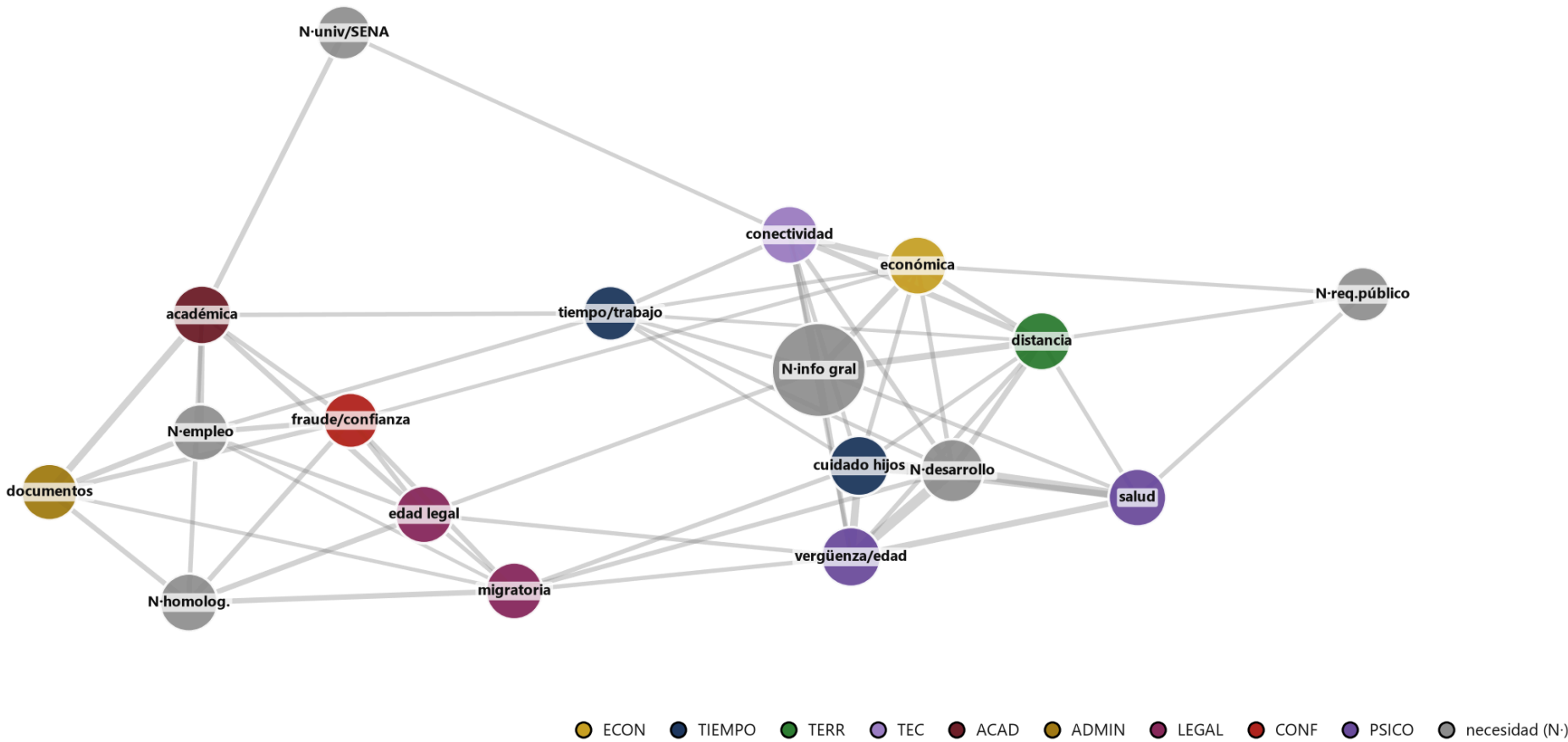


La red de barreras: costo, tiempo/trabajo, distancia, documentos, miedo académico

NETWORK OF BARRIERS · M76 · COLOMBIA

Cuidado hijos × vergüenza/edad: 305 conversaciones (3.0%), el par de barreras que más co-ocurre

Tamaño = frecuencia · grosor = co-ocurrencia en la misma conversación · color = familia de barrera (N= necesidad). Top 60 aristas con peso ≥ 3.

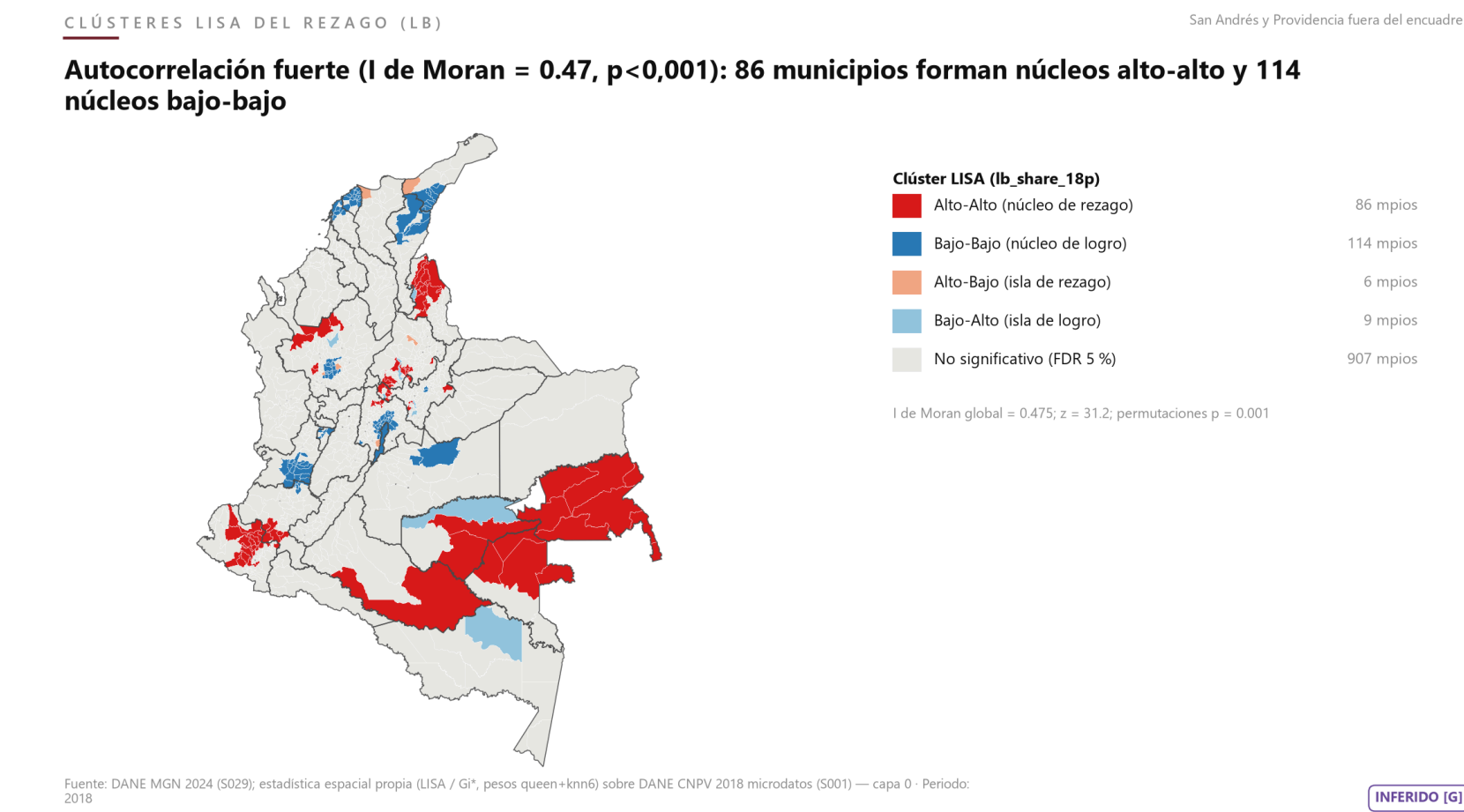


Fuente: corpus social Hyperion One (n=10,053 conversaciones públicas no-proveedor; etiquetado automático e5-small) · share of collected conversations · E

Barrier network colombia.

SEÑAL SOCIAL corpus social (share of collected conversations)

El rezago se agrupa: clústeres LISA y hotspots

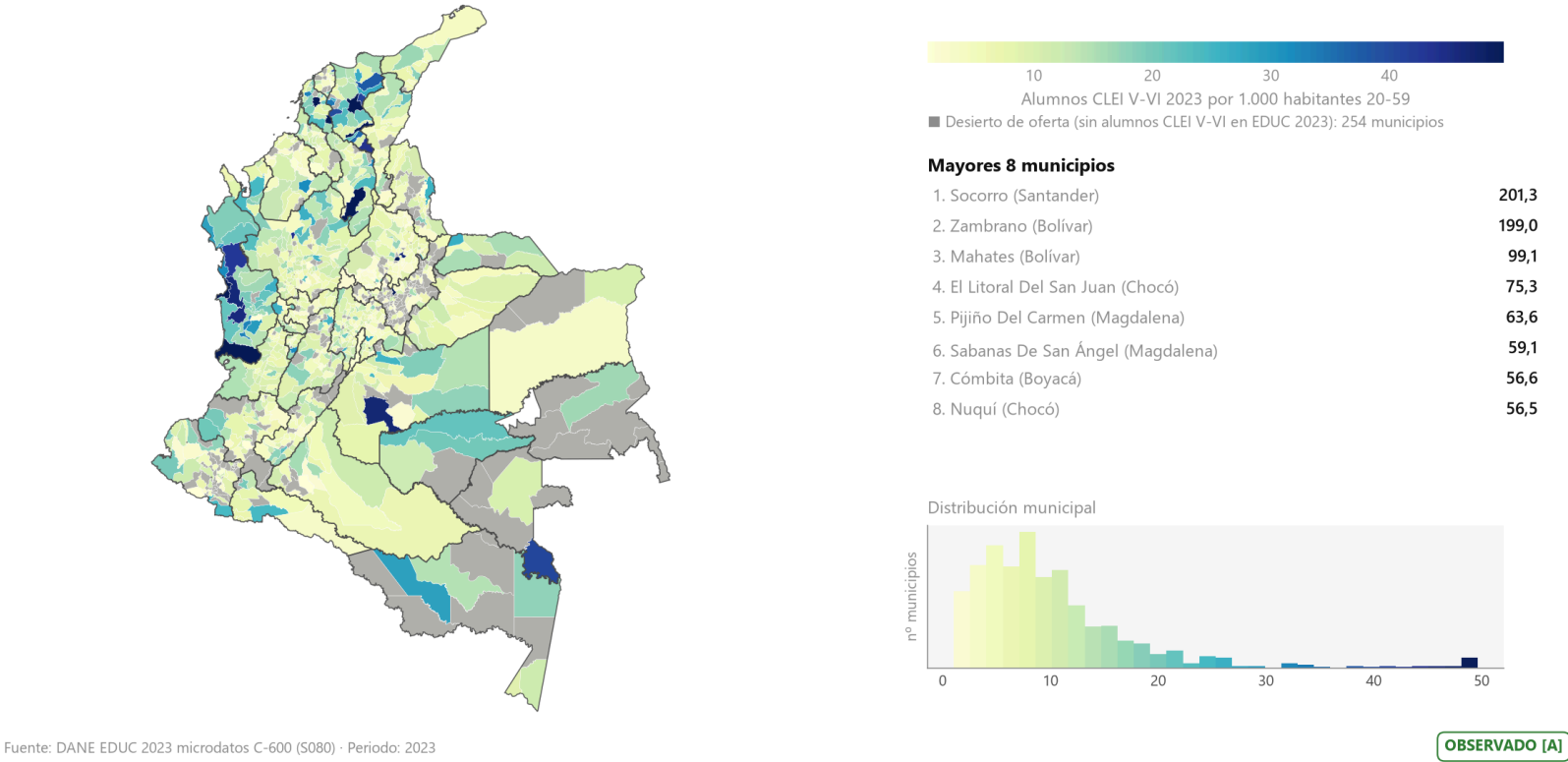


Desiertos de oferta: municipios sin CLEI V-VI

OFERTA CLEI V-VI POR 1.000 ADULTOS

San Andrés y Providencia fuera del encuadre

253 municipios son desiertos sin ningún alumno en CLEI V-VI (2023); allí viven 0,84 M adultos sin media (5,8 % del target)



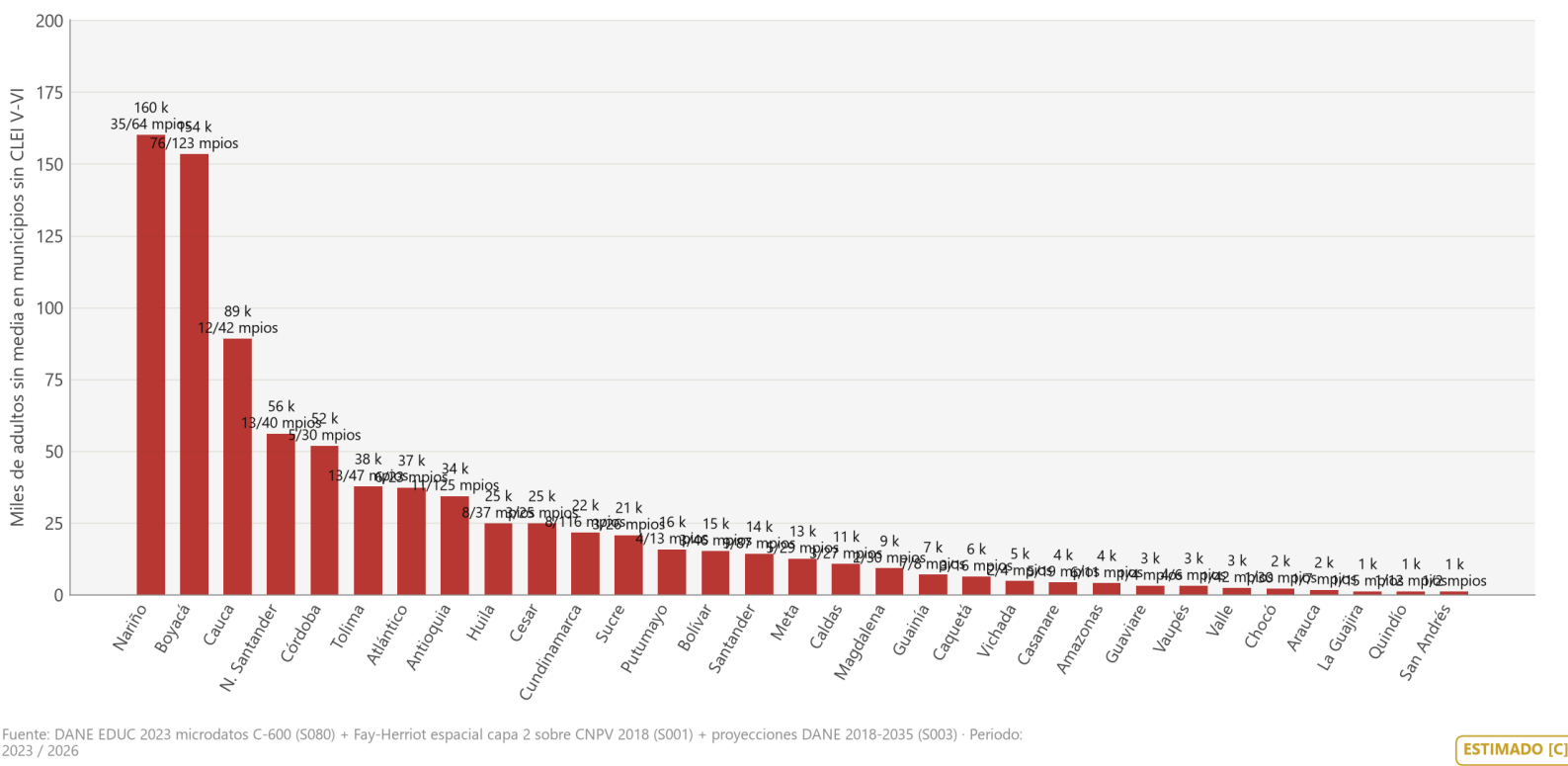
Oferta CLEI V-VI por 1.000 adultos. 253 desiertos de oferta CLEI V-VI con 835.554 adultos sin media

OFICIAL ['S080', 'S001']

DESIERTOS DE OFERTA CLEI V-VI POR DEPARTAMENTO

Sin accessibility_mpio.parquet: desierto = municipio sin registros CLEI V-VI en EDUC 2023

253 municipios sin ningún alumno en CLEI V-VI (2023) albergan 835.554 adultos sin media (5,8 % del target); Nariño concentra 160.239 de ellos en 35 municipios



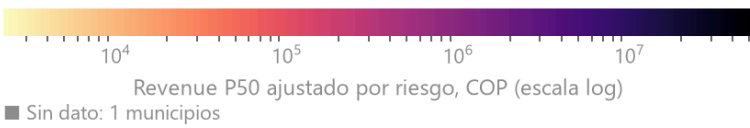
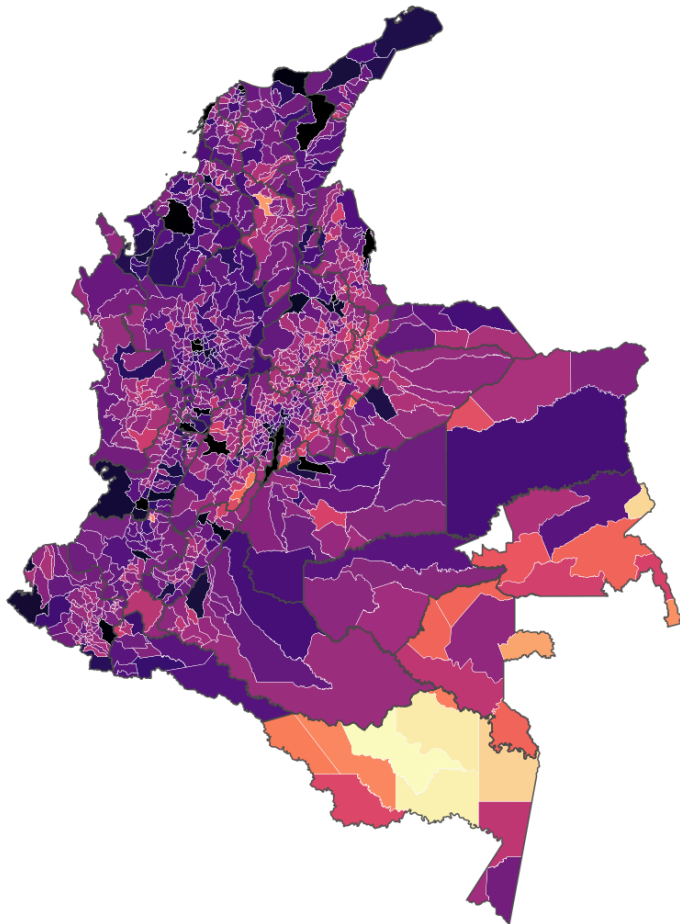
Desiertos de oferta CLEI V-VI por departamento. 253 desiertos con 835.554 adultos sin media (5,8 %)

ESTIMADO ['S080', 'S001', 'capa2_fh']

Corredores de oportunidad: mercados lógicos con hub

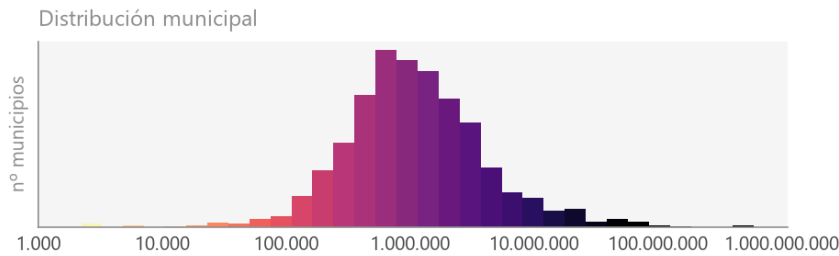
INGRESO P50 AJUSTADO POR RIESGO

Los 20 municipios líderes explican 47,9 % de los 4,5 mM COP de ingreso P50 ajustado por riesgo (escenario base, 12 meses)



Mayores 8 municipios

1. Bogotá, D.C. (Bogotá)	599 M COP
2. Medellín (Antioquia)	406 M COP
3. Soacha (Cundinamarca)	140 M COP
4. Santiago De Cali (Valle)	119 M COP
5. Barranquilla (Atlántico)	112 M COP
6. Bello (Antioquia)	81 M COP
7. Villavicencio (Meta)	80 M COP
8. Bucaramanga (Santander)	74 M COP



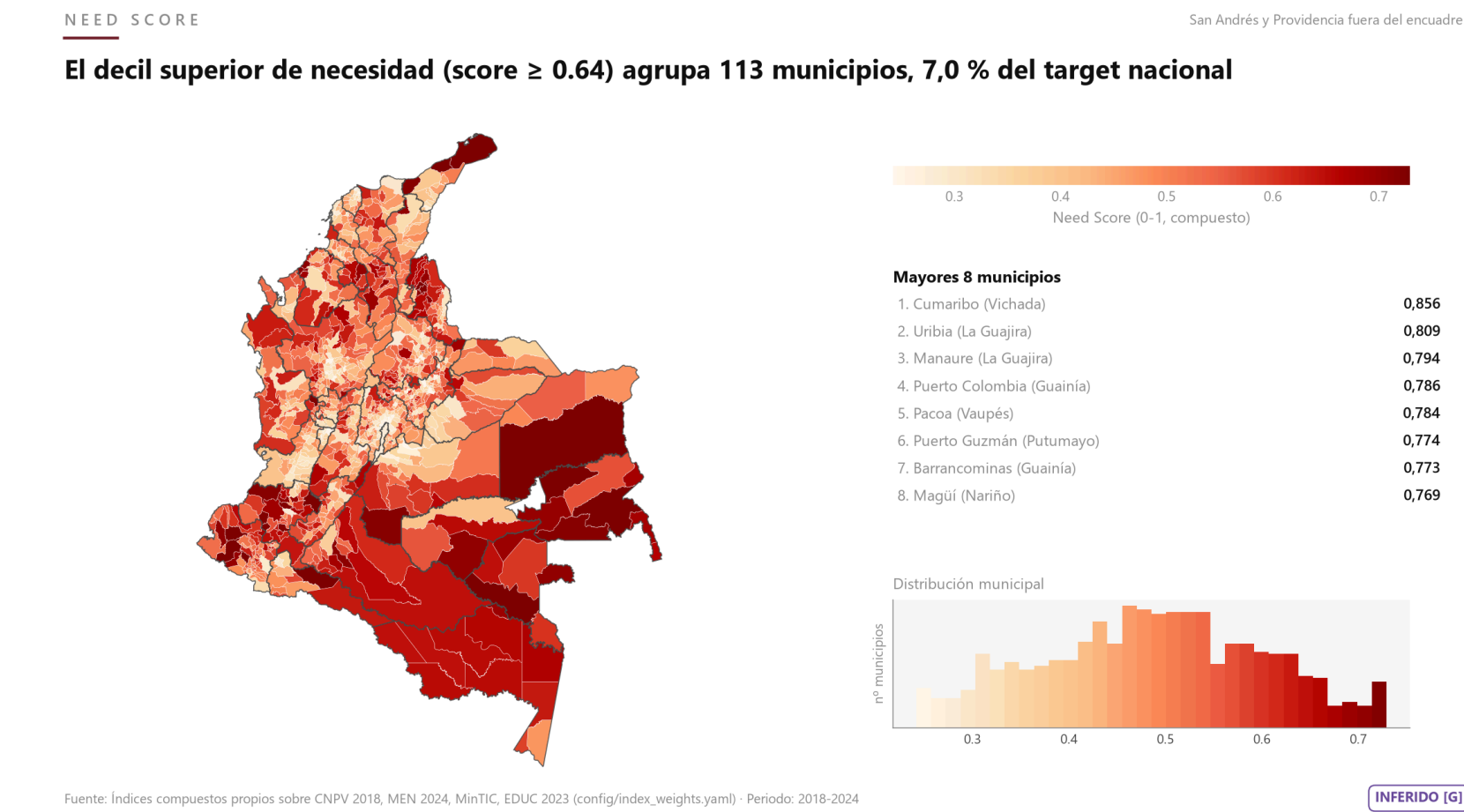
Fuente: Motor comercial Monte Carlo (config/business_assumptions.yaml, ASSUMPTION) sobre target FH 2026 · Periodo: 12 meses desde 2026

ESTIMADO [C]

Ingreso P50 ajustado por riesgo. Top 20 municipios = 47,9 % del ingreso P50 ajustado (4,5 mM COP)

ESTIMADO ['capa2_fh', 'business_assumptions']

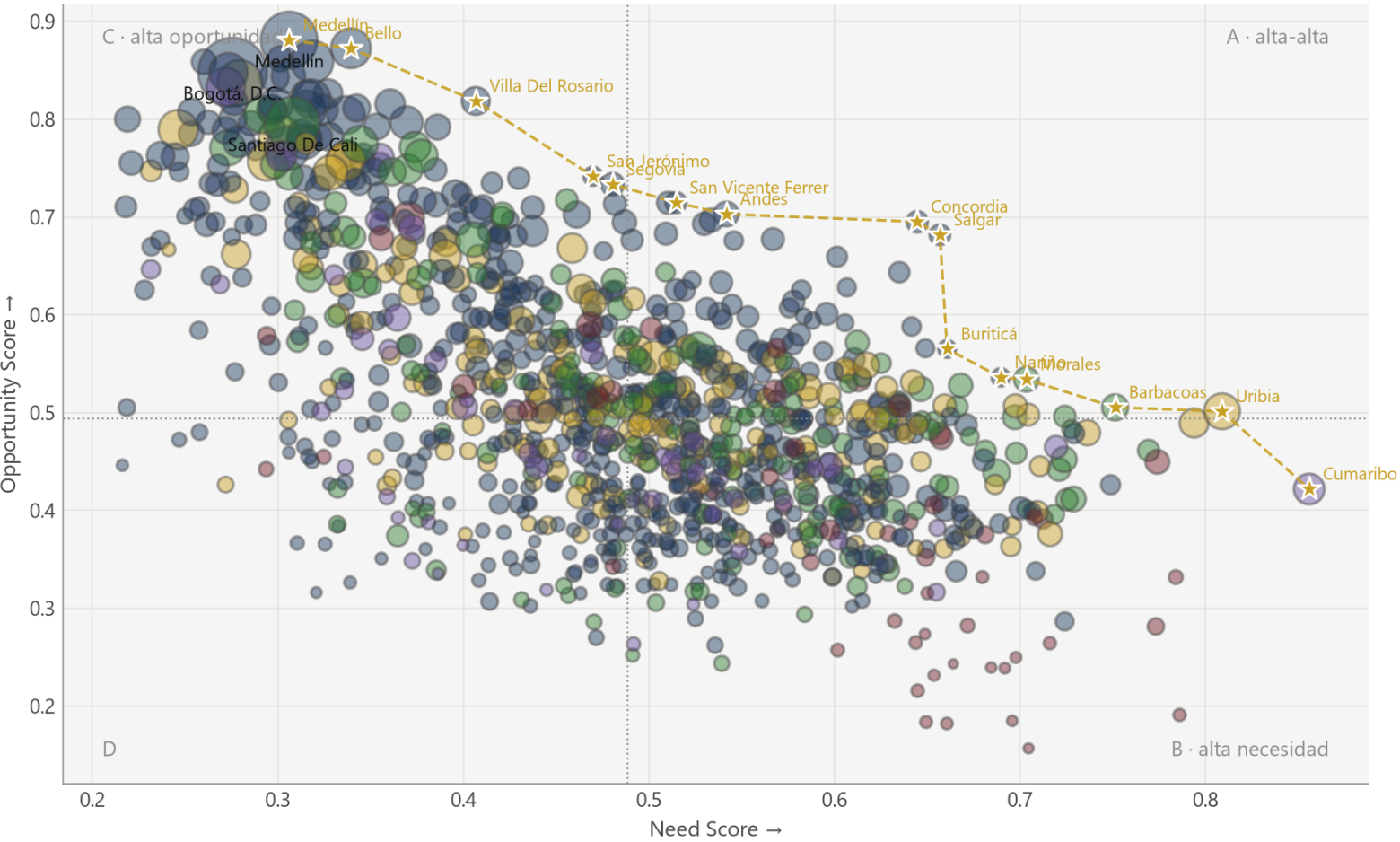
Necesidad → Oportunidad → Confianza



El 2×2 que ordena la decisión

NECESIDAD × OPORTUNIDAD: EL 2×2 QUE DECIDE

Cuadrante A (177 municipios) reúne 13,1 % del target; 15 municipios forman la frontera de Pareto (no dominados en ambas dimensiones)



Fuente: Índices compuestos propios sobre CNPV 2018, MEN 2024, MinTIC, EDUC 2023 (config/index_weights.yaml) · Periodo: 2018-2025

INFERIDO [G]

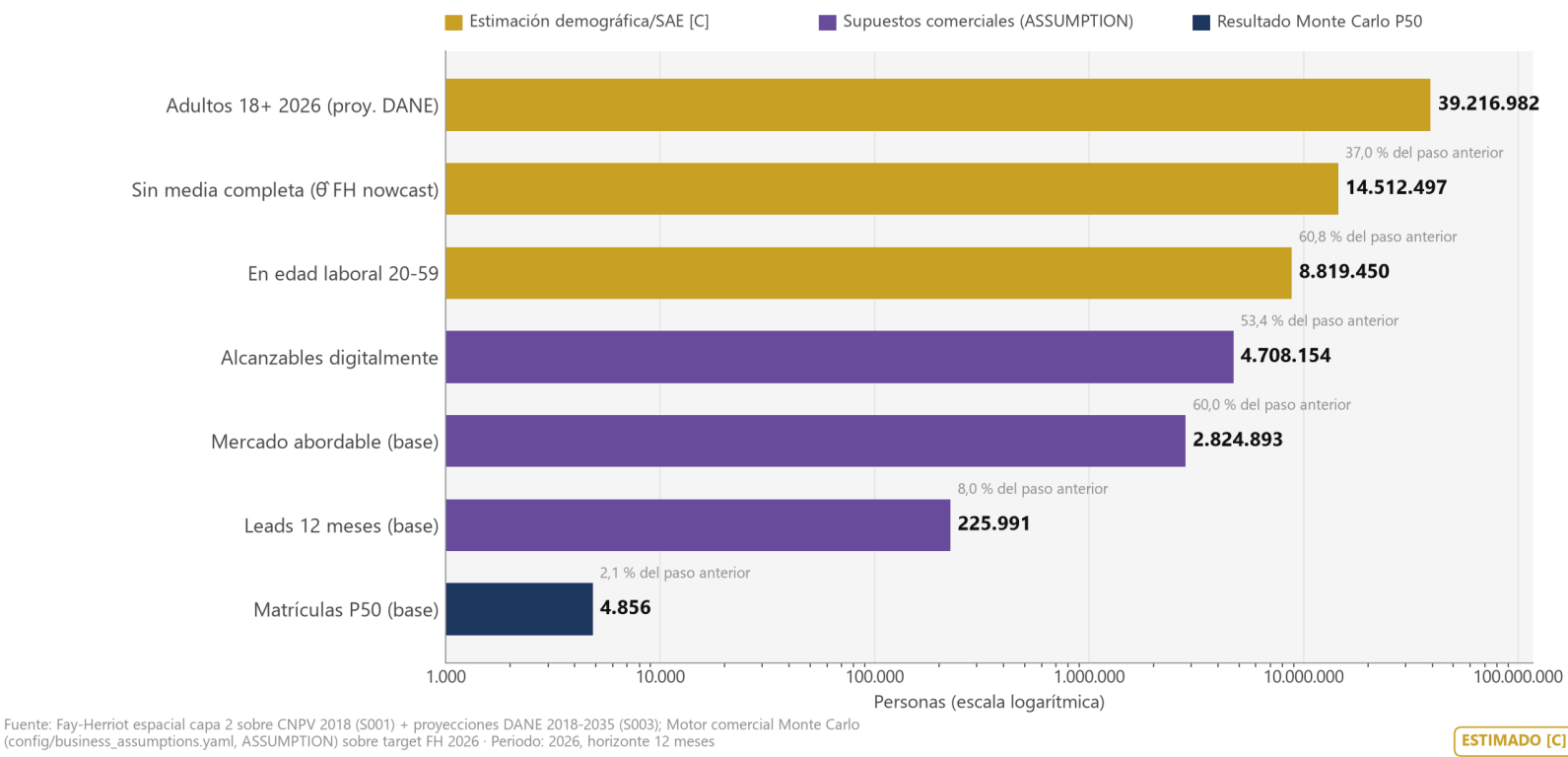
Necesidad × Oportunidad: el 2×2 que decide. Cuadrante A: 177 municipios; frontera de Pareto: 15

INFERIDO ['S001', 'S012', 'S026', 'S080']

Money Map: del déficit a la matrícula, eslabón por eslabón

MONEY MAP NACIONAL

De 39,2 M adultos, 14,5 M no terminaron la media; el embudo comercial base convierte 225.991 leads en 4.856 matrículas P50 en 12 meses

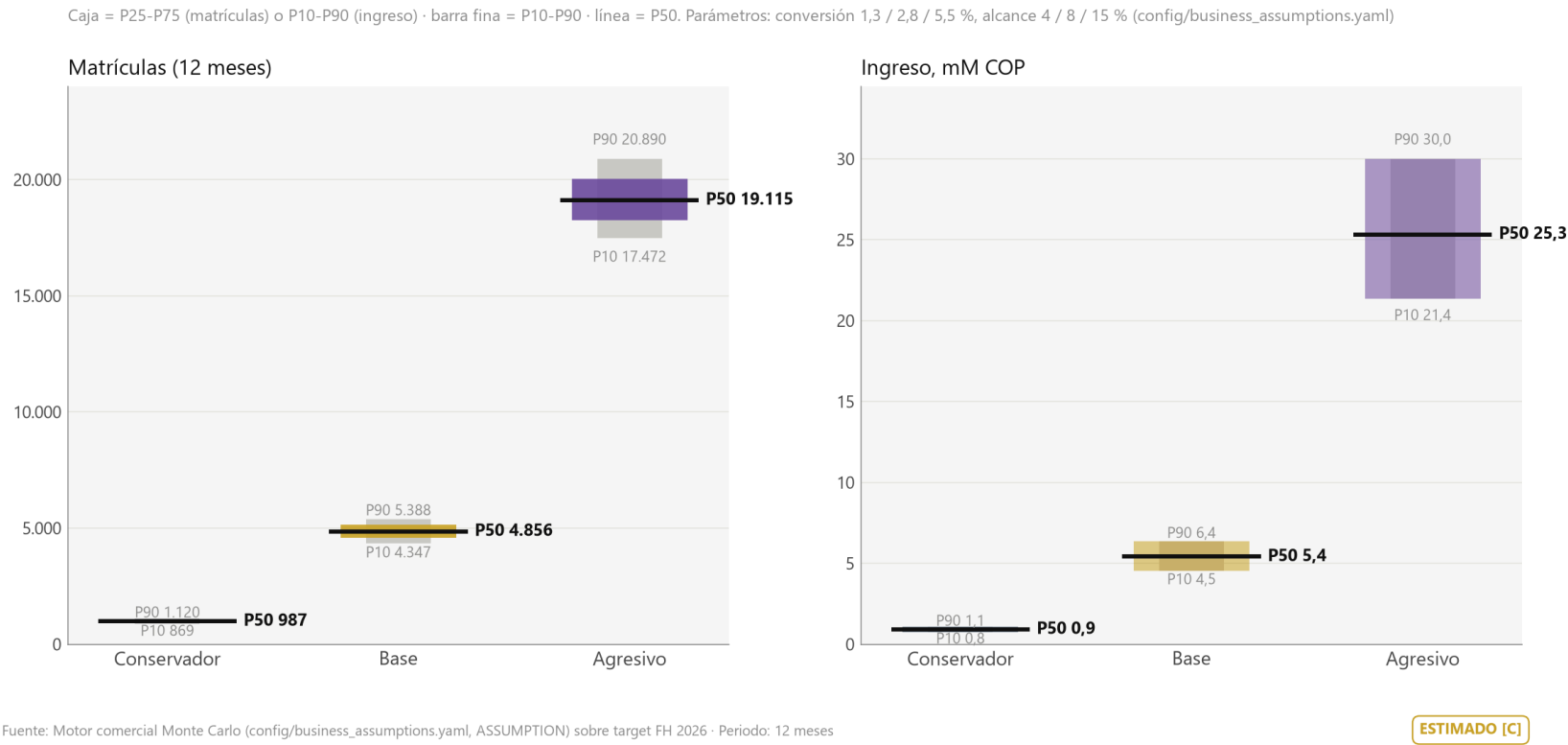


Escenario base P50: 4.856 matrículas/año · 5,4 mil M COP · P10–P90 4,5 mil M COP–6,4 mil M COP. Supuestos en business_assumptions.yaml.

ESTIMADO ['S001', 'S003', 'capa2_fh', 'business_assumptions']

ESCENARIOS MONTE CARLO (10.000 SIMULACIONES)

Escenario base: 4.856 matrículas y 5,4 mM COP P50 en 12 meses (P10-P90: 4,5 mM COP – 6,4 mM COP); probabilidad de superar 5 mM: 71 %



Escenarios Monte Carlo (10.000 simulaciones). Base P50: 4.856 matrículas, 5,4 mM COP; agresivo P50 25,3 mM COP

ESTIMADO ['business_assumptions', 'capa2_fh']

Top-20: dónde entrar primero

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

El canal se ajusta al territorio: ChannelFitScore

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

field: 205 municipios whatsapp: 162 municipios government: 157 municipios tiktok: 127 municipios education_partners: 108 municipios referral: 98 municipios employer: 91 municipios google_search: 84 municipios call_center: 51 municipios meta_ads: 39 municipios

El score dice dónde encaja un canal, no cuánto convierte. Las tasas por canal serán experimentales.

Tres motores: B2C, empleadores, institucional

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

D2C · TAM: 14.512.497 (DERIVED) D2C · SAM: 2.824.893 (ASSUMPTION) D2C · SOM: 4.864 (ASSUMPTION) Employer · TAM: 3.448.191 (DERIVED) Employer · SAM: 559.337 (DERIVED) Employer · SOM: None (NOT_AVAILABLE) Institutional · TAM: 8.607.587 (DERIVED) Institutional · SAM: 7.137.560 (DERIVED) Institutional · SOM: None (NOT_AVAILABLE)

Por qué esto hace dinero

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

DATA → menos incertidumbre de mercado → mejor selección de territorio → mejor ajuste de canal → menos adquisición desperdiciada → mayor densidad de leads calificados → mejor unit economics. Cada flecha se apoya en una métrica existente: DataConfidence, NeedScore/OpportunityScore, ChannelFitScore, asignación LP, intención social. Lenguaje: permite priorizar, permite optimizar, reduce incertidumbre. No: 'reducirá el CAC un 37 %'.

Cómo se construyó (una sola lámina)

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

81 fuentes · 744.782 observaciones con linaje · 182 variables por municipio · 12.723 conversaciones · 4.519 documentos web. Censo 2018 → split GEIH 2025 → Fay-Herriot espacial → nowcast 2026. Moran/LISA/Gi*, H3, OSM, gravedad, Louvain. K-Means, LightGBM CV espacial, SHAP. e5 + BERTopic + zero-shot. Monte Carlo 10.000 × 3 escenarios. LP. Reproducible: make all. Auditable: observations en DuckDB.

No buscamos personas al azar

HYPERION ONE · corte 2026-09-08 · BORRADOR_IA

Sabemos dónde está la necesidad. Sabemos qué la explica. Sabemos dónde existe oportunidad. Y ahora podemos decidir dónde invertir.

DATA → TERRITORY → PEOPLE → OPPORTUNITY → REVENUE